NOAA Technical Report ERL 436-FSL 2

DAR³E-I Evaluation: An Overview

Kenneth F. Heideman
Denice C. Walker
John A. Flueck

March 1989

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Environmental Research Laboratories

NOAA Technical Report ERL 436–FSL 2

DAR³E–I Evaluation: An Overview

Kenneth F. Heideman
Denice C. Walker
John A. Flueck

Forecast Systems Laboratory
Program for Regional Observing and Forecasting Services
Boulder, Colorado

March 1989

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary

National Oceanic and Atmospheric Administration
William E. Evans, Under Secretary for Oceans and Atmosphere/Administrator

Environmental Research Laboratories
Boulder, Colorado
Joseph O. Fletcher, Director

Contents

DAR^{3}E-I EVALUATION: AN OVERVIEW

Kenneth F. Heideman*
Denice C. Walker
John A. Flueck*

ABSTRACT. The unique DAR^{3}E-I (Denver AWIPS-90 Risk Reduction and Requirements Evaluation, Part I) system combines an advanced, interactive workstation with special meteorological data sets and provides the capability to demonstrate and test many of the concepts central to the AWIPS-90 (Advanced Weather Interactive Processing System for the 1990's) component of the planned modernization of the National Weather Service (NWS). This overview of six evaluation reports focuses on the following areas: (1) forecaster assessment of the system, (2) forecaster product usage, and (3) forecaster performance as measured by forecasts and warnings. Forecasters agreed that the integration of standard data sets with new (primarily mesoscale) data sets into a single workstation, and the capability to manipulate the data in ways not previously possible, make DAR^{3}E-I a substantial improvement over the current AFOS (Automation of Field Operations and Services) system. Forecasters also identified some problem areas and they are being addressed in DARE^{3}E-II. Product usage analysis shows that DAR^{3}E-I provides sufficient flexibility in the daily product inventory to allow forecasters to focus on the contrasting forecasting problems presented on the synoptic and subsynoptic scales in both the cool and warm seasons. Doppler radar imagery is a key component of the mesoscale product set available on the DAR^{3}E-I system. On the basis of the data available, assessment of tornado warnings revealed a number of substantial improvements following installation of DAR^{3}E-I, including increased lead time, and decreased size of area warned and false alarm ratio (FAR). Similar, though smaller, improvements were observed with regard to severe thunderstorm warnings. In contrast, the probability-of-detection (POD) scores declined for both types of warnings over the same period. There is a suggestion of improvement in cool and warm season 0-12 hour temperature forecasts whereas precipitation forecasts after the installation of DARE^{3}E-I do not show any notable changes in skill or reliability.

1. INTRODUCTION

In late 1986, the Program for Regional Observing and Forecasting Services (PROFS) installed an advanced, interactive workstation with special meteorological data sets at the Denver Weather Service Forecast Office (WSFO). This unique system, referred to as DAR^{3}E-I (Denver AWIPS-90 Risk Reduction and Requirements Evaluation, Part I), provides the capability to demonstrate and test many of the concepts central to the AWIPS-90 (Advanced Weather Interactive Processing System for the 1990's) component of the planned modernization of the National Weather Service (NWS) forecasting operations.

*Also affiliated with the Cooperative Institute for Research in Environmental Sciences.

As part of the full DAR³E effort, PROFS has been conducting an ongoing evaluation of the effects of the DAR³E-I system on forecast operations at the Denver WSFO. To date, this evaluation has focused on three areas: (1) forecaster assessment of the system, (2) forecaster product usage, and (3) forecaster performance as measured by accuracy of forecasts and warnings. The PROFS Evaluation Group has completed six major reports detailing its work and findings in each of these areas (Heideman, 1989a and b; McCoy and McGinley, 1989; McCoy and Kent, 1989; Walker 1988, 1989). The following is an overview of these six reports. For full details and insights into this work, we strongly recommend that the reader consult these reports.

2. REVIEW OF PRINCIPAL FINDINGS

2.1. Forecaster Assessment

The primary goal of the DAR³E-I forecaster surveys has been to provide the National Weather Service, and in particular those responsible for establishing AWIPS-90 system requirements, with information on the strengths and weaknesses of the DAR³E-I system as experienced by the Denver WSFO forecasters. To date, the results from two questionnaires submitted to the forecasters have been analyzed (Walker, 1988, 1989). The first questionnaire, given to the forecasters in late February 1987, covered the cool-season months of December 1986 through February 1987 and focused on the forecasters' satisfaction with the DAR³E-I system. The second questionnaire was given to the forecasters in March 1988, and contained two sections. The first dealt with the 1987 warm season (approximately May through August); the second gave the forecasters an opportunity to provide feedback on the system after approximately 15 months of operational use (December 1986 through March 1988). The focus of the second questionnaire was on the forecasters' assessment of the effect of the DAR³E-I system on their ability to perform their jobs. Twelve of 15 forecasters completed the first questionnaire; 12 of 14 completed the second questionnaire.[1]

The following is a discussion of the most significant findings from the first two questionnaires and a brief description of the actions taken to address problems (DAR³E-I) or of planned upgrades (DAR³E-II).

The forecasters agreed that the DAR³E-I system is a substantial improvement over AFOS:

---

[1] A third questionnaire was given to the forecasters in November 1988. It sought the forecasters' assessment of the effect of the DAR³E-I system on their jobs during the 1988 warm season, including their assessments of several new products and application programs added to the system just prior to the beginning of the warm season. Responses to that questionnaire are being compiled and are not included in this overview.

The DARE³E-I system provides forecasters with a view of the atmosphere not previously possible by integrating essential meteorological data sets into a single workstation and combining them with the system's overlay, looping, automatic update, and color-enhancement capabilities.

The forecasters left no doubt that what contributes most to making the DAR³E-I system a substantial and significant improvement over AFOS is the fact that it takes the standard meteorological data sets, adds new (primarily mesoscale) data sets, integrates them into a single workstation, and provides the ability to manipulate the data in ways not previously possible. The abilities to display and animate (loop) satellite or radar images, add color-coded graphic overlays, color-enhance satellite imagery, and animate graphics without imagery, substantially aid the forecasters in extracting significantly more information from the data available to them. Further, this ability to manipulate the data increases the utility of all products, particularly those the forecasters consider to be among the most useful (Walker, 1988, 1989): satellite and Doppler radar imagery, mesoscale products such as mesonet, profiler, and Mesoscale Analysis and Prediction System (MAPS) graphics, and the family graphics products. (Family graphics are sets of analysis and forecast products that are grouped together according to the model that produced them. They can be selected for display as if they were a single product, but once they are selected, the forecaster can display a variety of combinations of family members.) Additionally, integrating these data sets and capabilities into a single workstation eliminates the forecaster's need to monitor several disparate displays, while attempting to mentally integrate the data from one display into a meaningful context for the next. With the DAR³E-I system, the forecasters are able to assimilate more data in less time than was previously possible, and thus have more time to avail themselves of the full range of data (see Tables 1-3).

