U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SYSTEMS DEVELOPMENT
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 83-18

RESULTS OF THE AFOS MONITORING AND UPDATING FIELD TESTS

David J. Vercelli, Gene Norman, Jr., and Ward R. Seguin

December 1983

RESULTS OF THE AFOS MONITORING AND UPDATING FIELD TESTS

David J. Vercelli, Gene Norman, Jr., and Ward R. Seguin

1. INTRODUCTION

The operational field tests of the automation of field operations and services (AFOS) aviation terminal forecast (FT) monitoring and updating system have been completed at National Weather Service Forecast Offices (WSFO's) located in Cleveland (CLE), Des Moines (DSM), Birmingham (BHM), and Portland (PDX). This report is based on the data collected during these tests.

The field tests of the monitoring and updating system had three primary goals. The first goal was to determine the impact of the information provided by the programs on the current preparation, monitoring, and updating of FT's. The second goal was to determine the quality of this information, and the third was to assess the value of this information to the forecaster and the WSFO. For a detailed description of the field test procedures, see NWS (1983a; 1983b; 1983c).

2. SOFTWARE DESCRIPTION

The monitoring and updating system consists of two FORTRAN applications programs which were run in the background partition of the WSFO's Data General Eclipse S230 minicomputers. As applications programs, both needed to be initiated either manually or through the use of an AFOS console procedure.

The first program, MONITR (Vercelli, 1983), is designed to monitor the FT's within a WSFO's area of responsibility. This is done by comparing the elements in the FT's with the corresponding elements in the surface airways observations (SAO's) and with both the official terminal forecast amendment criteria and a set of alert criteria. The alert criteria are designed to provide warning that, although an amendment is not required at present, conditions are such that one may be required in the near future. If the program determines that a criterion has been met at one or more of the terminals, it generates an appropriate message for each of the problem terminals as part of the output display. A message is also generated if an SAO or FT decoding error is encountered by the program. In addition to detailing the problems, MONITR also summarizes the status of each of the remaining FT's.

It should be noted that the amendment and alert decisions made by MONITR are based only on the FT's prevailing conditions. Currently, the FT decoding program does not decode the FT remarks and, therefore, they do not take part in the monitoring process. They are displayed as a literal statement in the output product, however, for the forecaster's use.

The second program, GEM (Herrmann, 1983), which stands for Generalized Exponential Markov (Miller, 1981), provides an updated guidance forecast for those terminals where MONITR determines an amendment is required or may be required shortly. The guidance forecast for a particular terminal is generated based on the information contained in that terminal's SAO. The

forecasts are valid for projections of 1, 2, 4, and 6 hours from the SAO time with guidance provided for each of the surface observed elements. The forecaster can also generate guidance forecasts for any location for which an SAO is available in the local database by entering the proper command at the alphanumeric display module (ADM).

3. TEST DESCRIPTION

The first field test was conducted at WSFO CLE between April 26 and May 31, 1983. Development of the field test plan involved coordination between the Office of Systems Development's Techniques Development Laboratory (TDL), Office of Meteorology's (OM's) Aviation Services Branch, OM's Services Evaluation Branch, Eastern Region Headquarters, and WSFO CLE. The field test plan (NWS, 1983a) was approved by each of the organizations.

The field test plan was used at WSFO's DSM, BHM, and PDX with slight modifications to the questionnaires. The WSFO's and their respective regional headquarters approved the updated plans (NWS, 1983b--Central and Southern Regions; NWS, 1983c--Western Region). The DSM and BHM tests began on June 14, 1983. The tests ended on July 13, and July 14, 1983, for DSM and BHM, respectively. The PDX test began on October 10, 1983, and ended on November 10, 1983.

For the duration of the tests, both the MONITR and GEM programs were resident on the Eclipse minicomputer disk designated as DPOF. The programs were initiated automatically three times each hour through the use of an ongoing AFOS console procedure. The GEM program would automatically generate a guidance forecast for each terminal for which MONITR determined a problem or potential problem existed with an FT. In addition, the aviation forecasters were asked to run the GEM program manually for each terminal sometime during the 2-h period prior to each of the three regularly scheduled FT issuance times. The GEM program could also be run manually by the forecasters at any other time if they so desired.

