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

TDL OFFICE NOTE 83-13

RESULTS OF THE FIELD TEST OF THE AFOS-ERA FORECAST VERIFICATION PROGRAM

Mary M. Heffernan

August 1983

Table of Contents

RESULTS OF THE FIELD TEST OF THE AFOS-ERA FORECAST VERIFICATION PROGRAM

Mary M. Heffernan

1. INTRODUCTION

The Techniques Development Laboratory (TDL) with guidance from the Office of Meteorology (OM) and NWS Regional Headquarters has developed an AFOS-era forecast verification (AEV) program. The program is designed to provide automated collection and collation of forecasts and observations; a local, quality-controlled database; automatic transmission of data to a central site; a permanent central archive; and facilities to produce local, regional, and national summaries. Fig. 1 shows an overview of this program. Functions of the AEV program are performed at two levels. At the (local) Weather Service Forecast Office (WSFO) level, forecasts and observations are collected and collated, archived for local use, and transmitted to a central site for use in national summaries. Portions of this local processing require manual intervention. At the national level, the data sent from WSFO's are archived for use in the production of national and regional summaries. These functions are performed at the National Meteorological Center (NMC) with software provided by TDL.

The AEV program, in total, is a complex program which uses processing and computer capabilities of several NOAA systems. The field test which was conducted was not intended to test all functions of the AEV program. Specifically, the field test evaluated the local collection and collation software, the forecaster quality control and manual entry process, and the transmission of data to NMC. The functions of central archive and the production of regional and national summaries are being evaluated independently of this test. The design of the local, long-term archive and the production of local summaries can be adjusted to suit local needs. So, while versions of the software to perform the local archive and to produce local summaries were made available to the test sites, no efforts were made to evaluate these functions. See Dunn (1982) for a description of this software.

The field test of the AEV program had two goals. The first was to assess the impact of the AEV program on WSFO operations. This included not only forecaster workload but also local AFOS system resources. The second goal was to determine the reliability of transmitting the locally collected data from the WSFO to NMC.

The field test was conducted in two parts. In the first part, all 10 WSFO's participating in the test used the verification programs for a 1-month period. This portion of the test dealt primarily with the program being integrated into office operations. Based on the problems uncovered in this first portion and suggestions made by the test personnel, the software was revised and made available to the offices for further testing. In this second part of the test, two offices transmitted their verification data to NMC to test the transmission portion of the revised program.