Table 1. Distribution of the twelve forecasters' ratings of the DAR³E-I system

Item	Number of Forecasters Giving Rating							
	Inadequate (1)	Inadequate (2)	Marginal (3)	Adequate (4)	Effective (5)	Effective (6)	No Ans/ Other	Average Rating
Response time			3	5	4			4.1
Ease of use				2	7	2	1	5.0
Reliability				1	9	1	1	5.0
Utility					8	3	1	5.3

Table 2. Distribution of the twelve forecasters' ratings of the DAR^{3}E-I overlay capabilities

Item	Inadequate (1)	(2)	Marginal (3)	Adequate (4)	Effective (5)	(6)	No Ans/ Other	Average Rating
	Number of Forecasters Giving Rating							
Response time			4	4	4			4.0
Ease of use				5	3	4		4.9
Reliability				5	4	2	1	5.2
Utility				2	6	4		5.2
Quality of over-laid graphics				5	4	1	2	4.6

Table 3. Distribution of the twelve forecasters' ratings of the DAR^{3}E-I looping capability

Item	Inadequate (1)	(2)	Marginal (3)	Adequate (4)	Effective (5)	(6)	No Ans/ Other	Average Rating
	Number of Forecasters Giving Rating							
Response time		2	2	3	3	2		4.1
Ease of use				1	7	4		5.3
Reliability			1	2	7	2		4.8
Utility					8	4		5.3

The warning generation program (1) makes optimal use of DAR^{3}E-I's integrated data sets and data manipulation capabilities by enabling the forecasters to interact directly with the data during the process of issuing a warning, (2) speeds the process by automating much of it, and (3) reduces mistakes by requiring less typing.

The warning generation program on DAR^{3}E-I was rated very highly by the forecasters (see Table 4). The program uses either a radar or satellite image (or images in a loop) displayed on either of the two display screens. The images(s) may also contain one or more overlays of additional data the forecaster considers necessary in deciding whether or not, or when, to issue a warning. The warning generation program then enables the forecaster to

Table 4. Distribution of the twelve forecasters' ratings of the DAR^{3}E-I warning generation program

Item	Number of Forecasters Giving Rating: Inadequate (1)	Inadequate (2)	Marginal (3)	Adequate (4)	Effective (5)	Effective (6)	No Ans/ Other	Average Rating
Response time			1	2	5	4		5.0
Ease of use				3	4	5		5.2
Reliability				5	5	2		4.8
Utility				1	7	3	1	5.2
Compared with other methods				1	5	5	1	5.4

identify the storm for which the warning is being issued and to outline the threatened area directly on the display screen; as a result, warning areas are more precisely defined than was previously possible. The program automatically lists the counties, cities, and towns that are included in the outlined area. In contrast, the AFOS system has no imagery and does not provide forecasters with the <u>integrated</u> data sets necessary to assess, quickly and accurately, the appropriateness of a warning. Neither does AFOS provide the capability to describe the warning area graphically (and therefore more precisely).

The Denver forecasters also agreed that some aspects of the DAR^{3}E-I system detract from their ability to do their jobs:

The time the DAR^{3}E-I system takes to acknowledge the forecaster's input and to complete the command is much too long.

The forecasters repeatedly indicated that the overall utility of the DAR^{3}E-I system is degraded by the slow system response time (Table 1). On the basis of limited trials conducted at PROFS, the time to load an eight-frame satellite loop ranges from 45 to 80 seconds. The time to load a model family graphic runs from 21 to 44 seconds. Thus the forecasters are reluctant to clear an eight-frame satellite loop to look at additional, though perhaps not critical, data because of the time it will take to re-display the original satellite loop. The result is less than maximum utilization of the full range of products available on the DAR^{3}E-I system, <u>especially</u> in time-critical situations. Furthermore, the slow system response time may also be the primary reason most forecasters indicate they would prefer three display screens--one each for displaying a continuous satellite and radar loop, and the third for displaying all additional data.

Design changes planned for the DAR^{3}E-II system are expected to result in nearly instantaneous acknowledgment of forecaster input and equally fast completion of the command. Other planned features of the DAR^{3}E-II system that will make it faster include two independent display screens, a "smart" mouse, and "tear-away" menus. The two independent displays will increase the speed of the system by making the activity on one screen completely independent of the other. The smart mouse will allow the forecasters to have a variety of functions at their finger tips, eliminating the need to move through one or more menus to activate the desired function. Similarly, the tear-away menus will allow the forecasters to leave a portion of a larger menu continuously displayed on the screen, eliminating the need to move through the menu hierarchy to retrieve the desired sub-menu.

The DAR^{3}E-I VAXstation text interface lacks standard word processing features.

Persistent forecaster complaints about the lack of word processing features in the VAXstation text interface highlight the fact that the DAR^{3}E-I system offers them a text editor, when what they want is a word processor (Table 5). While most forecasters agree that the text editor on the DAR^{3}E-I system is substantially better than that on AFOS, they expect a state-of-the-art system such as DAR^{3}E-I to provide basic word processing features. The VAXstation text interface of the DAR^{3}E-II system will provide forecasters with basic word processing capabilities. It is worth noting that the text editor is only one of several capabilities and features that collectively form the text interface. Others include the SAO processing and handling capability, the alarm/alert feature, the browser (a series of layered menus that allow the user to retrieve products without knowing product identifiers or typing an AFOS command line), and the multiple-windows feature. Taken in its entirety, the text interface is viewed positively by the forecasters.

Table 5. Distribution of the twelve forecasters' ratings of the DAR^{3}E-I text editor

	Number of Forecasters Giving Rating							
	Inadequate		Marginal	Adequate	Effective		No Ans/	Average
Item	(1)	(2)	(3)	(4)	(5)	(6)	Other	Rating
Design		1	2	5	4			4.0
Response time			1	3	6	2		4.8
Ease of use			3	5	3	1		4.2
Reliability				5	6	1		4.7
Utility			2	1	8		1	4.6
Compared with others		1	1	4	6			4.3

On the DAR^{3}E-I system, pixel replication causes lines, letters, and numbers on a zoomed graphic to enlarge, making some graphics difficult to read once they are zoomed beyond a few magnification factors.

The forecasters' most consistent complaint about the zoom feature concerns the poor quality (readability) of zoomed graphics. This is reflected in Table 6, which shows that quality of zoomed products received the lowest rating of the five rated items. There is also no de-clutter feature to provide additional data as the zoom factor increases. Hardware limitations require that the zoom feature on the DAR^{3}E-I system be implemented using pixel replication. These limitations will not change for DAR^{3}E-II; thus, the zoom feature will continue to use pixel replication. However, the DAR^{3}E-II system will also provide a redraw capability, which will improve the readability, and hence utility, of zoomed graphics by redrawing lines, numbers, and text of a zoomed product in their original line weights and character sizes. The DAR^{3}E-II system's redraw feature will also provide de-clutter (which presents more data with increasing zoom), enhancing its overall utility.

Table 6. Distribution of the twelve forecasters' ratings of the DAR^{3}E-I zoom feature

Item	Number of Forecasters Giving Rating: Inadequate (1)	Inadequate (2)	Marginal (3)	Adequate (4)	Effective (5)	Effective (6)	No Ans/ Other	Average Rating
Response time				3	6	3		5.0
Ease of use				2	5	5		5.3
Reliability				3	5	4		5.1
Utility			1	4	5	2		4.7
Quality of zoomed product		1	2	2	4	1	2	4.2

The DAR^{3}E-I printers are only marginally adequate for operational use because (1) the laser printer does not make copies large enough for hand analysis, and (2) the color printer is much too slow for operational use.

