NOAA Technical Memorandum ERL FSL-19

A SECOND EVALUATION OF STATE-OF-THE-ATMOSPHERE VARIABLES GENERATED BY THE MAPS SURFACE ASSIMILATION SYSTEM

J.L. Mahoney
P.A. Miller
J.E. Ramer
T.L. Smith
M.M. Cairns
R.J. Miller

Forecast Systems Laboratory
Boulder, Colorado
April 1996

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

NOAA Technical Memorandum ERL FSL-19

A SECOND EVALUATION OF STATE-OF-THE-ATMOSPHERE VARIABLES GENERATED BY THE MAPS SURFACE ASSIMILATION SYSTEM

Jennifer Luppens Mahoney
Patricia A. Miller
James E. Ramer
Tracy Lorraine Smith
Mary M. Cairns
Ronald J. Miller

Forecast Systems Laboratory
Boulder, Colorado
April 1996

UNITED STATES DEPARTMENT OF COMMERCE
Michael Kantor
Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
D. JAMES BAKER
Under Secretary for Oceans and Atmosphere/Administrator

Environmental Research Laboratories
James L. Rasmussen
Director

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA Environmental Research Laboratories. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

For sale by the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 20061

CONTENTS

FIGURES

TABLES

A SECOND EVALUATION OF STATE-OF-THE-ATMOSPHERE VARIABLES GENERATED BY MSAS

Jennifer Luppens Mahoney, Patricia A. Miller, James E. Ramer,
Tracy Lorraine Smith, Mary M. Cairns, and Ronald J. Miller

ABSTRACT. This paper describes the second evaluation (E2) of state-of-the-atmosphere variables (SAVs) derived from the Mesoscale Analysis and Prediction System (MAPS) Surface Assimilation System (MSAS). This is part of an ongoing program at the Forecast Systems Laboratory (FSL), sponsored by the Federal Aviation Administration's Aviation Weather Research Program (AWRP). The evaluation was conducted from 22 February through 10 March 1992 using a verification dataset obtained from the Stormscale Operational and Research Meteorology-Fronts Experiment Systems Test (STORM-FEST). For verification purposes, MSAS produced analyses at specific surface station locations for 3 March through 10 March 1992. The results from MSAS are compared to surface analyses from MAPS.

In general, better results were obtained from the MSAS system than for MAPS for the surface fields such as altimeter, temperature, and dewpoint. The analyses of the surface winds were roughly the same for MSAS and MAPS. Two important differences between MSAS and MAPS analysis techniques may explain these results. The number of surface observations ingested into the MAPS system is roughly a third of all available observations. The majority of these missing observations occur over the western United States. MSAS, on the other hand, ingests nearly all surface observations available at the time of its analysis, including those over the western U.S. Therefore, the MSAS surface analyses more closely follow the observations in areas where MAPS is not using these observations. The inclusion or exclusion of the surface observations is tied to the difference between the actual surface pressure and the model pressure defined by the topography. The envelope topography used in MAPS produces a higher topography field that rides over the true terrain. This topography field is needed to sustain the predictive part of MAPS. MSAS, on the other hand, uses a minimum topography-minimum elevation method. These MSAS techniques were developed to extract consistent information from surface data collected in mountainous regions and take into account the physical blocking and channeling induced by the moutainous terrain, as well as, distinguish between different air masses (Miller and Benjamin 1992). Future improvements are expected for both systems.

1. INTRODUCTION

FSL, part of the National Oceanic and Atmospheric Administration, is involved in an FAA-sponsored project aimed at developing gridded analyses and forecasts of SAVs and aviation-impact-variables (AIVs) from numerical models and analysis systems (Kraus 1993). Most evaluations of numerical models and data assimilation systems are concerned with SAVs such as temperature, moisture, and winds. Although these variables are the basis for weather forecasting, it is often left to humans to interpret the output of SAVs and develop analyses and forecasts of AIVs such as ceilings, visibility, and precipitation type. The Aviation Division of FSL has created a Verification Program to evaluate the accuracy of the gridded systems and to assist in the development of the algorithms used to generate SAVs and AIVs. In this evaluation, only the SAVs were produced by MSAS.

Two verification exercises have been conducted, Exercise 1 (E1) and Exercise 2 (E2). E1 occurred between 1-10 April 1991 and is described in Cairns (1992) and Cairns et al. (1993). The datasets and verification methods are discussed in Miller and Cairns (1993). Exercise 1 provided a baseline from which to build E2 which occurred 22 February-10 March 1992. The E2 domain covered portions of the STORM-FEST domain.

To date, four systems have been evaluated: the Mesoscale Analysis and Prediction System (MAPS) (recently transferred to the National Centers for Environmental Prediction (NCEP), as the Rapid Update Cycle (RUC)), the Local Analysis and Prediction System (LAPS), MSAS (transfered to NCEP as the RUC surface analysis system, RUCS), and the NCEP's Eta model. The MSAS analyses of SAVs and a comparison between MSAS and MAPS are presented in this document. See Mahoney et al. (1995) for the results from the LAPS analysis system and Cairns et al. (1994a and 1994b) for results from the MAPS and Eta models.

2. SYSTEM DESCRIPTION

MSAS is an assimilation system built to exploit the surface data by providing timely and detailed surface analyses. The system is part of MAPS, which provides real-time guidance for forecasters in local nowcasting environments. MSAS provides analyses of nine variables on a 1-h cycle and uses persistence (the previous hourly analysis) as the background for the current analysis.