The aviation forecasters were asked to fill out two questionnaires during the test period. The first, Questionnaire A, was to be filled out hourly. The forecasters were instructed to fill it out at approximately the same time each hour. If they were unable to fill it out for a particular hour, they were not to go back and fill it out later; it was to be left blank for the missed hour.

The second questionnaire, Questionnaire B, was to be completed following a forecaster's tour of duty at the aviation desk which would be every 3 to 5 days depending on shift schedules. Some offices, however, decided to set specific dates when the questionnaire would be filled out because of the irregularity of some forecaster's schedules.

Upon examining the responses received from WSFO CLE, it became obvious that some of the questions were not worded properly and that the selection of possible answers did not cover all cases. Therefore, some modifications were made to the questionnaires prior to the DSM and BHM tests, and then again prior to the PDX test.

TDL was an active participant during each of the field tests. An AFOS console procedure was kept running on the TDL Eclipse minicomputer throughout each field test with the FT's and SAO's for each test terminal being archived and hard copy made of the MONITR program output. This was done to supplement data collected through the questionnaires and to allow rapid responses to field questions if any problems arose with the MONITR program. The GEM program was not run on a regular basis on the TDL system but could be run when necessary.

4. SUMMARY OF RESULTS FOR QUESTIONNAIRE A

Each question is summarized individually, and comments are included as appropriate. Responses are labeled with the WSFO call letters to which they apply. The percentages shown are based on the total possible number of hourly responses during the test period (CLE-824 hours; DSM-736 hours; BHM-712 hours; PDX-768 hours) except where noted. The percentages have been rounded to the next highest percent so that totals may not always add to exactly 100%.

The summary of the PDX data shows that the "no response/not applicable" category accounted for a high percentage of the test hours. There were two main reasons for this. First, we inadvertantly provided PDX with too few questionnaires at the start of the test. When the supply ran out, PDX copied additional forms but used the forms designed for the WSFO CLE test. When the problem was discovered, PDX stopped answering the questions until we sent a new supply of the correct questionnaires. Although the CLE and PDX questionnaires were similar, several questions had to be deleted from the PDX sample. The second reason was that, for the 12 day period between October 29 and November 9, the swing shift questionnaire was missing, as were five day-shift and two midnight-shift questionnaires.

Question: Was the MONITR procedure running this hour?

Responses: Percent of time

	CLE	DSM	BHM	PDX
Yes	93	90	88	56
No, degraded mode	1	0	9	1
No, other reason	3	5	3	7
No response	3	0	1	35
Other	1	5	1	1

Generally, if the procedure was not running for a reason other than degraded mode computer operations, it was because of severe weather outbreaks or AFOS crashes where the procedure was not restarted. The "other" category includes those responses where the procedure was running only for a portion of the hour or where the forecaster was running the programs manually rather than by using the procedure. The "no response" category is indicative of those times when the forecaster did not answer the questionnaire or when the questionnaire was missing.

Question: Did running MONITR have any noticeable impact on the use of AFOS by other forecasters?

Responses:

	Percent of time			
	CLE	DSM	BHM	PDX
Yes	1	1	1	3
No	93	91	88	56
Not applicable/no response	7	9	12	41

There were four predominant reasons given for the MONITR procedure to have had an impact on the use of AFOS. These were interference with message composition at the console where the procedure was running, disruption of typing at that console, delaying of other applications programs from running on time, and being a contributing factor in AFOS crashes during times of heavy computer use. One forecaster also indicated that MONITR itself was delayed by the running of other applications programs. There was also some interference with use of the Dasher due to the standard foreground program initiation or completion messages being printed.

The last category refers to those times when the MONITR procedure was not running or when other priorities (such as severe weather occurrence) prevented forecaster completion of the questionnaire.

Question: Did MONITR provide you sufficient information to make a firm amend or no amend decision?

Responses:

	Percent of time		
	CLE	DSM	BHM
Yes	45	11	23
No, required using log	14	2	0
No, required using log and other information sources	18	13	42
Not applicable	12	71	34
No response	11	4	1

The "not applicable" category includes those cases when no amendment or alert messages were issued by the program and when the MONITR procedure was not running. "No response" is interpreted to mean that the forecaster was too busy to respond to the question. The weather during the DSM test was reasonably good with few amendment messages being issued by the program; this may account for the high "not applicable" response.

Question: Did you amend any FT's earlier this hour because MONITR provided you with an earlier warning than you might have otherwise had?