Two printers accompany the DAR^{3}E-I system: a laser printer that makes black and white copies, and a color copier. These printers meet current AWIPS-90 specifications and cannot be upgraded without a change or waiver of the specifications. Nonetheless, they are barely adequate for operational use. (The color printer has been disconnected from the system and is not being used at all.) The forecasters consistently indicated that they need laser-printer copies of graphic products that are large enough for hand analysis. They also need a color printer that is fast enough to be useful in a real-time environment and one that makes good-quality copies of gray-shaded images (e.g., infrared or visible satellite) (see Table 7). In addition, they must be able to make several print requests at one time (queuing), a capability not provided by the DAR^3E-I system but planned for DAR^3E-II.

Table 7. Distribution of the twelve forecasters' ratings of the DAR^3E-I system's hard copy devices

Item	Number of Forecasters Giving Rating: Inadequate (1)	Inadequate (2)	Marginal (3)	Adequate (4)	Effective (5)	Effective (6)	No Ans/ Other	Average Rating
Response time	1		5	2	2	1	1	3.6
Ease of use			1	4	3	3		4.3
Reliability			1	6	2	2	1	4.5
Utility			3	4	4	1		4.3
Quality of copy		1	1	5	3	2		4.3

NEXRAD (Next Generation Weather Radar) base products—Doppler reflectivity and velocity images—are especially useful. However, the NEXRAD storm algorithm products are of little use in helping forecasters to identify or anticipate severe weather phenomena.

Prior to the 1987 warm season, NEXRAD-simulated Doppler reflectivity and velocity images (the so-called NEXRAD base products) were installed on the DAR^3E-I system. Also installed were the NEXRAD Storm Sequence algorithms and the NEXRAD Mesocyclone Detection algorithm. The resolution of the reflectivity and velocity data received from the NCAR (National Center for Atmospheric Research) CP-2 Doppler radar was degraded to match, as nearly as possible, the specified data resolution of NEXRAD. Similarly, the meteorological structure and integrity of the NEXRAD algorithms, as specified by NEXRAD, were retained in the process of fitting them into a real-time environment, with the exception of the mesocyclone algorithm. Thus, while the reader can be confident that the forecasters' assessments of the base products and the Storm Sequence algorithms are based on an accurate reproduction of the NEXRAD design, the same cannot be said of their assessments of the Mesocyclone Detection algorithm, which was discovered to have been inadvertently altered as the result of coding errors. Therefore, the following does not include discussion of the Mesocyclone Detection algorithm.

The forecasters commented very favorably on the utility of most products available on the DAR³E-I system. They are particularly enthusiastic about the Doppler reflectivity and velocity products (Table 8). The forecasters most commonly stated that reflectivity and velocity imagery improves their ability to anticipate convective initiation and to make improved short-term forecasts. They also indicated that these products substantially aid them in monitoring storm development, structure, and movement.

Table 8. Distribution of the twelve forecasters' ratings of the NEXRAD base products: Doppler reflectivity and velocity images

	Number of Forecasters Giving Rating							
	Inadequate		Marginal	Adequate	Effective		No Ans/	Average
Item	(1)	(2)	(3)	(4)	(5)	(6)	Other	Rating
Quality				2	4	5	1	5.3
Quantity		1		2	4	5		5.0
Reliability				2	6	4		5.2
Timeliness				2	4	6		5.3
Update frequency				2	3	7		5.4
Utility				1	3	7	1	5.6

In contrast, however, the forecasters indicated that the NEXRAD storm algorithm products do little to improve their ability to identify or anticipate severe hail, storm structure, or storm movement. The forecasters stated that (1) the products do not provide the guidance they need most; for example, they need information on hail size, but only get information on the likelihood of hail occurrence, (2) they do not trust the accuracy of the information provided by the products, and (3) there is not enough time to use the algorithms during active situations. In addition, most forecasters are reluctant to rely solely on the algorithm output; they also want to see the supporting reflectivity and velocity data. This is due to the forecasters' need to confirm the accuracy of the algorithm output and to their need to see, directly, the current state of the atmosphere.

One possible explanation for the lack of accuracy of the storm algorithms is that their design has been based largely on structure and behavior characteristic of Oklahoma-type severe thunderstorms. Thus the algorithms are not optimally suited to identifying and depicting severe weather phenomena occurring in Colorado storms.

Several NEXRAD hydrology algorithms were introduced in 1988: plots of 1-hour and 3-hour precipitation accumulations, and a storm-total precipitation image. Preliminary results from the 1988 questionnaires indicate that the forecasters continue to use the NEXRAD storm algorithm products infrequently. However, they are using the precipitation products much more regularly and find them to be helpful.

The forecasters also requested, and in some cases have already received new products and capabilities.

The DAR³E-I system now has the capability to overlay up to four skew-T products and supplies additional CP-2 Doppler radar imagery; these features provide the forecasters easy access to information about the vertical structure of convective storms. The additional CP-2 products given to the forecasters for the 1988 warm season included the plan position indicator (PPI) "packed Z/V" image (reflectivity and velocity combined into one image), composite PPI Z/V at Ø.75 km and 6.4Ø km altitude, the CP-2 cross section application program, and the WER (weak echo region) application program.

In both questionnaires the forecasters asked for the ability to overlay at least two images. Most often the request was to overlay radar and satellite, but several requests were for radar over a road-map image. To date, this capability has been available only by invoking a rather slow and cumbersome application program that allows the forecasters to combine two images. The DAR³E-II system will provide an enhanced combining capability that will be accessible by means of the mouse, allowing the forecasters to combine two images quickly and easily.

The forecasters also indicated that they would like to be able to specify the composition of family graphics. On DAR³E-I the composition of family graphics is a fixed set of eight graphics per family. DAR³E-II will provide forecasters with a limited ability to specify the composition of the family graphics. More importantly, however, DAR³E-II will provide a greatly enhanced graphic-product selection matrix, enabling the forecasters to easily select related groups of graphic products directly from the matrix.

Additional enhancements planned for DAR³E-II include the ability to simultaneously run three application programs per display, quickly re-display a previous screen (i.e., display) configuration, and customize the state of the display. The number of frames in a loop will be user-specified, up to a maximum of 32 frames, and there will be new WFO (Warning and Forecast Office) and sub-WFO product scales, a relational data base, extensive event and product monitoring, and user-defined procedures. Perhaps one of the most significant enhancements will be the new grid-to-graph facility, which will provide the forecasters greater flexibility in accessing and displaying NMC (National Meteorological Center) gridded data. These additional capabilities and features have not been planned in direct response to specific criticisms or requests of the forecasters. Rather, they represent essential upgrades and, in many cases, are necessary to meet AWIPS-9Ø specifications that were not met with the DAR³E-I system.

2.2. Product Usage

Analysis of the frequency with which products are requested provides information on how forecasters use the DAR³E-I system. With more than 1000 products from which to choose, forecasters constantly make decisions about which products provide information most germane to the particular weather situation or forecast problem at hand. Their choices can reveal a great deal about how system resources are utilized under various meteorological conditions, although the reader is cautioned that the number of requests for a given product is not necessarily proportional to the usefulness of the product. This information can also identify products that are being used less often than expected. The process of determining possible reasons for the lack of use can lead to corrective measures to improve the product and its configuration, or perhaps to improve forecasters' understanding of the product.