Orographic influences often complicate the analysis of surface data. MAPS surface analysis techniques, used in MSAS, were developed with an emphasis on extracting consistent information from surface data collected in mountainous regions. The analysis method was chosen for its ability to handle varying data density and is unique in its use of elevation and potential temperature differences to help model the spatial correlation of the surface observations. The resulting horizontal correlation functions take into account physical blocking and channeling by moutainous terrain, as well as distinguishing between different air masses. MSAS is described in further detail by Miller and Benjamin (1992) and Benjamin and Miller (1990).

This 60-km horizontal resolution version of MSAS evaluated for E2 is the operational version as of November 1993; it covers 48 contiguous states and neighboring areas of Mexico and Canada.

3. VERIFICATION DATA

Although E2 occurred from 22 February through 10 March 1992, this study covers only the period when the MSAS system was run, 3 through 10 March 1992. The main impetus for choosing this period was to utilize the enhanced observation datasets which were available during STORM-FEST. Four primary observation data sources were selected for the verification data: SAO, upper-air (UPA) rawinsonde, vertical wind profiler (PRF), and pilot reports (PIREPS). This study focuses on the surface observations only, since MSAS is a surface data assimilation system. For information on how the other fields, such as clouds, were used for verification, see Mahoney et al. (1995). Most of the raw observation data were obtained from the STORM-FEST field project office and those observations unavailable from STORM-FEST were obtained from FSL's Facility Division. Table 1 lists the SAVs verified in the evaluation.

Table 1: SAVs Evaluated by MSAS in E2

VARIABLE	UNITS
Altimeter	in. of mercury
Temperature	^oF
Dewpoint	^oF
Wind Speed	kt
Wind Direction	degrees

3.1. Surface Data

This section describes the verification data collection and quality control. For E2, 319 SAO stations, which manually recorded observations hourly, were selected for the surface verification dataset. Thirty-one of the 319 sites were not collected by STORM-FEST but were obtained from the FSL data archives. No special observations were used in the verification, only hourly observations. In addition to the manual SAOs, STORM-FEST collected automated surface observations within the STORM-FEST domain from AWOS, ASOS, and other automated networks (FSL Mesonet, NCAR PAM, High Plains Network, Illinois State Water Survey). With the exception of the ASOS sites, the automated observations were also used in the verification.

The SAO data obtained from STORM-FEST were already quality controlled. This was done by comparing the observations to analyses from MSAS (Miller and Benjamin 1992). Disagreements between the observations and analyses were flagged according to Table 2. In addition, a manual quality control check was performed on selected observations, which verified and/or adjusted data with unlikely and questionable quality flags. For this verification, only "Good" or "Questionable" data were used, and it should be noted that only the individual value was declared to be bad, not the entire observation.

For the SAO data obtained from FSL, two quality control checks were performed on the observations. The first was a simple check to make sure values were within acceptable ranges (e.g., -50° F < Temp < +130° F). For the second check, the time-series of the data were plotted and subjectively evaluated to detect any obvious errors. If a value of an observed element failed either check, it was declared as missing and no attempt was made to correct or reconstruct the value. As with the STORM-FEST data, only the value was declared to be bad not the entire observation.

Table 2: Quality Control Thresholds for STORM-FEST Surface Observations

PARAMETER	UNLIKELY	QUESTIONABLE
Station Pressure	> 10.00 mb	> 3.00 mb
Sea Level Pressure	> 10.00 mb	> 3.00 mb
Calculated Sea Level Pressure	> 20.00 mb	> 6.00 mb
Dry Bulb Temperature	> 8.00 °C	> 5.00 °C
Dewpoint Temperature	> 8.00 °C	> 5.00 °C
Wind Speed	> 20.00 ms^{-1}	> 5.00 ms^{-1}
Wind Direction	------------	> 90.00 deg

3.2. Instrument Precision

In any verification study, the raw observations that are being used as *truth* are assumed to be correct, after even minimum quality control has been applied. Although routine measurements of the atmospheric variables are often considered to be exact, there are known inaccuracies. Unfortunately, it is often difficult, if not impossible, to provide a standard against which a measurement can be verified. For example, if a rawinsonde reports a 500-mb temperature of -25°C, there is no other exact measurement by which we can judge the accuracy of the rawinsonde report. So, instrument precision is often described as the root mean square (RMS) difference between two identical collocated instruments. For example, two rawinsondes are attached to the same balloon, and the differences in their measurements are computed. Table 3 lists the surface and instrument RMS differences as defined by Hoehne (1980), NWS (1991a,b) and NOAA (1991). These numbers should be kept in mind when evaluating the relative accuracy of numerical models and data assimilation systems. Even though a perfect analysis or forecast is the ultimate goal, it can hardly be assumed that an analysis or forecast error will be less than the precision of the verifying instrument. In addition, errors of representativeness (Daley 1991) due to the spacing of the observational data can affect the results.