Responses:

	Percent of time
	PDX
Yes	7
No	31
Not applicable	18
No response	44

In several cases, the comments written indicate that the forecasters did not amend an FT, although MONITR indicated amendment was necessary, because the event was covered in the next hour's forecast. The ability of MONITR to "catch" specials was, in some cases, responsible for forecaster issuance of amended FT's earlier than they might have otherwise been.

Question: For those terminals where MONITR indicated an amendment may be required (i.e., ALERT), did you amend those FT's?

Responses: Percent of time

	CLE	DSM	BHM	PDX
All	3	1	4	1
Some	8	2	6	8
None	46	21	38	31
Not applicable/no response	43	76	52	60

Here, the last category includes those cases when the MONITR procedure was not running, the forecaster did not have time to respond, or no "ALERT" messages were issued by MONITR.

Question: Do the alert criteria need modifications?

Responses: Percent of time

	CLE
Yes	41
No	35
Not applicable/no response	24

Question: Did MONITR adequately handle all meteorological cases (i.e., decoding FT's and SAO's)?

Responses: Percent of time

	CLE
Yes	69
No	14
Not applicable/no response	18

Since the question did not ask for comments when a problem was observed, few comments were given. This question was modified prior to the other tests to request comments when a problem was detected. It is interesting to note that several forecasters responded "no" because of errors in the way the SAO or FT had been entered into the database (e.g., C3 BKN which should have been C30 BKN). In a few cases, the forecaster requested the observer to issue a corrected observation.

Question: Did you notice any errors in the MONITR output product?

Responses:

	Percent of time		
	DSM	BHM	PDX
Yes	2	3	5
No	89	85	50
Not applicable/no response	9	12	45

Several decoding errors were discovered during the tests. These were: MONITR not recognizing a report of thunder when no precipitation was occurring; decoding an FT visibility of 1 1/2 mi as 11 1/2; decoding a cold frontal passage as a ceiling; and not handling a corrected SAO or a delayed FT properly. Occasionally, when a certain combination of events occurred, an incorrect multiplication resulted in an integer overflow error. All of these errors have since been corrected.

It should be noted that the decoders expect to see the SAO's and FT's in a certain format. If the products are not entered properly by the forecaster or the observer, then the decoders may lose track of what they are decoding resulting in error messages that may not be related to the actual error. This can best be handled by more care being exercised when the products are entered into the computer.

Another problem mentioned by forecasters was MONITR producing an "SAO OUTSIDE FT VALID TIME" message. This occurs only when the forecaster issues a new or amended FT which has a valid time later than the current clock time. If MONITR runs prior to that valid time and a new observation is found (either an hourly or a special), MONITR will produce the above message. This occurs because MONITR looks only at the latest forecast in the database and, under these circumstances, it is the "previous version" of the FT that is technically still valid.

One additional error was discovered during the PDX test. If the forecasted visibility is 7 miles or greater, then it does not actually appear in the forecast. An example would be the following:

C50 BKN RW- 3520G35.

The error arises when the FT decoder used by MONITR looks for the visibility. If a visibility is found, then it looks for a weather group. If no visibility is found, but a weather group exists, then the decoder erroneously treats everything between the weather and the end of the FT sentence (i.e., the period) as a remark. We are working on this problem, and it should be corrected when Version 2.0 of MONITR is issued.

Question: For ZZV (Zanesville) FT only, where MONITR indicated an amendment was necessary, did you not amend because, in your judgment, you did not believe the amendment condition would persist?

Responses:

	Percent of time
	CLE
No amendment, condition temporary	4
No amendment, other reason	3
No response	93

There were few occurrences when MONITR indicated an amendment for ZZV was necessary. The "no response" category includes those cases when MONITR indicated an amendment was necessary but the forecaster did not have time to answer the question, the procedure was not running, or when MONITR did not issue an amendment message for ZZV.

Question: For ZZV FT only, if you did amend the FT, did you find the GEM guidance useful?

Responses:

	Percent of time
	CLE
Yes	1
No	3
Not applicable/no response	96

Here, as with the last question, few responses were received from the forecasters. Due to the very small sample sizes, no attempt will be made to draw conclusions from either question.

Question: Where MONITR indicated an FT amendment was necessary, in how many cases did you not amend because you felt the amendment condition would not persist?