Several limitations arose from constraints on the software used in the counting process. The name, scale, and time of each product request was logged, but it was not possible to determine the number of times application programs were actually run. Another limitation involved the automatic update feature that was used extensively with radar and satellite imagery. Only the initial request for the product could be recorded. Subsequent automatic updates could not be included in the number of requests. In addition, there was no way to count the number of requests for individual products when forecasters were viewing family graphics. This resulted in significant undercounting of the use of many AFOS graphics products. The reader should bear these limitations in mind when reviewing the numerical results.

To date, product usage has been analyzed on a seasonal basis (1 December 1986 to 28 February 1987, representing the cool season, and 1 June to 31 August 1987, representing the warm season) and in association with specific snow and convective events. The results provide a baseline for future studies, and some of the findings have already motivated changes in current or projected DAR³E-II system capabilities. An analysis of forecaster product usage on the DAR³E-I workstation during the warm season of 1988 is in progress.

The following information highlighting forecasters' use of the DAR³E-I system comes from analysis of the workstation product inventory and of the forecasters' product usage.

Product inventory shows that the DAR³E-I system provides products, over and above AFOS, that enable forecasters to monitor mesoscale processes effectively. Integration of Doppler radar imagery with workstation capabilities makes up a large part of this mesoscale resource.

Table 9 incorporates the number of products in a given category with the update frequencies (the frequency with which new data become available) for these products. For example, most numerical model graphics are updated every 12 hours (twice per day), and satellite imagery on the state scale is normally updated every 30 minutes (48 times per day). Although more numerical model products than satellite images are available on the DAR³E-I system, the update

Table 9. Daily inventory of products available on DAR^3E-I as of early 1988

Product Scale	Product Type									Total
	Models	Satellite	Radar	Surface	Upper Air	Skew-T/ Profiler	Back-grounds	Appli-cation	Misc.	
Local	0	128 (1.0)**	4896 (37.8)	384 (2.9)	0	0	19 (0.2)	9 (0.1)	579 (4.4)	6015 (46.4)
State	0	228 (1.8)**	2736 (21.1)	168 (1.3)	0	0	19 (0.2)	8 (0.1)	1 (0.0)	3160 (24.4)
Regional	0*	672 (5.2)	480 (3.7)	24 (0.2)	0	0	11 (0.1)	8 (0.1)	0	1195 (9.2)
West 2/3	0	0	0	24 (0.2)	0	0	3 (0.0)	0	0	27 (0.2)
National	848* (6.5)	72 (0.6)	24 (0.2)	24 (0.2)	192 (1.5)	0	5 (0.0)	1 (0.0)	105 (0.8)	1271 (9.7)
N. Amer.	408* (3.1)	32 (0.2)	0	14 (0.1)	192 (1.5)	0	12 (0.1)	0	96 (0.7)	754 (5.8)
N. Hemis	224* (1.7)	8 (0.1)	0	0	0	0	3 (0.0)	1 (0.0)	0	236 (1.8)
Vertical	0	0	0	0	0	208 (1.6)	2 (0.0)	8 (0.1)	0	218 (1.7)
Other	0	0	1 (0.0)	0	0	0	74 (0.6)	6 (0.0)	5 (0.0)	86 (0.7)
TOTAL	1480 (11.4)	1140 (8.8)	8137 (62.8)	638 (4.9)	384 (2.9)	208 (1.6)	148 (1.1)	41 (0.3)	786 (6.0)	12962 (100.0)

Note: Numbers in each category represent the number of individual products in the category multiplied by the number of times each product is updated with new data in a 24-hour period. Percent of total product inventory is given in parentheses. There is no update frequency for Backgrounds or Application programs, so multiplication factor is 1. Update frequency for local radar varies between 3 and 6 minutes and was arbitrarily assigned at 5 minutes for the purposes of this table. Rows and columns may not sum to 100% because of round-off error.

*AFOS Graphics may be displayed on all of these scales, but this table accounts only for the scale on which the product is stored on the DAR^3E-I workstation. AFOS Graphics are not available on the state or local scales.

**Numbers for local and state satellite imagery include 8 hours of rapid-scan capability, when new images are available every 5 minutes instead of the usual interval of 30 minutes. The decision to implement rapid-scan is that of the National Severe Storm Forecast Center in Kansas City, Missouri.

frequency becomes an overriding factor. The update frequency is generally a function of the scale of a product; conditions on the mesoscale change more rapidly than those on the synoptic scale and thus necessitate more frequent updates. Thus, product types and their associated update frequencies cannot be separated when the range of data available on the DAR³E system is assessed. The inventory shows that the DAR³E-I system provides every product available on AFOS and adds a considerable mesoscale dimension, both spatial and temporal, to the tools available to the forecaster.

Product usage analysis revealed disproportionately greater use of reflectivity over velocity imagery during the 1987 warm season and resulted in introduction of the "packed Z/V" radar product just prior to the summer of 1988.

Data on radar product usage during the warm season of 1987 are presented in Table 10. The most important finding is the imbalance in requests for PPI reflectivity over velocity. The ratio of reflectivity requests to velocity requests was about 5 to 1 on the state scale and almost 2 1/2 to 1 on the local scale. Since forecasters do not dispute the need for velocity data (see Walker, 1988, 1989), a possible reason for the disparity is the forecasters' lack of experience with Doppler velocity data. (Velocity products have been available to forecasters only since the summer of 1987, but Limon reflectivity imagery has been available at the Denver WSFO for many years.) Since velocity imagery is arguably more difficult to interpret than reflectivity data, supplemental forecaster training may be needed in this area. Finally, it may be that physical limitations of the system have contributed to the limited use of velocity imagery. For example, in convectively active situations forecasters often dedicate one of the two display screens to radar imagery. Most frequently they choose to have four or eight frames of automatically updating images of reflectivity on that screen. Because of the slow system response, too much time would be lost if the forecaster replaced reflectivity with velocity and then wanted to reload reflectivity on the same display. Thus forecasters may be reluctant to load velocity images, so the product is used less.

This problem was addressed with the creation of the "packed Z/V" radar products, which display both reflectivity and velocity on the same image--one or the other appears on the screen depending on the position of a toggle switch. The packed product makes it easier for forecasters to get a sense of the location of velocity features relative to reflectivity features. The first summer during which the packed Z/V radar products were available was in 1988. Table 11 quantifies the use of most radar products during that summer. The overwhelming preference for the new Z/V images is obvious: more than 90% of all requests for PPI radar data were for the combined products. Although there was no way to determine the exact proportion of time forecasters spent looking at reflectivity versus velocity when they used the combined product, preliminary results from the third questionnaire show that forecasters estimated an average use of roughly 30% for velocity imagery. This represents a modest increase in use of velocity products over the previous summer when packed Z/V imagery was not available.