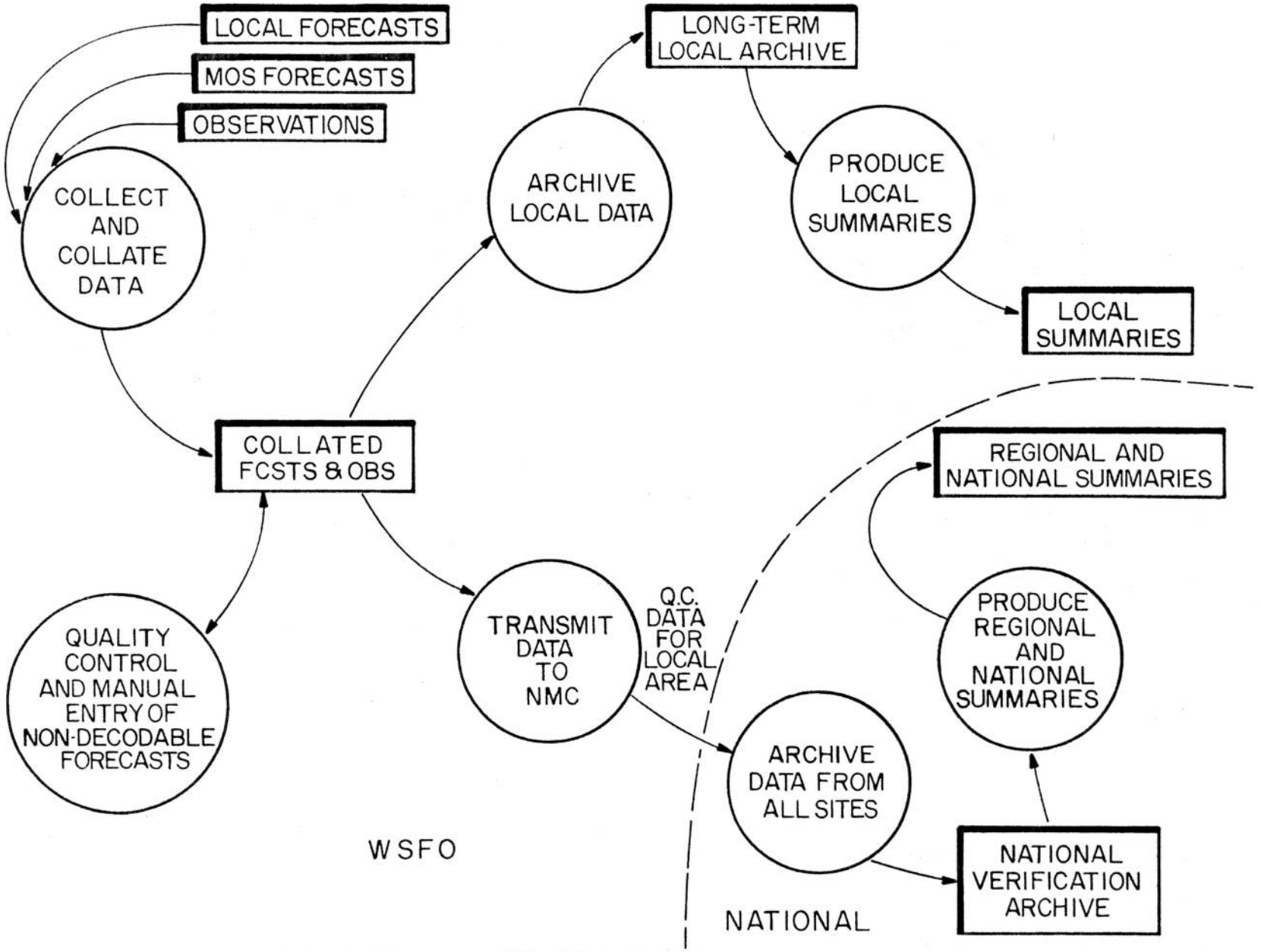

Figure 1. Overview of the AFOS-era verification program.

Overall, the field test was successful. The impact of the software on local operations is not judged to be significant. The amount of time required to run the verification system for two verification sites is about 10 minutes twice a day. With the new version of local software, no system impacts have been reported. The transmission of the verification data to NMC for the central archive was successful; of the 306 messages transmitted, 96 percent were received properly. Based on these test results, a decision was made by OM to implement the AFOS-era forecast verification program with minor modifications beginning October 1, 1983.

2. BACKGROUND

The Office of Meteorology operates two national forecast verification programs. The first, which verifies probability of precipitation (PoP) and maximum/minimum temperature forecasts, has been in existence for more than 10 years (NWS, 1982). It has been automated to the extent that local forecasts are collected on NOAA's IBM 4341's, Model Output Statistics (MOS) guidance forecasts are taken from TDL's archives, and observations are obtained from the National Climatic Data Center in Asheville, N.C.

The second program, which verifies forecasts of cloud amount, ceiling height, visibility, surface wind, and precipitation type, has been in existence for 10 years (NWS, 1973). Local forecasts are furnished on mark sense cards by each WSFO for two stations--itself and one other for which it is responsible. MOS guidance forecasts and verifying observations are obtained in the same way as they are in the PoP/temperature program.

Summaries are produced from both of these programs twice a year and disseminated to the WSFO's on microfiche about 6 months after the end of a season. Because of the absence of rapid feedback, forecasters and local management derive little benefit from these programs; consequently, the forecaster's interest is rather low.

The major advantages of AEV program over the existing verification programs are:

- o Rapid feedback locally.
- o On-site quality controlled data.
- o Automation of nearly all of the verification process.
- o Simplification of the collation process.

The AEV program will completely replace the existing two verification programs.

3. TEST DESCRIPTION

The field test of the AEV program was conducted between December 1982 and March 1983. It involved operations at NWS Headquarters, the NWS Regional Headquarters for the Eastern, Central, Southern, and Western Regions, and 10 WSFO's from these regions. The plan for the field test was submitted by OM to the AFOS Change Management Board as part of a request for change (NWS 016) to implement the AEV program into the AFOS system. Permission to conduct the test was granted and based on the results of the test a decision was made to implement the AEV program. The offices participating in the test were:

Cleveland, Topeka, Boston, Milwaukee, Raleigh, Ann Arbor, Memphis, Salt Lake City, Little Rock, and Boise.