Table 3: Instrument Root Mean Square Error (RMSE) Differences

SURFACE VARIABLE	THRESHOLD
Altimeter	0.02 in. Hg
Temperature	1^{o} C (1.8^{o} F)
Dewpoint	from 1.1^{o} F for Td > 32^{o} F & T-Td < 11^{o} F to 7.9^{o} F for Td < 32^{o} F & T-Td > 54^{o} F
Wind Speed	2 kt or 5% (whichever is greater)
Wind Direction	5^{o} for wind speed > 5kt

4. MSAS ANALYSES

For verification purposes, MSAS analyses were rerun with the 1993 version of the system and interpolated to the observation locations described in Section 3. The MSAS output was formated to reflect the observational format. The SAVs directly available from the MSAS grids are altimeter, potential temperature, dewpoint, and the *u* and *v* components of the wind. A surface temperature was derived for each station from MSAS interpolated potential temperature and altimeter, and the actual station elevation.

The verification area was defined by approximately 25-50 ^{o}N and 65-130 ^{o}W. The MSAS system contains the area defined by 22.84 ^{o}N, 120.5 ^{o}W (lower left corner) to 45.99 ^{o}N, 60.83 ^{o}W (upper right corner). Within the verification area, MSAS was required to interpolate point analyses for the surface locations only. For this verification exercise, MSAS was run on an hourly cycle from 1200 UTC 3 March 1992 to 0000 UTC 10 March 1992 and produced hourly analyses. The MSAS system was evaluated on its surface analyses of altimeter, temperature, dewpoint and winds.

5. RESULTS

MSAS was designed to provide timely, detailed, and coherent analyses of surface data, even when the data are collected in rough terrain where station elevations differ widely and observations are often subject to local effects (Miller and Benjamin 1992). To measure the difference that these techniques make in the analysis of the surface SAVs, the results from MSAS are loosely compared to those produced by MAPS. Note that this comparison can not be truly a one-to-onecomparison since MSAS was run on an hourly cycle and MAPS on a 3-h cycle. Therefore over a 7-day period, MSAS produced roughly 165 analyses, while MAPS produced only 55.

Appendix A describes statistical measures used for this verification, while Appendix B lists summary statistics for MSAS and MAPS. The following subsections further discuss details of the MSAS evaluation. The reader is encouraged to refer to Appendix B, as each section references the statistics listed there.

5.1. Altimeter.

The altimeter errors are very small for both MSAS and MAPS (Appendix B). The correlation coefficients are close to 1 in both cases with only minor scatter evident (Figs. 1 and 2). The variability in the altimeter is shown by the box plots of altimeter error (Figs. 3 and 4). Note, however, the scales in Figs. 3 and 4 are not directly comparable. The plots show that the range of error between the 10th and 90th percentiles is .04 in. for an observed altimeter of 30.00 - 30.30 in. for MSAS and .12 in. for the same altimeter for MAPS. Geographically (Figs. 5 and 6), MSAS and MAPS show a slight positive bias over Arizona, Nebraska, and Missouri. In addition, MAPS tends to slightly underanalyze the altimeter over the northern tier of states, the Pacific Northwest states, and New England.

Overall, the RMSE differences for MSAS are 0.019 in. and 0.029 in. for MAPS. These differences between MSAS and MAPS are attributed to number of surface observations that get ingested into each system. For instance, only a third of the surface observations make it into the MAPS system, while MSAS uses all available surface observations to perform its analysis.

5.2. Temperature.

MSAS analyzes the surface temperature slightly too warm, while MAPS analyzes these temperatures too cool, as indicated by the biases (Appendix B). It is apparent from the scatterplots (Figs. 7 and 8) that the MSAS system more closely matches the observed temperatures that are greater than 30 ^{o}F than does MAPS. Thus, the MAPS cool temperature bias is the most prominant for observed temperatures greater than 30 ^{o}F. This variability in temperature is captured in the box plots of temperature error (Figs. 9 and 10). The box plots indicate that the MSAS temperature errors generally are smaller than errors for MAPS. The variability in the temperature errors for MSAS is also less than for MAPS. These temperature errors in the MAPS model are confined mainly to the Intermountain West (Fig. 11). Overall, the RMS error for MSAS is 2.22 ^{o}F and for MAPS it is 3.91 ^{o}F. The RMSE error for MSAS is only slightly above the instrument error, as shown in Table 3.

This cool bias and temperature error in the MAPS model is explained by the absence of surface data particularly over the western states. The inclusion or exclusion of surface temperatures into MAPS is closely tied to the differences between the actual surface pressure