Responses:

	Percent of time		
	DSM	BHM	PDX
All (100%)	6	29	13
Most (>50%)	0	3	2
Some (<50%)	1	3	7
None (0%)	11	11	5
Not applicable	82	54	73

There were no prevailing comments provided with the responses to this question. The "not applicable" category includes those cases when no amendment messages were issued by the program, when the forecaster did not respond to the question, and when the questionnaire was missing.

Question: Did you initiate GEM manually during the hour?

Responses:

	Percent of time			
	CLE	DSM	BHM	PDX
Yes	14	18	11	8
No	78	74	78	51
Not applicable/no response	8	8	11	41

The percentages for the following questions are based on the "yes" responses to the last question for CLE only (114 out of 824 responses).

Question: Did running GEM have any noticeable impact on the use of AFOS by other forecasters?

Responses: Percent of time

	CLE
Yes	4
No	87
Not applicable/no response	9

There was one case of GEM possibly contributing to an AFOS crash during a period of heavy use. The running of the program also delayed the running of another program used in producing the hydrological summary for transmission.

Question: Did you use the GEM guidance in preparing or amending any FT's besides ZZV?

Responses: Percent of time

	CLE
Yes	21
No	68
Not applicable/no response	11

Question: In your judgment, was the GEM guidance produced useful when balanced against the time it takes to produce it (cost/benefit)?

Responses: Percent of time

	CLE
Yes	16
No	67
Not applicable/no response	18

Question: Was this the planned GEM run for all terminals prior to a regularly scheduled FT issuance?

Responses: Percent of time

	CLE
Yes	70
No	23
Not applicable/no response	7

The forecasters were asked to run GEM manually for all stations prior to each of the three regularly scheduled FT issuance times, and they were encouraged to run it any time they wished to see updated guidance. This question was designed to distinguish between the scheduled runs and the optional runs.

5. SUMMARY OF RESULTS FOR QUESTIONNAIRE B

Questionnaire B was to be filled in by each forecaster at least twice during the test. The questionnaire was designed to allow the forecasters an opportunity to comment on specific aspects of the monitoring and updating system and to provide information on areas not covered by Questionnaire A.

The numbers shown below for each question are the actual number of responses received. They have not been converted to percentages. As with Questionnaire A, some modifications were made to some of the questions between the CLE test and the other tests. The call letters of the WSFO's to which the responses apply are included as appropriate.

Question: Did you feel that MONITR was a useful tool for you?

Responses:

	CLE	DSM	BHM	PDX
Yes	17	14	12	7
No	9	2	1	5

Question: If you did not feel it was useful, why not?

Responses: Many of the forecasters pointed out the need to decode the remarks. Without including the remarks in the monitoring process, the alerts and warnings issued by the program are more frequent because important information is not taken into account. Another forecaster did not feel the program gave sufficient information for decision making. Also, some forecasters answered "no" because they didn't feel that they had had enough experience in using the program to properly judge it. There were two "no" responses which indicated confusion on which program was being referred to in the question. Although the question stated MONITR, both forecasters talked about the persistence of the forecasts (i.e., GEM), but one added that the "alerts on stations requiring amendments is good, however." The confusion may be related to the fact that the procedure used to initiate both MONITR and GEM was itself called MONITR.

Question: Did MONITR save you any time you were able to use in preparing or amending FT's or in performing other duties?

Responses:

	CLE	DSM	BHM	PDX
Yes	3	7	3	2
No	22	11	10	10

Question: If your answer was yes, how did you spend the time?

Responses: Most forecasters spent the extra time analyzing the current weather situation and preparing amendments. A few indicated they used the time to complete this questionnaire.

Question: Did the program significantly delay you from performing your duties and in what way?

Responses: Most forecasters indicated that answering the frequent MONITR alarms did delay them. Another important response was that these programs were run as part of a procedure which tied up a console. This highlights a need for an AFOS foreground function to automatically trigger background programs by time or receipt of a product or other information. This would allow MONITR to start upon the arrival of an SAO and would not tie up the console.

Several responses referred to the additional time necessary to fill out this questionnaire.

Question: Are there any changes you would like to see made in the program (e.g. method of running the program, display format, information displayed, etc.)?