The field test was conducted in two parts. In the first part, the 10 WSFO's used the programs daily in their operations. At the end of a 5-week period, the offices were asked to respond to the questionnaire shown in Fig. 2. Based on their responses and suggestions, the local software was modified and distributed to the offices again for testing. The first version of the software automatically transmitted the verification data to NMC for central archival. Because of problems described in Section 4 of this report, it was not possible to test the transmission portion of the program in the first part of this test.

In the second part of the test, two WSFO's received software which automatically transmitted the data. Detailed testing was conducted on the data transmitted by the two offices and on its archival. The remaining eight WSFO's received the same software with the transmission portion "turned off." At the end of a 4-week period, all 10 offices were asked to comment on this new version of the software.

---

1. Was the documentation provided for implementation of the software adequate? If not, how can it be improved?

2. Was the documentation for operating and maintaining the software adequate? If not, how can it be improved?

3. Was there disk space on your system to accommodate the software?

4. What times of day did you generally run the software?

5. What was the average runtime?

6. What amount of operator time was spent in running the software and quality controlling the data?

7. Were there problems encountered in editing the PVM and AVM products in message composition? If so, what were they?

8. Do you know or have reason to believe that the AEV application was associated with any system crashes at your site?

9. If any forecast cycles were missing for the period of the test, please give the reason(s).

10. In general, describe the impact of the AEV application on your site's operation.

Figure 2. Questionnaire used at WSFO's to evaluate the operation of the locally implemented software.

A. Software Description

In order to perform the functions of local collection and collation of forecasts and observations, of quality control and manual entry of non-decodable forecasts, and of transmission of data and receipt at NMC, software unique to the AEV program is required both at the WSFO's and at NMC. Additionally, the AEV program relies on the programmed capabilities of the AFOS system and the communications system at NMC. Only the interaction of this system software and the AEV program was considered; no attempt was made to validate the functions of existing system software. For example, no testing of the message composition capabilities of the AFOS system was undertaken; however, the functioning of that capability as needed by the AEV program was monitored.

Fig. 3 shows major components, database products, and data files of the locally implemented software. The first of two programs, MERGE, reads and decodes the forecast products and outputs a single product of the decoded forecasts. This product, RDG, can be edited by using the AFOS message composition capability to quality control the forecasts and to add the non-decodable forecasts. It is also possible to edit these same data further along in the process.

The second software program, COLLATE, accepts the output product, RDG, combines the forecasts with weather observations, and creates or updates a public verification matrix (PVM) product and an aviation verification matrix (AVM) product. These products contain, for the most recent 5 days, the collated sets of public and aviation forecasts and verifying observations. These data may be quality controlled with the AFOS message composition feature. Two forecast issuances (0940 GMT and 2140 GMT) per day are verified, so for 5 days there are 10 forecast cycles. The tenth forecast cycle is automatically transmitted to NMC the next time COLLATE is executed, which allows 5 days for quality control of the data. This information is also stored in a disk file for access by local archival routines. A more detailed description of this software is contained in Heffernan et al. (1983).

Fig. 4 shows the steps for archiving at NMC the data received from the WSFO's. Messages are generated at the WSFO's and transmitted on the Regional Distribution Circuits (RDC's). SMCC forwards these messages to the IBM 4341 computers. On the 4341's, the verification messages are routed with some other non-related messages to the IBM 360/195 computers where they are stored in one file. On the 360/195's, a program, STORAFO, selects the verification messages from the file and writes them to a verification file. STORAFO was written especially for the verification; it runs twice a day to collect new verification data that has been received. Weekly, a program, ARCHAFO, is run to write the data in the verification file to the central archive files which are stored on magnetic tape.

B. Local Office Operations

The tasks performed at the WSFO level during the field test included:

- Implement software into AFOS system.