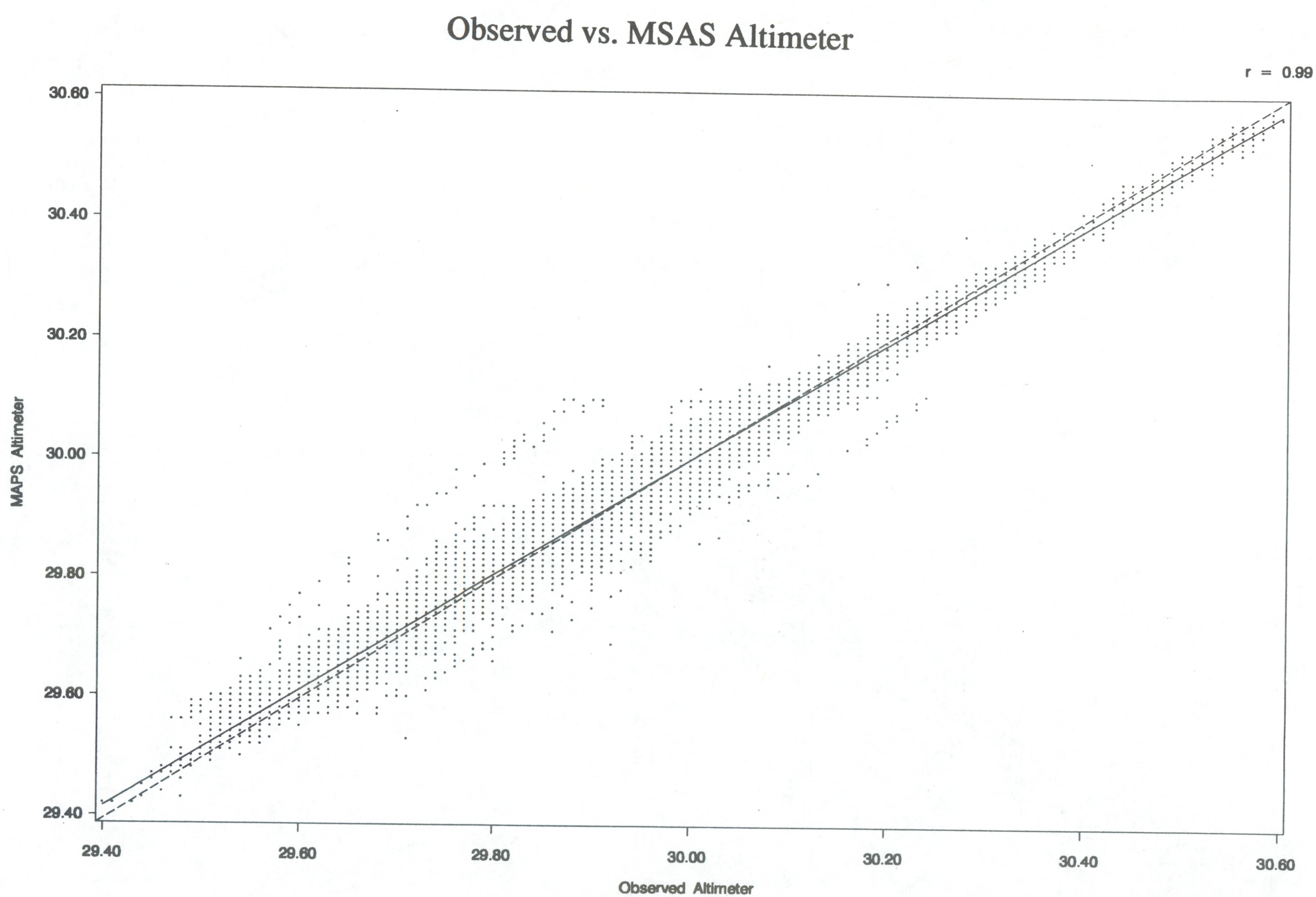


Figure 1. Scatterplot of observed altimeter (in) versus MSAS analysis.