Table 10. Use of Doppler radar reflectivity and velocity imagery on the state and local scale during the warm season of 1987 on the DAR³E-I workstation

Product	Elev. Angle (degrees)	Number	Percent
LOCAL SCALE			
Reflectivity	0.2	554	9.4
	0.7	1129	19.1
	4.0	632	10.7
Velocity	0.2	431	7.3
	0.7	482	8.1
	4.0	84	1.4
STATE SCALE			
Reflectivity	0.2	225	3.8
	0.7	443	7.5
	1.8	189	3.2
	2.9	119	2.0
Velocity	0.2	101	1.7
	0.7	91	1.5
	1.8	19	0.3
	2.9	16	0.3

Note: Use is given by number of requests and percent of all radar requests, which includes radar products not included in this table, such as WSR-57 imagery from Limon, CO, and Cheyenne, WY, and NEXRAD algorithms. Total radar use is given in Table 12b.

It appears that forecasters requested only separate reflectivity or velocity images (instead of the packed image) in order to retain maximum horizontal and depth resolution for 8, 16, or 32 frame loops. The loss of horizontal and/or depth resolution for packed Z/V images longer than 8 frames will not be a problem on DAR³E-II. Therefore, separate reflectivity and velocity images will not be needed.

Table 11. Use of Doppler radar reflectivity and velocity products and Doppler packed Z/V products during the warm season of 1988 on the DAR³E-I workstation

	Product	Elev. Angle (degrees)	Number of Requests	Percent of All Requests*
LOCAL SCALE				
	Reflectivity	0.2	1	0.0
		0.7	59	1.4
		4.0	29	0.7
	Velocity	0.2	1	0.0
		0.7	19	0.5
		4.0	2	0.1
	Packed Z/V	0.7	1282	29.3
		1.8	388	8.9
		4.0	615	14.1
		6.2	30	0.7
STATE SCALE				
	Reflectivity	0.2	30	0.6
		0.7	22	0.5
		1.8	21	0.5
		2.9	15	0.3
	Velocity	0.2	1	0.0
		0.7	1	0.0
		1.8	0	0.0
		2.9	0	0.0
	Packed Z/V	0.2	425	9.7
		0.7	309	7.1
		1.8	152	3.5
		2.9	74	1.7

* "All requests" includes radar products not shown in this table, such as WSR-57 imagery from Limon, CO, and Cheyenne, WY, and NEXRAD algorithms.

The DAR³E-I system provides the flexibility necessary for forecasters to focus on synoptic- and subsynoptic-scale processes in both the cool and warm seasons, as changing weather conditions dictate.

In general, large-scale processes dominate the evolution of significant weather systems during the cool season, and subsynoptic processes are more dominant during the warm season. This is reflected in the product usage profiles for the 90-day warm and cool seasons (Tables 12a and 12b). For example, although the Nested Grid Model (NGM) was the preferred model by only a small margin during the cool season, it was clearly the model of choice during the warm season, garnering 43% of all model requests. This is consistent with the fact that the NGM was developed with improved diabatic physics and resolution for more accurate representation of the weak dynamic and highly convective situations characteristic of summer. Smaller-scale visible imagery was used considerably more frequently, relative to infrared imagery, during the warm season than during the winter. An obvious reason for this is that days are longer in the summer. A more subtle reason may be that visible imagery allows closer scrutiny of small-scale, low-level outflow boundaries that can initiate and focus convection. During the cool season, when large-scale features have an important influence on developing weather systems, North American-scale upper-air analyses were requested >3 times more often than during the warm season. This is not to suggest, however, that large-scale processes are important only during the cool season or that subsynoptic-scale processes need to be addressed only during the warm season.

Table 12a. Analysis of product usage on the DAR³E-I workstation

Category	Cool Season 1986-87		Warm Season 1987	
	Number	Percent	Number	Percent
Models	25448	49.8	14824	35.2
Satellite	6711	13.1	6768	16.1
Vertical	6746	13.2	6189	14.7
Radar	1074	2.1	5917	13.9
Surface	3344	6.5	3932	9.3
Upper Air	1775	3.5	550	1.3
Other	6020	11.8	3974	9.4
TOTAL REQUESTS	51118	100.0	42154	100.0

Table 12b. Expanded analysis of product usage on the DAR^{3}E-I workstation

Category	Cool Season 1986-87 Number	Cool Season 1986-87 Percent	Warm Season 1987 Number	Warm Season 1987 Percent
MODELS				
NGM	7197	28.3	6370	43.0
LFM	5564	21.9	2489	16.8
MAPS	1217	4.8	2176	14.7
Spectral	5508	21.6	1896	12.8
MRF	4851	19.1	1603	10.8
Other	1111	4.4	290	2.0
Total Requests	25448	100.0	14824	100.0
SATELLITE				
IR	4790	71.4	2762	40.8
Visible	1092	16.3	2223	32.9
Combo IR/VAS	483	7.2	1129	16.7
Combo Vis/IR	60	0.9	350	5.2
VAS	88	1.3	217	3.2
Other	198	3.0	87	1.3
Total Requests	6711	100.0	6768	100.0
VERTICAL				
Raob skew-t	4284	63.8	3174	51.3
Profiler skew-t	92	1.4	771	12.5
Profilers	2329	34.5	2097	33.9
Other	41	0.6	147	2.4
Total Requests	6746	100.0	6189	100.0
RADAR				
Total Requests	1074	100.0	5917	100.0
SURFACE				
Mesonet Products	1735	51.9	2658	67.6
SAO	1030	30.8	1086	27.6
Other	579	17.3	188	4.8
Total Requests	3344	100.0	3932	100.0
UPPER AIR				
Total Requests	1775	100.0	550	100.0
OTHER	6020		3974	
TOTAL REQUESTS	51118		42154	

Note: Use of radar products during the 1987 warm season is further stratified in Table 10. OTHER consists of products that could not be grouped in the other stratifications.

For example, a snow event is very often a mesoscale event and demands mesoscale products. Product usage analysis shows that DAR^{3}E-I provides the flexibility necessary to address any forecast problem, regardless of the season or primary scale of the weather event.

Nine of the fifteen most frequently requested products during both the cool and warm seasons were available only on the DAR^{3}E-I system.

The products requested most frequently during both seasons are listed in Table 13. In general, the greater demand is for data on larger spatial and longer temporal scales during the cool season, in contrast to the greater demand for smaller scale and more rapidly updating products during the warm season. Nine of the fifteen most-requested products during the cool and warm seasons were DAR^{3}E-only products, that is, products available on DAR^{3}E-I but not on AFOS. These include all satellite and radar images and MAPS products. In addition, although all the graphics contained in numerical model families on DAR^{3}E-I are available individually on AFOS, only DAR^{3}E has the capability to view all of the model data quickly and easily through use of the family graphic concept. The mesonet plot could be added to the category of DAR^{3}E-only products because, although it can be called up on AFOS, it cannot be animated (looped) as easily, and forecasters routinely take advantage of the DAR^{3}E-I looping capability when displaying the mesonet plot. It can be inferred from the popularity of DAR^{3}E-only products that the DAR^{3}E-I system is filling several needs of operational forecasters in Denver that were previously unmet. Granted, radar and satellite imagery were available to forecasters in limited form for a long time before DAR^{3}E-I was conceived, but the DAR^{3}E system provides for integration and manipulation of the data with other meteorological fields in a manner never before possible.

Product usage analysis provided the following insights into the forecasting process:

Numerical models generally dominated product selection on the days just before and after snow events. However, on snow event days, forecaster emphasis shifted from prognosis to diagnosis and from the larger scale to the mesoscale.