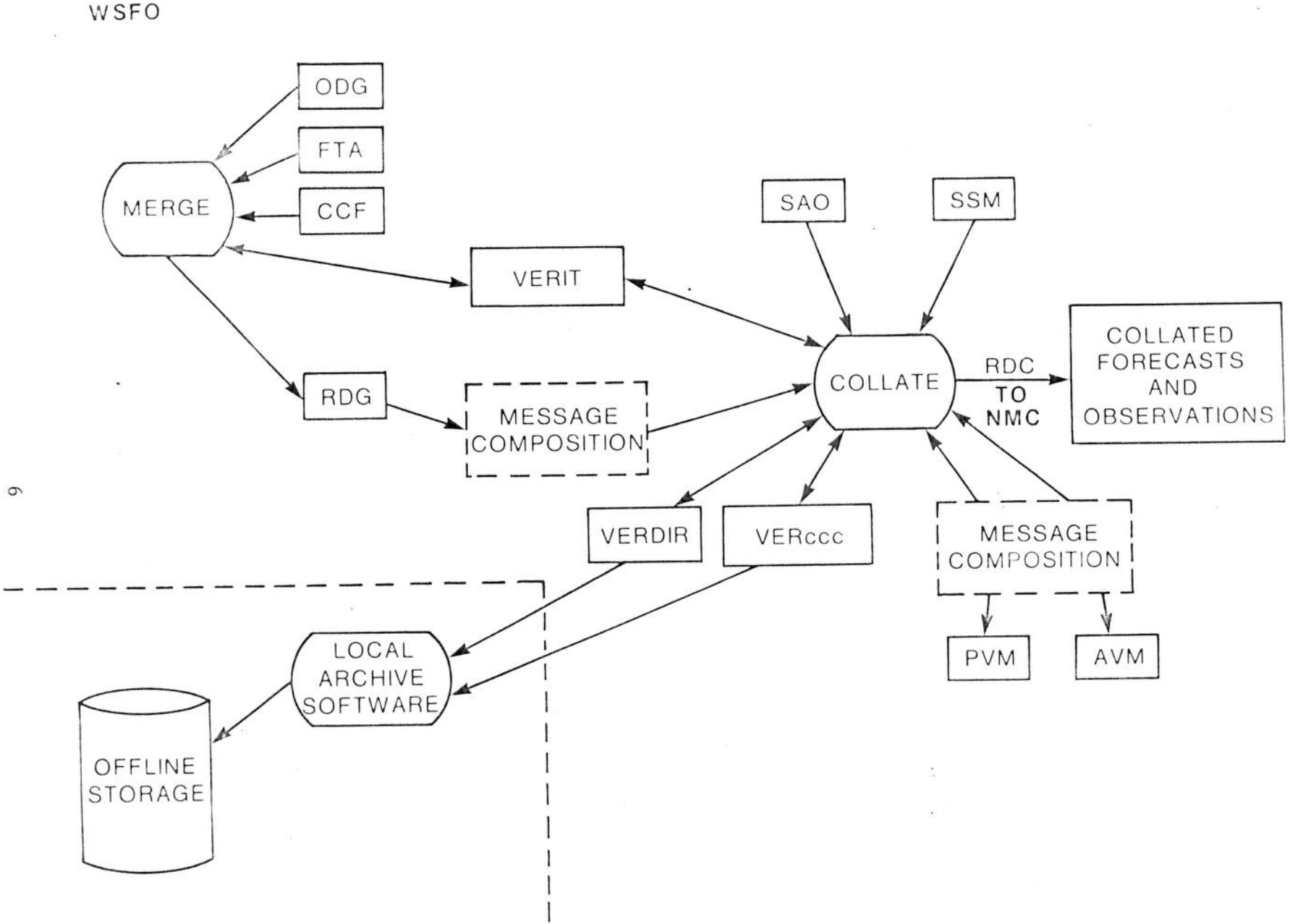

Figure 3. Design flow for the locally implemented verification software.