MAPS NATL MSAS USING TD
03MAR92:12:00 – 10MAR92:09:00, Analysis

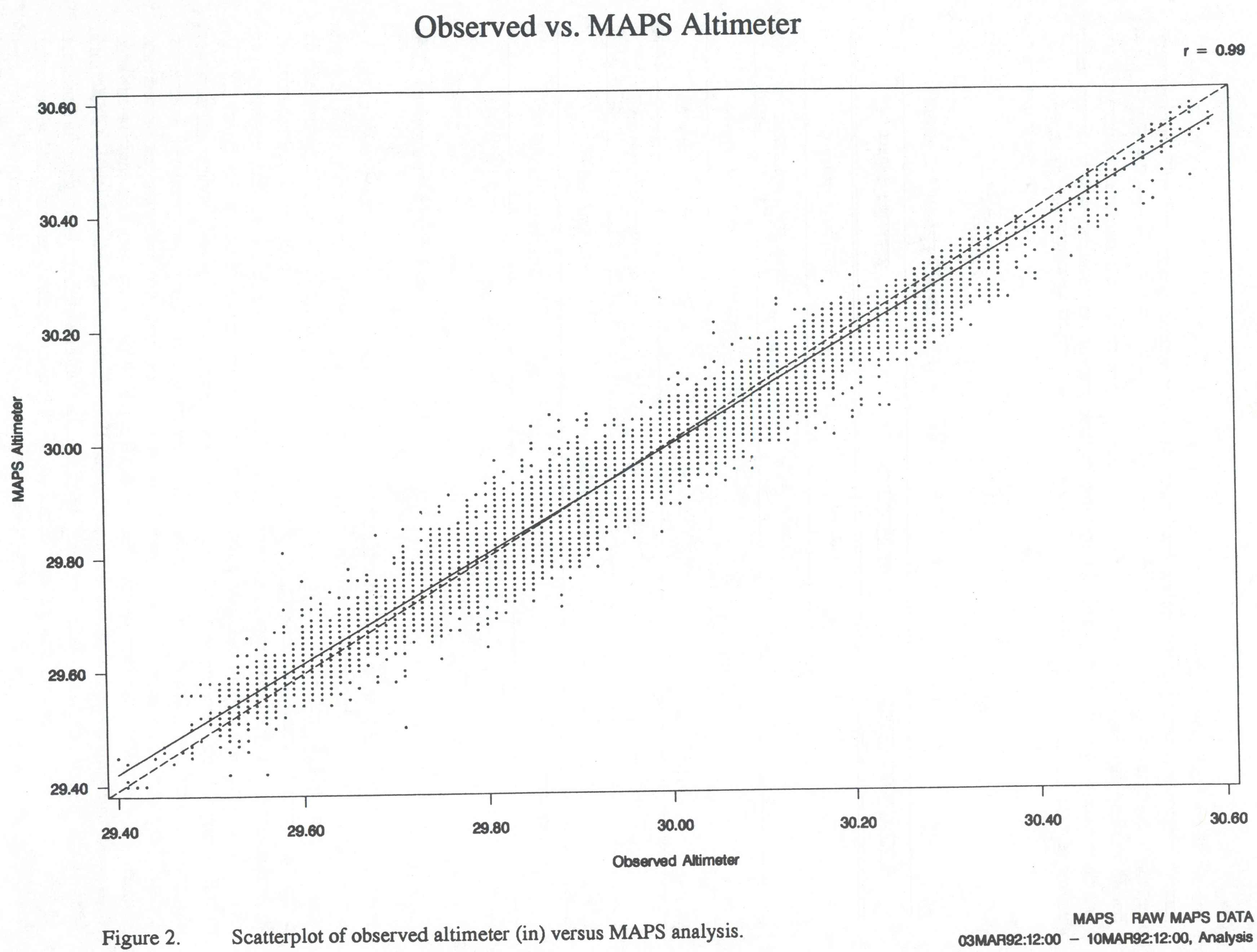


Figure 2. Scatterplot of observed altimeter (in) versus MAPS analysis.

Box plots of MSAS Altimeter Errors

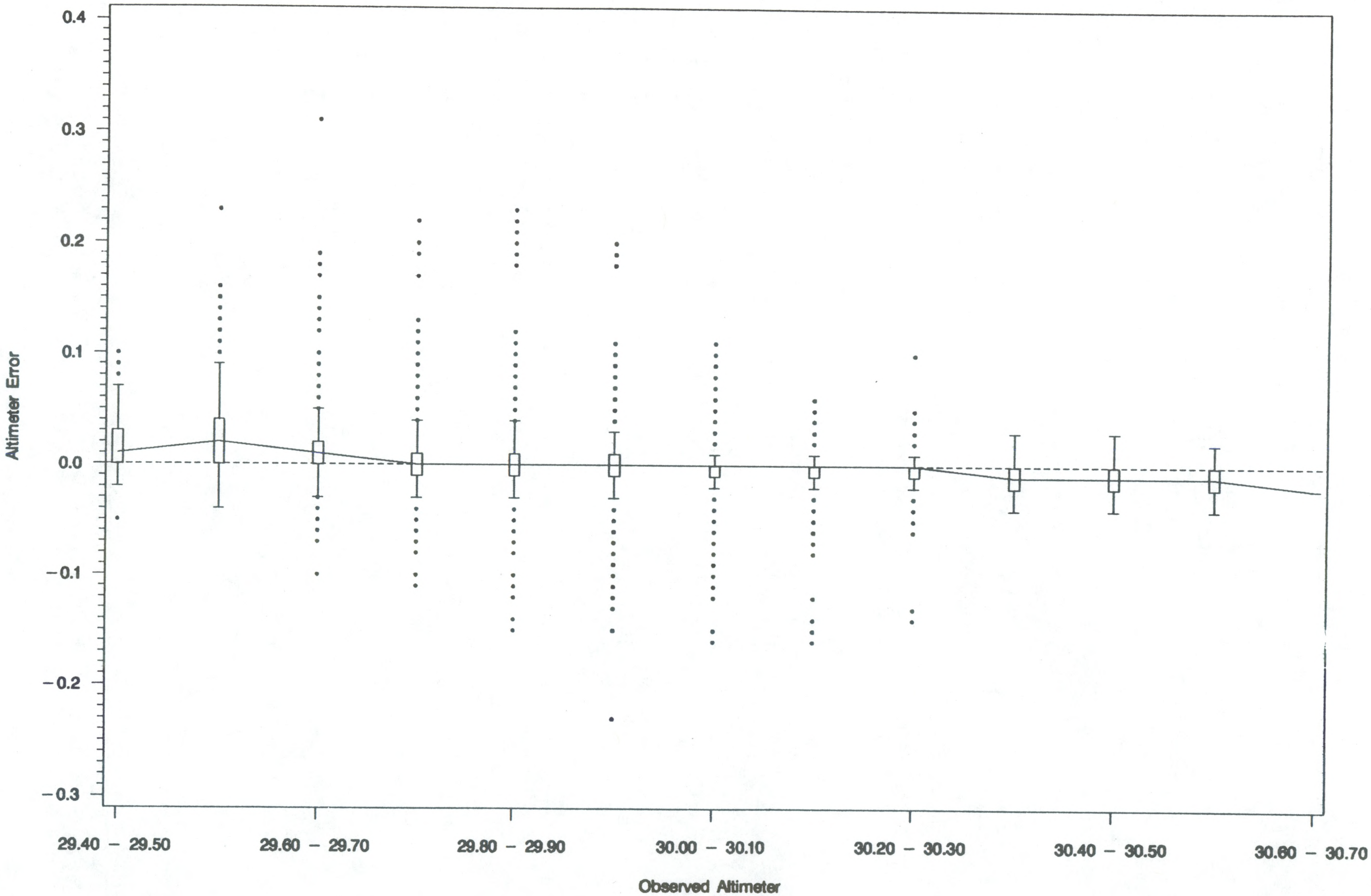


Figure 3. Box plots of MSAS altimeter (in) errors.