Product usage profiles for nine snow events of 1 inch or more were compared with a control group of eight fair-weather days during the winter of 1986-87. Because the effect of a snowstorm on a forecast office extends before and after the event, the analyses were performed over a 4- or 5-day "window" around each event. The number of requests generated every hour during these days was recorded to determine diurnal fluctuations in system demand. Requests for numerical model guidance dominated all selections on the days just before and after snow events. However, requests for diagnostic products such as satellite imagery, profilers, skew-t's, upper-air plots, and Doppler radar imagery (available operationally beginning in late March 1987) were maximized on the day of snow events. Indeed, once forecasters determined that a storm was probable, their emphasis shifted from prognosis to diagnosis and from the larger scale to the mesoscale, primarily in an effort to determine

Table 13. Most frequently requested products on DAR³E-I during cool and warm seasons, ranked by number of requests

Product	No. of Requests	Update Frequency
COOL-SEASON 1986-87		
*1. Infrared Satellite N. Amer. Scale	1538	3 h
*2. Infrared Satellite National Scale	1469	1 h
3. Mesonet Plot	1423	5 min
*4. Infrared Satellite Regional Scale	1219	30 min
*+ 5. LFM Family	1071	12 h
*+ 6. NGM Family	1034	12 h
7. Stapleton Profiler Wind Time/Ht X-sect	912	1 h
*8. Spectral Family	801	12 h
9. Plotted Satellite Info, Western U.S.	588	6 h
10. SAO Plot, Regional	586	1 h
*11. 500 mb Comparison Family	575	12 h
*12. Visible Satellite Regional Scale	543	30 min
^*13. Low-Lev. Radar Refl. on Local Scale	517	5 min
14. 120 h MRF 500 mb Height	491	24 h
15. Denver, CO, Skew-T Plot	473	12 h
WARM-SEASON 1987		
1. Mesonet Plot	2154	5 min
*2. Local-Scale Radar Refl. 0.7 deg.	1129	6 min
*3. IR Satellite National Scale	1010	1 h
*4. Visible Satellite State Scale	920	1 h
*5. Visible Satellite Regn'l Scale	832	30 min
*6. IR Satellite Regn'l Scale	814	30 min
7. Denver Profiler Skew-T Plot	771	1 h
8. Stapleton Profiler Wind Time/Ht X-sect	748	1 h
9. Denver Skew-T Plot	658	12 h
*10. National VAS WV/IR Combo	645	1 h
11. SAO Plot, Regional	643	1 h
*12. Local-Scale Radar Refl., 4.0	632	6 min
*13. Local-Scale Radar Refl., 0.2	554	6 min
*+ 14T. NGM Family Analysis	487	12 h
14T. Grand Junction, CO, Skew-T Plot	487	12 h

* DAR³E-Only Products
\+ Output from this model is available on AFOS but family graphics capability is available only on DAR³E
^ Available only small percent of the time
Update frequencies given for radar products are for surveillance (non-storm) mode

the spatial variation of snowfall amounts and the time when the event would begin and end.

On control days, forecasters apparently were looking well ahead to try to determine when a significant change might occur, because the Medium Range Forecast model (which provides forecasts out to 132 hours) was the most-requested numerical guidance (see Heideman, 1989b). Diurnal variations in system demand showed maximums in requests for products around shift changes, when outgoing forecasters briefed their replacements and incoming forecasters tried to get an understanding of the forecast problem of the day. There was a large variation in demand on the system throughout control days, but as storms approached, demand became less variable and forecasters used the system much more consistently throughout their shifts.

Overall, the results showed that forecasters modified their product selection patterns as changing weather situations dictated, and routine scheduled obligations had a large effect on diurnal fluctuations of system use.

During the convective season, forecasters tended to base their selection of workstation products on whether they expected conditions to become severe, and not on the extent of a severe weather outbreak.

McCoy and McGinley (1989) showed that Denver WSFO False Alarm Ratio (FAR; the ratio of the number of incorrect forecasts of occurrence to the total number of forecasts of occurrence) during the warm season of 1987 is smaller (suggesting an improvement) on very active severe weather days than on only marginally active severe weather days. Heideman (1989b) sought to determine whether differences in product selection on the DAR³E-I system could have been part of the reason for the improvement in FAR scores. Three stratifications (similar to those used by McCoy and McGinley) were used in the study (i.e., active severe, marginal severe, and nonsevere used as a control) which included cases occurring between 1 June and 31 August 1987. The difference between the active and the marginal severe stratifications was a function of the number of severe weather warnings and special weather statements issued during the day. Control days were clear days in most cases, when no warnings or special weather statements were issued. In terms of general product type, forecasters selected radar and numerical model guidance most frequently for all three stratifications. As might be expected, however, the proportion of all requests attributed to the numerical models was considerably less in summer than in winter, when synoptic-scale disturbances could often be tracked for 1 to 2 days upstream of Colorado. The disparity between use of radar reflectivity and velocity on all scales during the warm season of 1987 (Table 10) was evident in this study as well; even in severe weather situations forecasters used velocity images only sparingly. The NEXRAD algorithms were another class of products used only sparingly. On the 12 severe days included in this study, four NEXRAD algorithms, specifically designed for use in severe weather situations, were <u>collectively</u> requested only eight times. This is consistent with the overall dissatisfaction with these algorithms, expressed by forecasters in the questionnaires (see Walker, 1988, 1989).

When the products requested on severe days were grouped together (regardless of the marginal and active stratification) and compared with those requested on control days, several significant differences were revealed. For example, the proportion of requests for visible as opposed to infrared satellite imagery was much greater on severe days than on nonsevere days, and requests for 4-degree Doppler radar were much more frequently mixed with requests for low- and middle-elevation angles on severe days in an effort to assess the vertical structure of thunderstorms (a capability enhanced in 1988 with the introduction of the Weak Echo Region application program). However, there were no significant differences between products requested on marginal versus active severe weather days. Thus it appears that forecasters based their selection of workstation products on whether conditions would become severe, and not on the magnitude of a severe weather outbreak.

2.3. Forecast Evaluation

The two reports summarized below evaluate forecasts and warnings issued using the DAR³E-I system at the Denver WSFO. A number of forecasts and warnings issued during the pre-DAR³E-I years of 1983 and 1985 are compared with those issued during 1987, the first full year DAR³E-I was operational. The WSFO standard verification data were supplemented in each of these three years with observations collected by severe-storm chase teams participating in the PROFS real-time forecast experiments.

In the full report (McCoy and McGinley, 1989) the pre-DAR³E years of 1983 and 1985 are treated separately, but in this summary they are combined into a single data base, using weighted means, to be compared with the results from 1987. This enables us to compare forecaster performance before and after DAR³E-I was installed. Unfortunately, this comparison is confounded by the fact that the introduction of the DAR³E-I system was not the only change that occurred during the evaluation period. For example, the severe storm climatology varied significantly during the three years in the study. In addition, there were substantial changes (1) in the quantity and extent of verification data, (2) in forecaster staffing at the Denver WSFO, and (3) in the size of the study area. Furthermore, it must be remembered that the pre-DAR³E-I period of the study contains only two years and the DAR³E-I period only one year. Nevertheless, it still is constructive to perform a before-and-after comparison to investigate possible changes in forecaster performance.

The before-versus-after evaluation of severe weather warnings yielded the following findings.