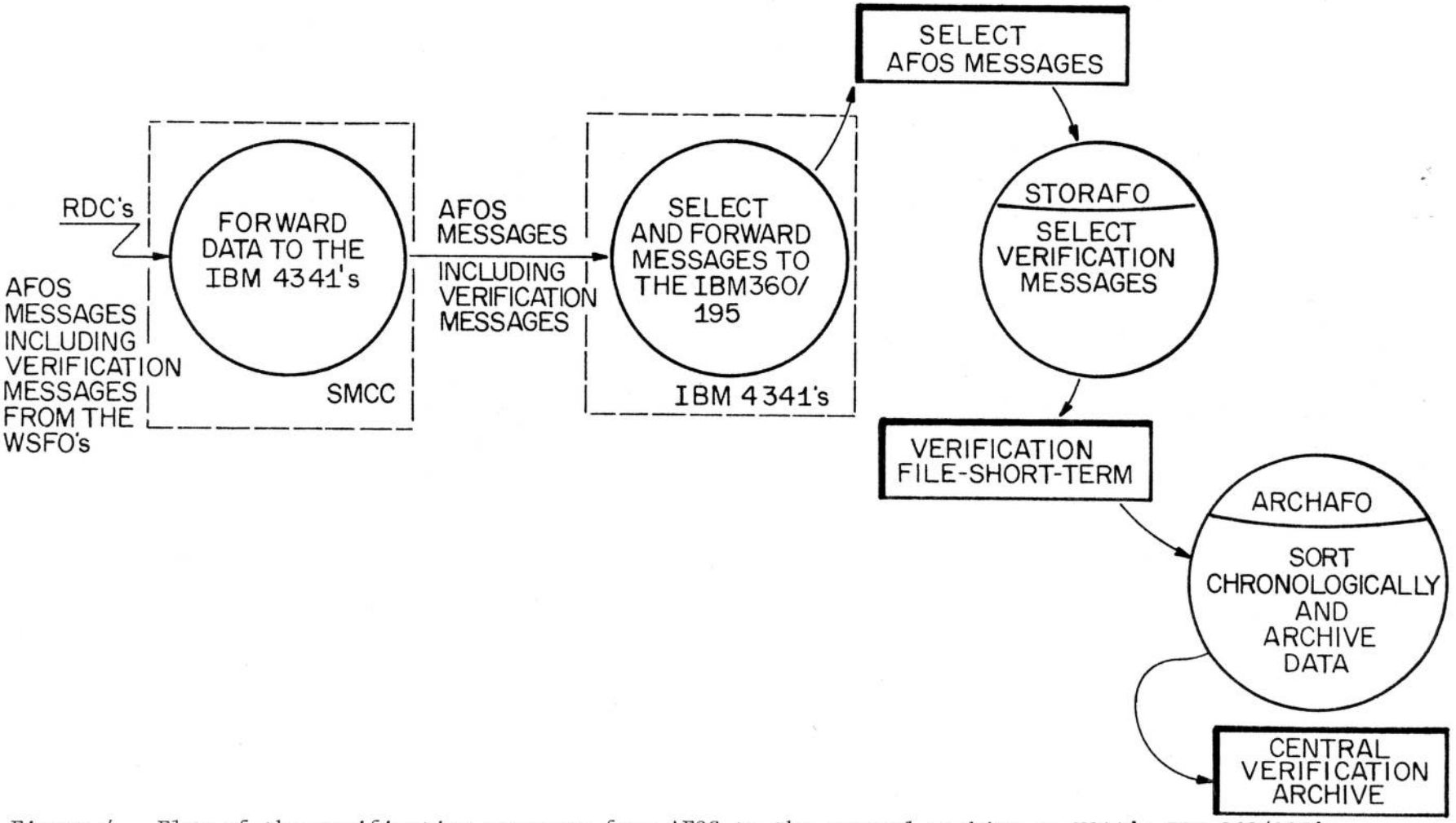

Figure 4. Flow of the verification messages from AFOS to the central archive on NOAA's IBM 360/195's.

- o Execute software twice a day.
- o Edit the PVM and AVM products to enter non-decodable forecasts and to quality control forecasts and observations.
- o Complete summary information about field test (see Fig. 2).

Also, personnel were encouraged to report any errors or problems encountered during the test period. The Meteorologist-in-Charge (MIC) assigned personnel to perform these tasks. Generally, it was the forecasters working on shift who executed the software and edited the matrices.

C. Regional Operations

The Scientific Services Division (SSD) of each region was responsible for accepting the software from Headquarters personnel and delivering it to the test sites in its region. The SSD's also coordinated the correspondence between the WSFO and national levels. Each region had the option to run the software on its Headquarters system for familiarity.

D. National Operations

TDL was responsible for designing, coordinating, and reporting on the test. Personnel were available during normal working hours to answer questions and provide guidance. At NMC, TDL personnel monitored the incoming data by using the IBM 360/195 software.