MAPS NATL MSAS USING TD
03MAR92:12:00 − 10MAR92:09:00, Analysis

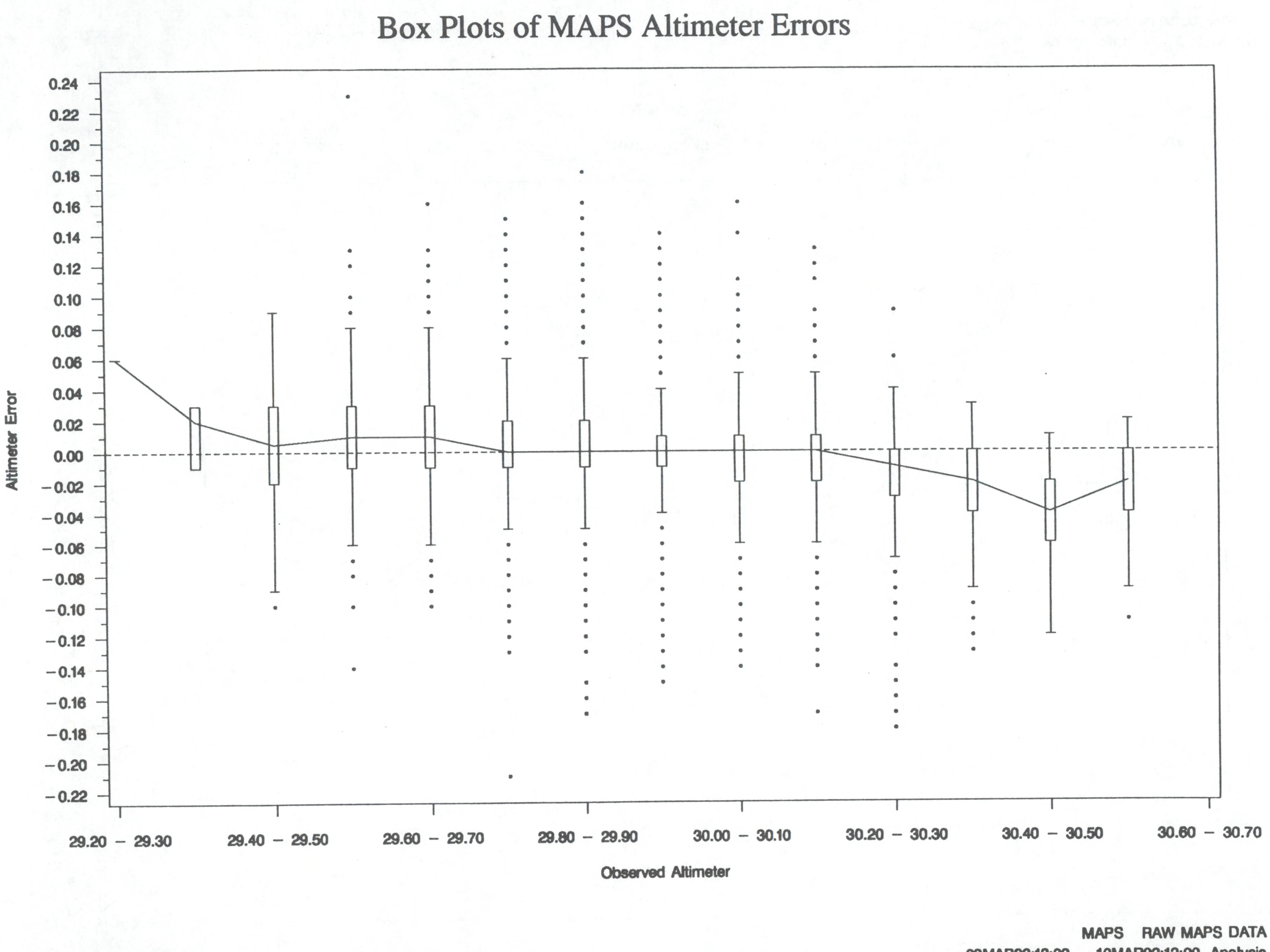


Figure 4. Box plots of MAPS altimeter (in.) errors.