Responses: There were several suggestions some of which have already been mentioned. First, decode the remarks and use them in the monitoring process. This would eliminate some of the alarms. Secondly, change the display so that the current FT and recent SAO's are displayed simultaneously. This would allow the forecaster to examine the weather trends, something he or she is now able to do with the manual log sheets. Thirdly, initiate the program from foreground upon the receipt of SAO's so as not to tie up a console. Lastly, eliminate the alarm light on the alert message or eliminate the alert message totally. In the case of PDX, the procedure was running at the focal point's console and several suggested it be moved to the aviation console. A couple of forecasters suggested that MONITR should be more tolerant when the observed conditions are better than the forecast. This is possible as far as the alert criteria are concerned since they are not official criteria. More tolerance should not be allowed on the amendment criteria, however; these are defined specifically in the NWS Operations Manual (see NWS, 1983d).

Question: Were the alert messages (i.e. amendment <u>may</u> be required) helpful to you?

Responses:

	CLE	DSM	BHM	PDX
Yes	14	10	12	5
No	11	6	1	6
Not sure	0	0	0	1

Question: Should the alert criteria be modified in any way?

Responses:

	CLE	DSM	BHM	PDX
Yes	14	7	6	6
No	9	10	7	1
Not sure	0	0	0	5

Question: If your answer was yes, how?

Responses: Again, the predominant response was to decode the remarks and use them in the monitoring process. Another forecaster recommended that the alert criteria be eliminated altogether. Two offices mentioned the need to improve the handling of thunderstorm conditions. Thunder alone, for example, was not being picked up by the program. This was a program bug which was corrected prior to the PDX test.

Question: Were there an unreasonable number of errors?

Responses:

	CLE		DSM		BHM		PDX	
	yes	no	yes	no	yes	no	yes	no
FT decoding errors	3	22	1	15	0	12	0	11
SAO decoding errors	4	20	6	11	2	9	1	10
Other	0	13	0	18	0	8	2	8

Question: Were the diagnostics printed at the Dasher (CLE test) or ADM (DSM, BHM, and PDX tests) helpful?

Responses:

	CLE	DSM	BHM	PDX
Yes	9	12	11	3
No	17	4	2	9

Question: If your answer was no, why not?

Response: The principal response from CLE was that the forecasters did not understand the error messages since the error codes were not defined. The fact that they were printed out on the Dasher was an inconvenience. For the DSM and BHM tests, the messages were defined better and displayed at the ADM as part of the MONITR output product. Although the messages for the PDX test were the same as for the DSM and BHM tests, the forecasters appeared not to understand what they meant. When an FT or SAO decoding error was detected, the message briefly stated in words what the decoder error return code was. For example, "temp/press error" would mean that, in decoding the SAO, a problem was detected while searching for the temperature or pressure variables. Despite the abbreviated error messages, we felt this method was preferred to our printing out the error return code number at the

Dasher. The latter method would then require the forecaster to look up the corresponding error message in a table.

Question: Should the FT remarks be decoded and included in the monitoring process or left as a literal in the display?

Responses:

	CLE	DSM	BHM	PDX
Left as is	5	6	3	2
Decoded	20	11	10	8

Question: Can the MONITR display replace your FT log?

Responses:

	CLE	DSM	BHM	PDX
Yes	3	3	4	5
No	22	14	9	4
Blank	0	0	0	3

Question: If your answer was no, would you like to see MONITR modified so that it could replace the log?

Responses:

	CLE	DSM	BHM	PDX
Yes	5	4	4	0
No	17	9	3	3
Blank	0	0	0	9

Question: How could it be modified?

Response: The CLE, DSM, and BHM offices responded by suggesting the display of the current FT and the last several SAO's. The PDX office has done away with the log to which this question referred.

Question: Do you have any additional comments on the MONITR program?

Responses: Answers to this question were mixed. One response suggested the elimination of the alert messages. Another expressed appreciation for it. But all offices agreed that the remarks should be decoded and used in the monitoring process.

Another commenter expressed appreciation for being able to catch special observations earlier than normal and stated that MONITR was particularly useful during busy weather situations. Finally, one forecaster did not like the program running as part of a procedure.

Question (CLE): Did you feel that GEM guidance was useful?

Response: There was one "yes" response and twenty-three "no's."

Question (CLE): If your answer was no, why not?

Response: The predominant response was that GEM followed persistence too closely and didn't pick up changes. One forecaster remarked: "I saw only one front. It had a sharp wind shift which GEM blew completely."