4. TEST RESULTS

The two goals of the field testing of the AEV program were to assess the impact of the program on local office operations and to determine the reliability of transmitting data from the WSFO's to NMC. The following subsections describe the results of this testing. The office evaluations of the field test can be found in Heffernan (1983).

A. Local Office Operations

To determine the impact of the AEV program on local office operations, five areas were examined during the testing. They were:

- o Local Software
- o Documentation
- o System Resources
- o User Interaction
- o System Interface

The questions the offices were asked to respond to focused on these areas.

Local Software

The first few weeks of testing uncovered several errors in the local software. In addition, suggestions were made to make the programs more flexible and less susceptible to operator error. Based on experience in the first part of the test, TDL produced a new version of the local software.

The changes made in this new version can be placed into three categories. They are (1) improvement of the interface between the user and the program, (2) enhancement of the decoders for database products such as terminal forecasts and surface observations, and (3) correction of errors. In the first category changes included:

- o Alerting the AFOS console upon completion.
- o Prohibiting an attempt to process data into the future.
- o Chronological sorting and matching of data in the PVM and AVM.
- o Capability to collate new observations on demand.
- o Addition of a switch to enter aviation forecaster number.
- o Defaulting certain non-decodable forecasts to their most frequent value.

A detailed list of changes made to the software is given in Appendix I.

There were two recommendations made for improving the local software that were not addressed with the new version produced. They were:

- o to take into account corrected observations and forecasts and
- o to use hourly observations for max/min temperature and precipitation amount when synoptic observations are not available.

Implementing each of these would save time for the user. In the case of observations and forecasts that are corrected, the responsibility fell to the user to quality control the data. For WSFO's that wish to verify stations that do not take synoptic observations, the job of entering the temperature and precipitation observations was left to the station personnel. Both of these enhancements were made to the AEV software prior to national implementation.

Documentation

Documentation to install and run the local software was provided with the software. In general, the users felt this was good, although cases of program malfunction could be handled better. An example would be where the software hung and the reason (unknown to the operator) was a bad database product. The improvements in this area would include some experienced advice for tracking down problems. The documentation was subsequently expanded based on these comments.

Specific instructions for the user interfacing with the program need to be developed. This would include such things as when to enter the non-decodable forecasts, the time to run the software etc. Because of differences in office practices, these are best specified locally. An example developed in the Western Region is given in Appendix II.

System Resources

The offices participating in the test had sufficient system resources to support the AEV program. The program requires 350 RDOS blocks, 28K of core, and less than 3 minutes of clock time to process data for two stations.

Offices participating in the test did not attempt to run it while in degraded mode.

User Interaction

The primary task of the site personnel was to quality control the data and enter the non-decodable forecasts. The software was executed by a procedure that required minimal effort. However, the editing of the PVM and AVM products was time consuming. To process data for two stations, the total time required was between 10 and 15 minutes for both the public and aviation forecasts. The editing step required paging through a five-page preformatted product for both public and aviation data for each station to be verified. In the second phase of the test, a simpler procedure to edit the matrices was discovered which allowed paging only as far into the five pages as the operator needed to go. Thus for most instances, only a one or two page product had to be edited. The time savings was about 5 minutes.

Offices chose different methods of editing the products. Some offices assigned one or two people to run the software and edit the matrices. In this case, the local forecaster usually entered his or her non-decodable forecasts on a sheet of paper to be used by the person doing the editing. This person then ran both forecast cycles together at the same time each day. On the other hand, some offices assigned the responsibility for running the software and editing the matrices to the public and aviation forecasters on duty. In this case, the software was run twice a day--generally around 1100 and 2300 GMT. Two sites in the Central Region--Ann Arbor and Topeka--did not enter the non-decodable forecasts.

A frequent suggestion received from the local offices was to find a way to eliminate the message composition step. This requirement is somewhat relieved by reducing the amount of time required to edit the matrices. However, it should be a goal for a verification program to automatically retrieve and decode all forecasts needed for verification.

The amount of user time required by the AEV program during the test was strongly influenced by the problems described under the local software and documentation subsections. During the first part of the test, when the software did not function properly and the documentation did not adequately support the user, a great deal of time was lost by the user trying to figure out what was wrong. Every office reported a significant decrease in time required for the AEV program with the new version of software.