> ***Assessment of tornado warnings revealed a number of substantial improvements in forecaster performance: warning lead times increased dramatically, the size of warned areas decreased, duration of warnings decreased, and the false alarm ratio declined. However, the probability of detection also declined.***

The evidence shows that in 1987 tornado warnings truly were becoming predictions rather than simple extrapolations of reports of funnel clouds or tornado touchdowns. Table 14 presents some of this evidence in terms of forecast lead time for "observed" warnings (i.e., warnings classified by the authors as being verified by qualified observers) as well as size of areas warned and warning duration for all warnings during the PROFS study period. Note that the mean lead time increased from 0.6 to 6.0 minutes (900%), the size of areas warned decreased from 1142 to 732 km^2 (36%), and the warning duration decreased from 35 to 28 minutes (20%). Some forecasters suggested that the improvement in tornado warnings can be attributed to the array of Doppler radar products available on DAR^{3}E-I and to the improved warning

Table 14. Mean forecast lead times, median size of warned area, and mean duration of warning for tornadoes and severe thunderstorms, before and after DAR^{3}E-I installation

	Before			After	% Change
	1983	1985	Avg. Weighted		
LEAD TIME (min)					
Tornadoes	1.1	0.0	0.6	6.0	+900
	(5)	(4)	(9)	(9)	
Thunderstorms	18.0	7.7	14.3	10.8	-24
	(23)	(13)	(36)	(18)	
AREA SIZE (km^2)					
Tornadoes	1,580	748	1,142	732	-36
	(9)	(10)	(19)	(16)	
Thunderstorms	2,820	2,088	2,543	1,940	-24
	(56)	(34)	(90)	(30)	
DURATION (min)					
Tornadoes	40	30	35	28	-20
	(9)	(10)	(19)	(17)	
Thunderstorms	60	60	60	59	-2
	(56)	(34)	(90)	(30)	

Note: The values in the before column are weighted means based on 1983 and 1985 data. LEAD TIME values are based on a subset of "Observed" Warnings. AREA SIZE and DURATION values are based on all warnings in PROFS study period.

Parentheses enclose the number of warnings used in the analysis.

generation program. In any case, the improvements should be cautiously interpreted in that some sample sizes were small (e.g., lead time) and there were some sizable fluctuations in the pre-DAR³E-I years (e.g., warned-area size). Nevertheless, if these results can be reproduced in coming years, it will represent an important improvement in the ability of the forecaster to target a potential tornado and warn of its presence prior to touchdown.

Table 15. False alarm ratio scores before and after DAR³E-I installation. Denver, full year, National Verification Program Scores are also presented.

	Before			After	Change (%)
	1983	1985	Avg. Weighted		
ONLY "OBSERVED" WARNINGS					
Tornadoes	.38 (8)	.6Ø (1Ø)	.5Ø (18)	.47 (17)	-Ø6
Thunderstorms	.43 (4Ø)	.55 (29)	.48 (69)	.4Ø (3Ø)	-17
ALL PROFS STUDY PERIOD WARNINGS					
Less Active (<5W)	.52 (23)	.82 (17)	.65 (4Ø)	.53 (3Ø)	-18
Active (≥5W)	.38 (42)	.5Ø (28)	.43 (7Ø)	.18 (17)	-58
Combined			.51 (11Ø)	.41 (47)	-2Ø
NATIONAL VERIFICATION PROGRAM SCORES					
Severe weather	.77	.77	.77	.68	-12

Note: The values in the before column are weighted means based on 1983 and 1985 data.

Parentheses enclose the number of warnings used in the analysis.

Further evidence of ability to predict tornadoes is given in Table 15 where the false alarm ratio shows a decline of 6% (representing an improvement) for observed warnings. However, the associated probability of detection (POD; the ratio of the number of correct forecasts to the total number of tornado events) also shows a decrease (representing a deterioration) of 31% (Table 16).

Table 16. Probability of detection scores before and after DAR³E-I installation. Denver, full year, National Verification Program Scores are also presented.

	Before			After	Change (%)
	1983	1985	Avg. Weighted		
ONLY "OBSERVED" WARNINGS					
Tornadoes	.38 (8)*	.5Ø (1Ø)	.45 (18)	.31 (17)	-31
Thunderstorms	.42 (4Ø)	.37 (29)	.4Ø (69)	.33 (3Ø)	-18
ALL PROFS STUDY PERIOD WARNINGS					
Less Active (<5W)	.5Ø (23)	.27 (17)	.4Ø (4Ø)	.29 (3Ø)	-28
Active (≥5W)	.52 (42)	.82 (28)	.65 (7Ø)	.58 (17)	-11
Combined			.56 (11Ø)	.39 (47)	-3Ø
NATIONAL VERIFICATION PROGRAM SCORES					
Severe weather	.53	.47	.5Ø	.48	-Ø4

Note: The values in the <u>before</u> column are weighted means based on 1983 and 1985 data.

Parentheses enclose the number of warnings used in the analysis.

The simultaneous improvement in FAR and decline in POD may simply be a reflection of the fact that, after DAR³E was installed, forecasters designated smaller warning areas and shorter durations. Although these changes may have contributed to the improvement in FAR, there is some evidence that POD was adversely affected by warning boxes that were slightly too small and durations that were slightly too short. These results suggest that the forecasters need more experience with the high degree of precision made possible by the DAR³E warning generation program. Forecaster comments suggested that much of the improvement in tornado warnings that has already been achieved can be attributed to the array of Doppler radar products available on DAR³E-I and to the improved warning generation program.

Assessment of severe thunderstorm warnings indicated some sizable improvements (though typically smaller than improvements to tornado warnings) after the DAR³E-I System was installed: the size of warned areas decreased, and the false alarm ratio decreased. However, the probability of detection also declined.

The severe thunderstorm warnings in 1987 show some improvement, with a 24% reduction in size of area warned (Table 14). The change in the FAR is a 17% decrease for observed warnings (Table 15) with an associated decline in the POD score of 18% (Table 16).

A further analysis combines all tornado and severe thunderstorm warnings issued during the PROFS study period and stratifies them into "active warning days" (i.e., days with five or more warnings issued) versus "less active warning days" (days with fewer than five warnings issued). Now the FAR score shows a decline of 58% for active days, a decline of 18% for less active days, and a decline of 20% for the combined data set (Table 15). The associated POD scores for the combined data set (Table 16) show a decline of 30%. These partial summertime results are consistent in direction and to a lesser degree in magnitude with full-year severe weather scores published by the NWS National Verification Program; the NWS FAR score declined by 12% and the POD score declined by 4% (P. Leftwich, 1987, Personal Communication from NSSFC).

The decline in both FAR and POD scores may again be partially due to the selection of smaller warned-area sizes. However, in 1987 forecasters monitored vertical storm structure by comparing lower with higher elevation scans in an effort to distinguish growing storms from dissipating storms (see Heideman, 1989b), and hence refrain from over-warning. These two actions may account for the relatively large change in the FAR score during active warning days (i.e., -58%, Table 15).

Verification analyses of temperature and precipitation forecasts were also undertaken. These studies investigated the possible effects of the DAR³E-I system on temperature and precipitation forecasts for a cool (1986-87) and a warm (1987) season in contrasting climatic regimes (i.e., Denver and Grand Junction, Colorado). The comparative results for the periods before and after installation of the DAR³E-I system can be summarized as follows:

Cool and warm season 0-12 hour temperature forecasts show a small and consistent reduction in error from the overall 20-year linear trend, after installation of DAR^{3}E-I.