Question (CLE): If yes, under what conditions did you find GEM guidance to be most useful?

Response: One forecaster noted that the diurnal wind variation showed winds higher in the afternoons as would be expected.

Question (DSM, BHM, and PDX): GEM utilizes persistence very much. Were you helped when it deviated from persistence?

Responses:

	DSM	BHM	PDX
Yes	1	3	2
No	12	10	10
Other	4	0	0

Question (DSM and BHM): If GEM deviated from persistence, what did it do best?

Responses:

	DSM	BHM
Turning them off	0	0
Turning them on	1	2
Both	0	1

One forecaster noted that forecast changes by GEM were in the wrong direction.

Question (PDX): If GEM deviated from persistence, under what circumstances was it most useful?

Responses: Several forecasters commented that it did not deviate from persistence, but one forecaster found it useful when ceilings were below 1000 feet.

Question (DSM, BHM, and PDX): GEM can be run at any time because it doesn't need model output, upper air, spatial, or temporal information. Does the immediacy of the forecasts overshadow the fact that such information is lacking?

Responses:

	DSM	BHM	PDX
Yes	4	2	2
No	8	6	5
No opinion	4	5	5

Question (DSM, BHM, and PDX): Which of the forms of output would you prefer?

Responses:

	DSM	BHM	PDX
Present categorical	7	7	3
Probabilities	1	0	0
FT format	5	3	3
Combination of these	6	3	4

Question (DSM and BHM): Did you feel that the GEM guidance was useful when balanced against the time needed for it to be produced?

(PDX): Did you feel that the GEM guidance was useful?

Responses:

	DSM	BHM	PDX
Yes	7	4	3
No	10	9	9

Question (DSM, BHM, and PDX): If your answer was no, why not?

Responses: Again, the principal complaint was that GEM followed persistence too much. One forecaster pointed out that GEM forecasted fog with a large temperature-dew point difference. This was caused by a software bug which has now been corrected.

Question (DSM, BHM, and PDX): If your answer was yes, under what conditions did you find the GEM guidance to be most useful (e.g., static weather situation, gradually or rapidly changing weather situation, etc.)?

Responses: Most people felt that GEM would adequately handle static or gradually changing conditions. One forecaster felt that it had problems with trends. Another forecaster said: "A forecaster, if he's up on the synoptic situation should, through the use of satellite pictures, analyses, and other tools, be able to make an FT as quickly and more accurately than GEM." At the other end of the spectrum was the forecaster who stated, "Situations when I wasn't sure what to forecast myself."

Question (CLE, DSM, BHM, and PDX): Are there any changes you would like to see made in the GEM program (e.g. method of running the program, display format, information displayed, etc.)?

Responses: Forecaster response to this question varied considerably and no single response dominated. Comments ranged from "format is good" to "I'd like to see it in FT format." One individual wanted the software set up so that it could run off any observation and another remarked: "Get rid of that silly advertising at the top. We know TDL did it." There were valid operational reasons for removing the "advertising." It's removal provided more room on the screen for the overlaying of such data as SAO's. The display format was changed as requested prior to the PDX test.

Question (CLE, DSM, BHM, and PDX): Were there an unreasonable number of errors?

Responses:

	CLE	DSM	BHM	PDX
Yes	3	2	2	1
No	21	16	11	9

Question (CLE, DSM, BHM, and PDX): If your answer was yes, what type of errors were most troublesome?

Responses: Two specific errors were mentioned. The first error was that GEM forecasted dust and visibility restrictions when it was obvious that there wouldn't be any, and the second was that fog was forecast with a 15° temperature-dew point difference. Some forecasters also mentioned incorrect forecasts due to GEM's strong reliance on persistence. For example, thunderstorms were carried for 6 hours whenever it appeared in the observation. Another commented that there were many GEM RDS errors. These are related to reading data. The cause of this has not been determined yet.

Question (CLE, DSM, BHM, and PDX): Do you have any additional comments on the GEM program?

Responses: On the positive side, one forecaster offered: "Has good potential. With enough climatological data incorporated into the problem, the forecasts should be much better." And there were some negative comments. For example: "Totally useless. For example, GEM forecast scattered clouds most places on May 2nd, the day that had the most significant outbreak of severe thunderstorms and tornadoes in Ohio in several years." Another forecaster expressed concern over GEM's inability to handle local effects but offered, "Perhaps equations developed for single stations would be better. This would account for at least some local effects."