System Interface

Three areas where there is a potential system impact were identified. These are message composition, database, and the operating system. The majority of the offices (7 out of 10) reported no problems at all. One office, Boston, reported system crashes and database problems associated with the AEV program, but subsequently found an incorrect PILEDIT had occurred on the database. After correcting this problem and implementing the new version of software, they experienced no system impact. One Central Region office reported a "trap" occurred after running the software three consecutive times. This happened only once and could not be reproduced in the AFOS Experimental

Facility (AEF). The third test office that reported system problems was Little Rock. They reported . . . "a couple of occasions when running the verification programs caused an AFOS crash. Most of these occurred during the first couple of weeks and may have been the result of the staff's unfamiliarity with the program." They noted "a time or two" when editing the PVM's and AVM's during a busy period caused a "console hang or crash." They also reported that end of file remarks began to appear more often resulting in the need to run MODIFY more often. They have "no direct correlation (with the AEV program) . . . but the symptoms appear coincidental with the software's installation." All of these problems were with the initial version of the local software. With the new version, they reported stable operations. This may be due to both the new software and increased local experience.

Both Boston and Little Rock noted that the editing of the preformatted PVM and AVM caused the system response time to increase--the system slowed down. Little Rock cited occurrences of PVM's or AVM's being garbled after a message composition step or being lost after COLLATE produced them. The garbling is caused by paging back and forth in message composition. This was demonstrated in the AEF. In the case of the lost PVM's and AVM's, the files are produced by COLLATE, but the call to FSTORE to store them in the database is not successful. The new version of software handles this problem.

B. Transmission Testing

In the first few days of testing, a problem was uncovered in the transmission of the data product from the WSFO's to NMC. The product was encoded in binary form to save characters on the network. It was determined that the message transmission might be cut short or lost because of the particular binary value a datum might assume. A decision was made to proceed with the test with the binary message format but to redesign the format to ASCII and conduct a seperate more limited test at two offices after completion of the first part of the test.

WSFO's Cleveland and Milwaukee received the new software which included the ASCII message transmission. A test was conducted between these offices and NMC for a 6-week period. The message receipt was 96% (291 out of 306). One problem in the decoding software on the IBM 360/195 was uncovered and corrected during the test period. By the end of the test period, another problem in reading the verification file on the IBM 360/195 was noted and was being worked on.

5. CONCLUSIONS

The field test of the AEV program was successful. While problems were encountered during the first part of the test, the second part has shown no adverse impact on local offices operations. Transmission of the data to the central archive was highly reliable.

The local software has been improved such that few problems were encountered in the second part of the test. The suggestions were made to enhance the software to (1) use hourly observations when synoptics are not available and (2) handle corrected observations. The documentation, while adequate, should be improved based on experience gained from the test.

The test results showed there were adequate system resources to support the program, and the operation of the program did not adversely affect the AFOS system. While functions that AFOS provides, such as editing preformatted products and storing products from a background program, did not work perfectly, the reliability of these functions was acceptable.

The amount of time required to support the AEV program locally is about 10 minutes twice a day. While that is more time than required for the present mark-sense card system, the benefits of having a system that affords rapid feedback should heavily outweigh the time factor involved.

The transmission of the verification data and its storage at a central site was highly reliable for the test period.

6. REFERENCES

Dunn, L. B., 1982: AFOS-Era automated forecast verification. Preprints Ninth Conference on Weather Forecasting and Analysis, Seattle, Amer. Meteor. Soc., 35-37.

Heffernan, M. M., 1983: Field test of the AFOS-era verification program. Techniques Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce, 50 pp.

_____, M. C. Newton, and R. L. Miller, 1983: AFOS-Era forecast verification program, NOAA Techniques Development Laboratory Computer Programs TDL CP 83-3, National Weather Service, NOAA, U.S. Department of Commerce, (in press).

NWS, 1982: Public forecast verification. National Weather Service Operations Manual, Chapter C-71, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 8 pp.

NWS, 1973: Combined aviation/public weather forecast verification. National Weather Service Operations Manual, Chapter C-73, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 14 pp.

APPENDIX I

Changes and Enhancements Made to the Local Software
Based on Part 1 of the Test

GENERAL

AVM's and PVM's can be edited when less than full. The date/time fields of the unused cycles must be blanked out.