The 0-12 hour (1st period) forecast was chosen because it is one of the items in the National Verification Plan and because an extensive data base already exists (McCoy and Kent, 1989). Table 17 presents the rank of the mean absolute error (MAE) for approximately 720 maximum and minimum temperature forecasts during the cool and warm seasons of 1986-87 versus the linear trend of the historical yearly means (1966-1985). We see that irrespective of season and location, the rank of the MAE for temperature forecasts in 1987 is smaller than the linear trend and represents some of the largest yearly decreases (implying an improvement). If these results persist in subsequent years, one implication is that temperature forecasts can be improved by the additional mesoscale information available on the DAR^{3}E-I system.

Table 17. Comparison of the 0-12 hour cool-season (1986-87) and warm-season (1987) temperatures and precipitation forecasts with the pre-DAR^{3}E-I (1965-1985) trend of values for Denver and Grand Junction

		Cool Season	Warm Season
A. TEMPERATURE (RANK OF CHANGE IN MEAN ABSOLUTE ERROR)			
DEN		3rd largest decrease	2nd
GJT		2nd	2nd
B. PRECIPITATION (CHANGE IN STANDARDIZED SKILL SCORE)			
DEN		+7	-2
GJT		-2	-4
PRECIPITATION (RELIABILITY; d > |0.10|)			
DEN			
	1986-87	4	2
	1970-71	2	1
GJT			
	1986-87	3	2
	1970-71	2	1

Cool and warm season 0-12 hour precipitation forecasts do not show any notable changes in skill or reliability after installation of DAR³E-I.

The mean performance scores (i.e., half-Brier Score, standardized with respect to climatology) for both the cool and warm seasons were compared with the linear trend of approximately 20 pre-DAR³E-I scores. Table 17, Part B, presents a summary of these results, and one notes that the changes in skill scores from the past trend typically are small. Even the increase of +7 for Denver in the cool season ranks only sixth in the nine largest increases. Furthermore, the changes for Denver are not consistent over the two seasons. Finally, compared with the variation about the trend, the changes after the installation of DAR³E-I appear to be in the noise level.

In a further validation effort, reliability diagrams were constructed for each of the locations and both seasons. Part B, Table 17, presents a quantitative summary of these results as measured by the number of differences (d) between the forecast probability and the observed frequency of occurrence that were greater than ±10 points (i.e., d > |0.10|). The counts in the pre-DAR³E-I period (1970-1971) and the DAR³E-I period (1986-1987) are rather similar (e.g., 2 versus 1 for the warm season at both locations), and thus there is little evidence of a change in the reliability of the precipitation forecasts.

3. SUMMARY AND CONCLUSIONS

From the principal findings of the six PROFS evaluation group reports come the following conclusions. Forecasters view the DAR³E-I system as a substantial improvement over the AFOS technology. DAR³E-I provides them with a view of the atmosphere not previously possible, by integrating essential meteorological data sets into a single workstation and combining them with overlay, looping, automatic update, and color-enhancement capabilities. Product usage statistics show that the system also provides the flexibility necessary for forecasters to monitor meteorological processes on all scales of motion. An inventory of products available on DAR³E-I reveals that most of the product sets unique to the system, particularly the suite of Doppler radar products, enable forecasters to monitor mesoscale processes effectively. In addition, the DAR³E-I system provides an improved warning generation program that allows forecasters to issue warnings that are more area specific and time specific than is possible on AFOS. Furthermore, this program enables forecasters to issue warnings more quickly and easily than before.

Though DAR³E-I has been operational only since December 1986, there is evidence that integrating extensive data sets with advanced data-manipulation capabilities has already had a positive effect on forecasts and severe weather warnings. For example, assessment of tornado warnings showed substantial improvement in the following areas: lead times increased dramatically, the size of warned areas and duration of warnings decreased, and the false alarm ratio declined. Assessment of severe thunderstorm warnings revealed similar, albeit less dramatic, improvements. Finally, the accuracy of 0-12 hour temperature forecasts has improved slightly following the installation of DAR³E-I, though no such improvement is evident for comparable 0-12 hour precipitation forecasts.

Forecasters did indicate a few areas where DAR³E-I needs improvement. Most important, the speed of the system in acknowledging and executing commands must be faster, and the VAXstation text interface must provide standard word processing features. These problems are being addressed and corrected in DAR³E-II. Furthermore, the present decline in the POD scores highlights the need to continue to evaluate the DAR³E system and understand how forecaster performance can be improved.

Based on data for only 1 year following the installation of DAR³E-I, the findings are encouraging but by no means conclusive. The limited sample and the problems inherent in making the before and after comparisons clearly require that the evaluation effort continue and expand. As part of this effort, responses to a third questionnaire focusing on the forecasters' assessment of DAR³E-I during the summer of 1988 are currently being analyzed. Further questionnaires will sample forecasters' views and provide feedback to designers and NWS management during implementation of DAR³E-II. Enhanced analysis of product usage on DAR³E-II will provide a full accounting of every important interaction between the forecasters and the DAR³E-II system, including use of automatic updates, application programs, and the use of individual AFOS graphics within model families. The effort to evaluate forecasts and warnings will continue when DAR³E-II becomes operational, including an attempt to improve the local verification data base and develop new scoring techniques commensurate with the forecasting capabilities made possible by DAR³E technology.

This overview is not exhaustive; the findings discussed generally highlight inherent DAR³E-I system capabilities. For additional details and insights concerning work already completed, the reader is urged to review the referenced reports.

ACKNOWLEDGMENTS. We wish to thank the many PROFS personnel who offered advice and comments on earlier versions of this report. Particular thanks go to P. Amstein, C. Bullock, H. Grote, and D. Small. We also wish to thank L. Dunn, R. J. Wagoner, and D. Walts, NWS, for their advice. Finally, special thanks to M. McCoy for data review and comments.

4. REFERENCES

Heideman, K.F., 1989a (Revised): Report on cool and warm season operational use of DAR³E-I workstation products. PROFS Evaluation Group report, Boulder, Colo., 66 pp.

Heideman, K.F., 1989b (Revised): Use of DAR³E-I workstation products at the Denver WSFO in association with snow events and convective outbreaks. PROFS Evaluation Group report, Boulder, Colo., 46 pp.

McCoy, M.C., and T. Kent, 1989 (Revised): A meteorological assessment of DAR³E-I, Part 2: Temperature and precipitation verification results. PROFS report, Boulder, Colo., 37 pp.

McCoy, M.C., and J.A. McGinley, 1989 (Revised): A meteorological assessment of DAR^{3}E-I, Part 1: Severe storm warning verification results. PROFS report, Boulder, Colo., 41 pp.

Walker, D.C., 1988 (Revised): An analysis of forecaster satisfaction with the DAR^{3}E-I system. PROFS Evaluation Group report, Boulder, Colo., 24 pp.

Walker, D.C., 1989 (Revised): Forecaster assessment of the DAR^{3}E-I system PROFS Evaluation Group report, Boulder, Colo., 54 pp.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1988—673-002/92,039 REGION NO. 8