Question (CLE, DSM, BHM, and PDX): In general, how did you feel the GEM guidance forecasts were for the following?

Responses: (See Table 1.)

Table 1. Field responses to the question: In general, how did you feel the GEM guidance forecasts were for the following?

Element	Cleveland					Des Moines			
	Good	Fair	Needs Work	Other	No Reply	Good	Fair	Needs Work	Other
Temperature	4	16	3	1	1	9	7	1	0
Dew Point	4	16	3	1	1	6	9	2	0
Visibility	1	6	14	2	2	1	8	8	0
Weather	1	5	17	0	2	2	2	13	0
Wind direction/speed	0	10	13	0	2	1	10	6	0
Sea level pressure	1	14	4	3	3	3	8	4	2
1st cloud layer amount	1	9	12	1	2	3	9	5	0
1st cloud layer height	0	11	11	1	2	2	10	5	0
2nd cloud layer amount	1	8	13	1	2	2	8	6	1
2nd cloud layer height	0	8	13	2	2	2	8	6	1
Total sky cover	3	11	9	0	2	2	8	7	0
Ceiling	0	8	14	1	2	2	9	6	0

Table 1. Continued.

Element	Birmingham Good	Fair	Needs Work	Other	Portland Good	Fair	Needs Work	No Reply
Temperature	1	7	4	1	0	1	4	7
Dew Point	0	6	6	1	0	1	4	7
Visibility	0	2	11	0	1	1	5	5
Weather	0	1	11	0	0	2	4	6
Wind direction/speed	0	5	7	1	0	1	4	7
Sea level pressure	3	5	2	3	0	1	4	7
1st cloud layer amount	0	4	9	0	1	1	5	5
1st cloud layer height	0	4	9	0	1	1	4	6
2nd cloud layer amount	1	7	4	0	1	1	4	6
2nd cloud layer height	1	6	6	0	1	1	4	6
Total sky cover	2	7	4	0	0	1	5	6
Ceiling	1	5	7	0	0	1	5	6

6. CONCLUSIONS

The MONITR procedure, which was used to automatically initiate the MONITR and GEM programs, was running a very high percentage of the time during the CLE, DSM, and BHM field tests, and most of the time during the PDX test. Severe weather outbreaks, AFOS crashes where the procedure was not restarted, and operations under degraded mode conditions were the primary reasons for the procedure not to be running.

The impact of the procedure and the programs on WSFO operations was very minimal. The effect of running an AFOS procedure at a console is well known and is not limited to this particular procedure. The programs, themselves, may have contributed to system instability and overloads during periods of heavy use since they, along with other applications programs, would be queued by AFOS for processing.

The MONITR program does not by itself provide all the information a forecaster needs to make a decision on amendment of the FT's. Forecasters at CLE, DSM, and BHM still rely on their FT/SAO log sheets as well as on other available information. The log sheets are still important to the forecasters because of the trend information available on them. PDX was the only test site which has eliminated the regular use of the logs.

One of the main benefits of MONITR was its ability to pick up special SAO's and alert the forecaster. This enabled the forecasters in several instances to issue FT amendments 20 to 30 minutes earlier than they might have otherwise.

The program also acted as a quality control for both the SAO and FT products. Typographical errors in the products were discovered by the decoders, and appropriate error messages were then included in the output display. In several instances, the forecasters requested that the observers issue corrected SAO's or, if the error was in the FT, they corrected it themselves.

The alert messages were designed to inform the forecaster when conditions were such that an amendment might be required in the near future but was not required now. Although the program did perform this function adequately, the forecasters generally did not amend the FT's based on the alert message. They did, however, feel that the information was useful. Several forecasters suggested that MONITR show more tolerance when the observed values are better than the forecast values. This is possible with the alert criteria and may help to cut down on the number of alerts issued by the program. On the other hand, since the amendment criteria are nationally defined (NWS, 1983d), more tolerance should not be allowed.

For the most part, MONITR performed well in the area of decoding and interpreting the SAO's and FT's. In those instances when MONITR was not able to decode a product, it did, as previously mentioned, alert the forecaster to the errors it found. However, there were also cases where MONITR successfully decoded a product but the value it then used was wrong because the wrong value had been entered by the forecaster or the observer. For example, an observer may have entered "C3 BKN" when it should have been "C30 BKN." This type of error can best be prevented by using caution when entering the data into the system.