MSAS Altimeter Bias

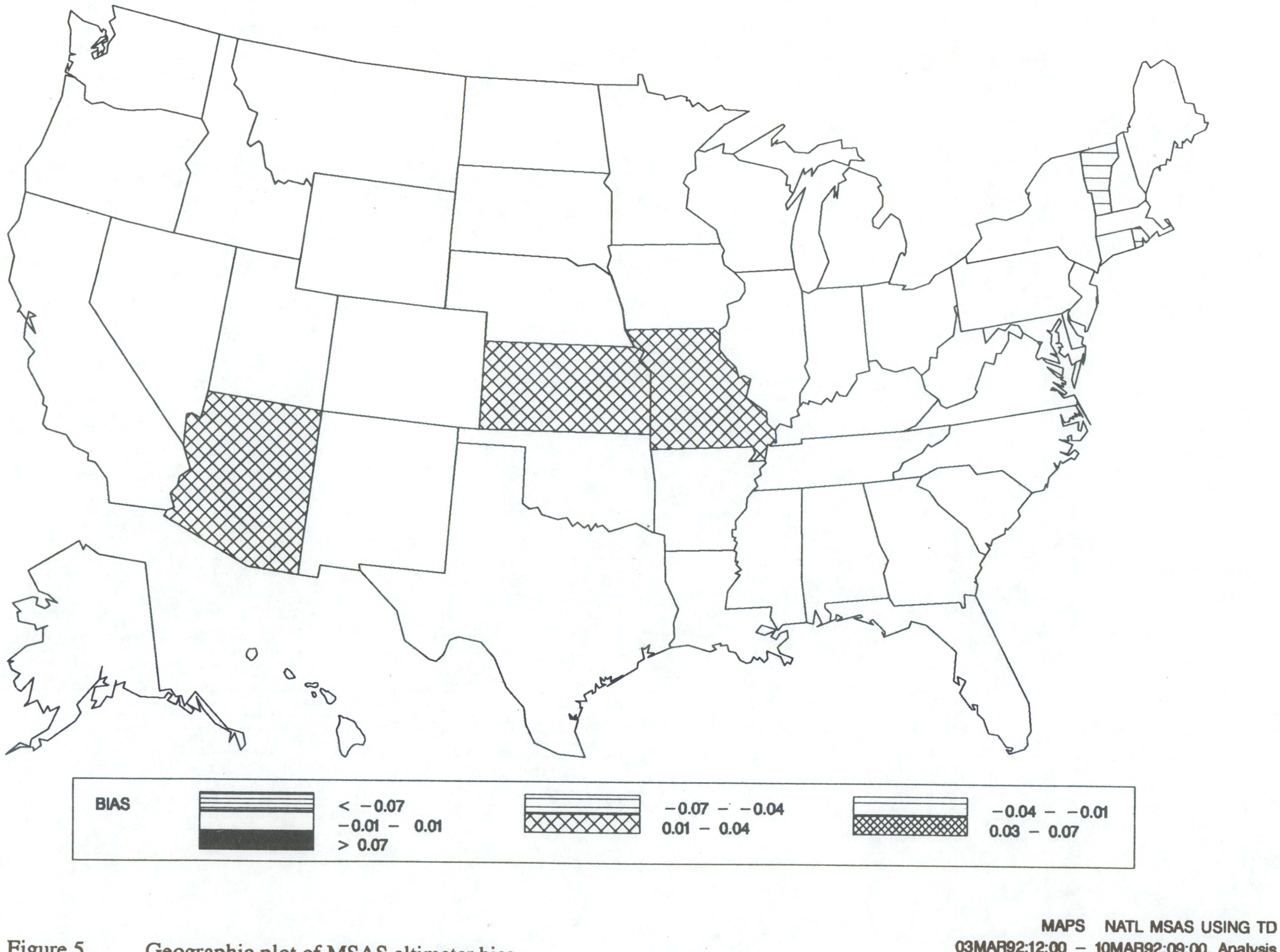


Figure 5. Geographic plot of MSAS altimeter bias.