Corrected display of precip type observed fields. Freezing and liquid categories were reversed.

MOS PoPs are rounded to the nearest ten percent. The 2 and 5 percent are preserved.

MERGE and COLLATE now alert the ADM.

CCCVERCCC encoded in ASCII.

All decoders recognize "BAD VERSIONS" of products.

MERGE won't allow processing into the future. The 00Z (12Z) cycle can't be processed until 930Z (2130Z, 2230Z MTN/PAC). The override (/0) switch will process any cycle.

COLLATE sorts cycles of data and insures PVM and AVM have the same number of cycles. Duplicate cycles are discarded; the last version of the cycle entered is the one retained.

Debugger removed from both COLLATE and MERGE.

MERGE modified to accept a CCF from 9Z to 11Z and from 21Z to 23Z.

Correct the mishandling of negative temperatures from the ODG or CCF.

COLLATE now knows there is only 30 days in November.

COLLATE now processes whatever cycle MERGE has done. It previously checked the date/time of the cycle and if it was not new it did not process it.

Addition of a /C switch on COLLATE which only collates new observations; it does not add a new cycle of forecasts to PVM or AVM and it does not transmit data.

The override (/0) switch on MERGE is now in the form MMDDCC where MM is the month (1 or 2 digits), DD is the day and CC is the cycle time 00 or 12Z.

Unlimited ceiling heights are coded as 97 rather than 99.

Aviation forecaster number can be entered with a switch (/A) on the MERGE program. Number entered is truncated to the rightmost 2 digits.

Station name added on PVM and AVM product.

Local wind forecast at 42h and local forecast of snow amount 12-24 h are defaulted to zero.

SAO

Correct error numbers displayed.

Thin cloud amounts are now properly encoded as category 1.

Variable ceiling is now recognized as a valid ceiling.

Warren Sunkel's decoder is used.

FT

Correct error numbers displayed.

Recognizes frontal passages--previously a forecast group with a frontal passage was not used.

Number of forecast groups allowed increased to 10.

Corrected error of picking up remarks as the body of the forecast. This error caused several problems in decoding FT's in the past.

Handles whole and franctional visibilities with a space between them.

Corrected amended FT's are now recognized. They are not used for verification, however, this change should reduce the number of error messages printed.

Synoptic Code

Corrected errors in rounding precip amount.

Modified min temperatures archived from

Previous 24h min reported at 0600 GMT
Previous 18h min reported at 0000 GMT

to

Previous 18h min reported at 1800 GMT
Previous 12h min reported at 1200 GMT

Zeroes are now displayed in the observed field of the PVM when no snow occurred.

APPENDIX II

Example Instructions for Operation of the Local Software

1. Change purge parameters on CCF, SSM, SAO, FTA products and add data base products with appropriate preformat fields only according to TDL CP 82-3. Table 4.

2. Set up and run VERCREATE.

3. Run MERGE, check with D:RDGXXX to see if it ran correctly. Where XXX is the station ID.

4. Run COLLATE, check with D:PVMXXX and D:AVMXXX where XXX is the station ID of the stations being verified. Use the current date and 'ØØ' in the /O switch to make sure the first cycle run is OOZ.

5. If Step 4 is a success set up the following:

```
--A procedure--
    RUN:MERGE
    PAUSE 10
    RUN:COLLATE
    RUN:MERGE
    PAUSE 10
    RUN:COLLATE
    PAUSE 240
    RUN:@VERBU@
    END
```

   where VERBU is an indirect file that will maintain backup verification data files on the archive floopy. This will allow for continuous verification without data loss if backup disks are loaded.

   --- VERBU should be as follows:

   MOVE/A/R DP3 VERIT VERDIR VERCCC where CCC is your station ID

6. Run this procedure once a day between 10Z-14Z. It will process the 12Z cycle and then the OOZ cycle. If this time period is missed, or even an entire day forgotten, then run the procedure until the latest OOZ cycle has been done. The limitation on this is the number of versions of forecasts and observations that are in the data base.

7. The PVM and AVM products should be checked for quality, with missing information edited into the appropriate location. This includes missing forecasts and observations due to holes in the data base or garbled products, as well as the undecodeable information such as pcpn type, cloud amount, and aviation forecast number.