Several program "bugs" were uncovered during the tests. The problems were related to the decoders not recognizing certain character strings properly. Both the SAO and the FT decoders have been modified and are now handling the data correctly. We have one remaining bug to correct in MONITR. This is the erroneous use of all information including and following the weather group in the FT as a remark when visibility is 7 miles or greater. This will be fixed in Version 2.0 of MONITR.

MONITR does not use the remarks portion of the FT's during the monitoring process. This was the major shortcoming of the program. As a result, the program operated on the conservative side by triggering the alarm and alert messages much more frequently than was necessary. It should be noted, however, that even without decoding and using the remarks, all test sites continued using MONITR following conclusion of the test.

The overall consensus of the forecasters was that MONITR was a useful tool for them. With ideas and recommendations received as a result of the tests incorporated into the program, MONITR should become a much more useful tool to the forecasters in the future.

The forecasters at all four forecast offices indicated that GEM followed persistence too closely and, therefore, was not as useful as it might otherwise be. Forecasters were particularly critical where GEM did not forecast a break in morning fog or an end to afternoon thunderstorms. When forecasters were asked whether GEM was better at forecasting the initiation or the end of an event, there were only four responses between the DSM and BHM tests. Three were for forecasting the end of the event.

There were only a few software problems noted which have since been corrected. Aside from the GEM forecasts themselves, the forecasters found very few errors in the program.

GEM requires approximately 1000 RDOS blocks of disk space to operate. To accommodate GEM, most forecast offices had to remove files from the disks. This was deemed unacceptable by these offices for long term use. These offices stopped using GEM following their respective tests.

Finally, the tests required extra work at the WSFO's, but much was learned about the operational use of these programs and also how to conduct tests in the future. There is no substitute for testing in an operational environment because of the many activities which compete for AFOS resources and the forecaster's time. This is particularly true during severe weather events.

The level of effort on the part of the WSFO's is reflected in the results presented in this report. Forecasters offered many valuable, constructive criticisms and suggestions. Our future program development activities will focus on the remarks received during these tests.

7. ACKNOWLEDGEMENTS

We want to thank Bob Krebs, Jerry Uecker, and Bob McLeod for their assistance in the design of the field test procedures, and to Jerry Uecker and Mike Tomlinson for initiating the contacts between TDL, the Regional Headquarters, and the WSFO's. Our thanks are also extended to the staff at each of the participating WSFO's, and especially to the Meteorologists-in-Charge (MIC's), the Deputy MIC's, and the AFOS Systems Managers. Finally, we thank Alice Baker and Judith Morrison for typing the manuscript.

REFERENCES

Herrmann, W. C., 1983: Generalized Exponential Markov (GEM) updating procedure for AFOS. NOAA Techniques Development Laboratory Computer Programs NWS TDL CP 83-5, National Weather Service, NOAA, U.S. Department of Commerce, 12 pp., (in press).

Miller, R. G., 1981: GEM: A statistical weather forecasting procedure. NOAA Technical Report NWS-28, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 103 pp.

National Weather Service, 1983a: AFOS monitoring and updating field test plan--April 1983. National Weather Service, NOAA, U.S. Department of Commerce, 33 pp. [Available from Techniques Development Laboratory, National Weather Service, Silver Spring, Md.]

_____, 1983b: AFOS monitoring and updating field test plan--June 1983. National Weather Service, NOAA, U.S. Department of Commerce, 35 pp. [Available from Techniques Development Laboratory, National Weather Service, Silver Spring, Md.]

_____, 1983c: AFOS monitoring and updating field test plan--September 1983. National Weather Service, NOAA, U.S. Department of Commerce, 33 pp. [Available from Techniques Development Laboratory, National Weather Service, Silver Spring, Md.]

_____, 1983d: Aviation terminal forecasts. National Weather Service Operations Manual, Chapter D-21, Manual Issuance 83-14, NOAA, U.S. Department of Commerce, 41 pp.

Vercelli, D. J., 1983: AFOS monitoring of terminal forecasts. NOAA Techniques Development Laboratory Computer Programs NWS TDL CP 83-4, National Weather Service, NOAA, U.S. Department of Commerce, 17 pp.