MAPS Altimeter Bias

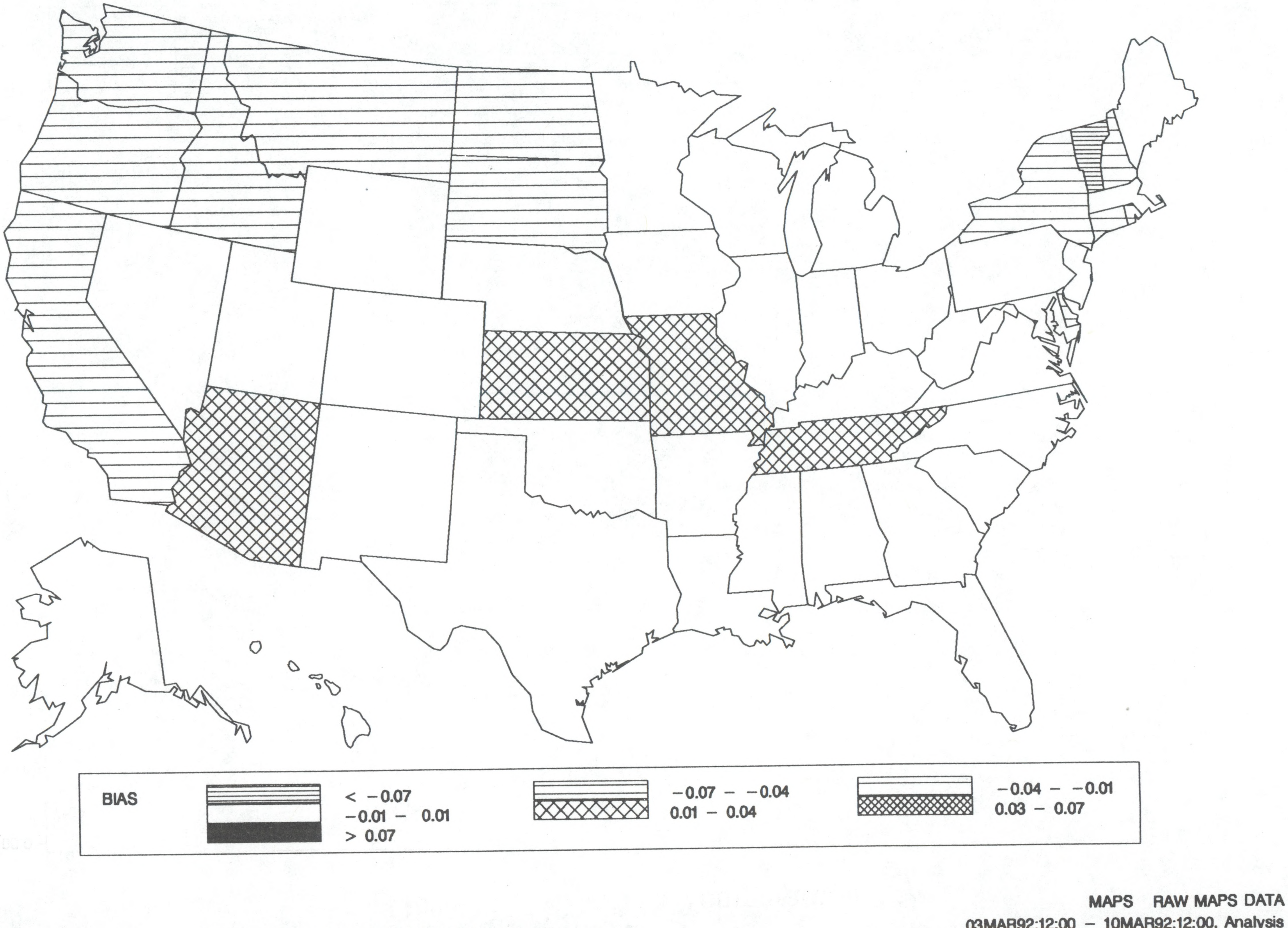


Figure 6. Geographic plot of MAPS altimeter bias.

Observed vs. MSAS Temperatures

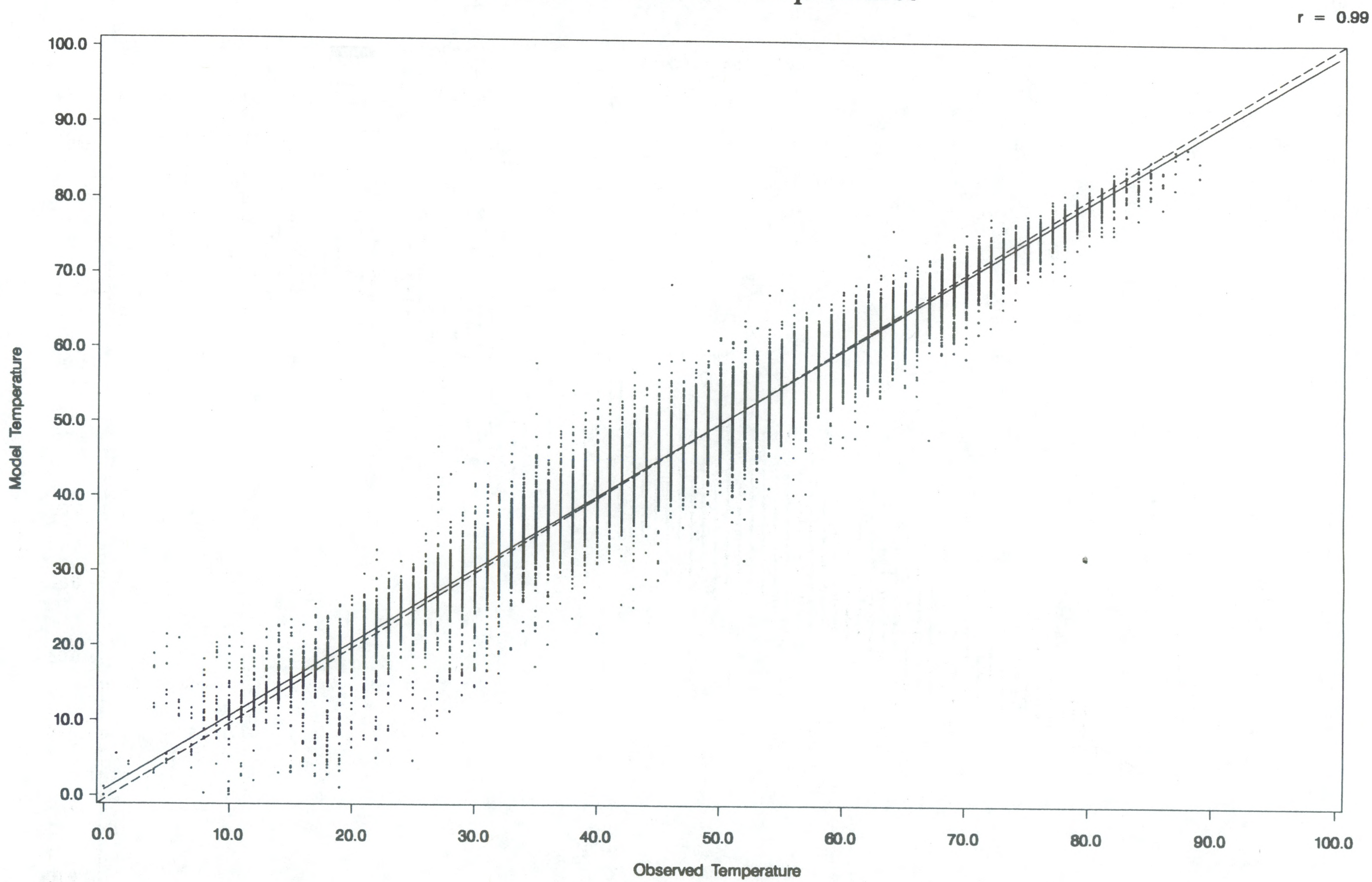


Figure 7. Scatterplot of observed temperature (F) versus MSAS analysis.

MAPS NATL MSAS USING TD
03MAR92:12:00 – 10MAR92:09:00, Analysis

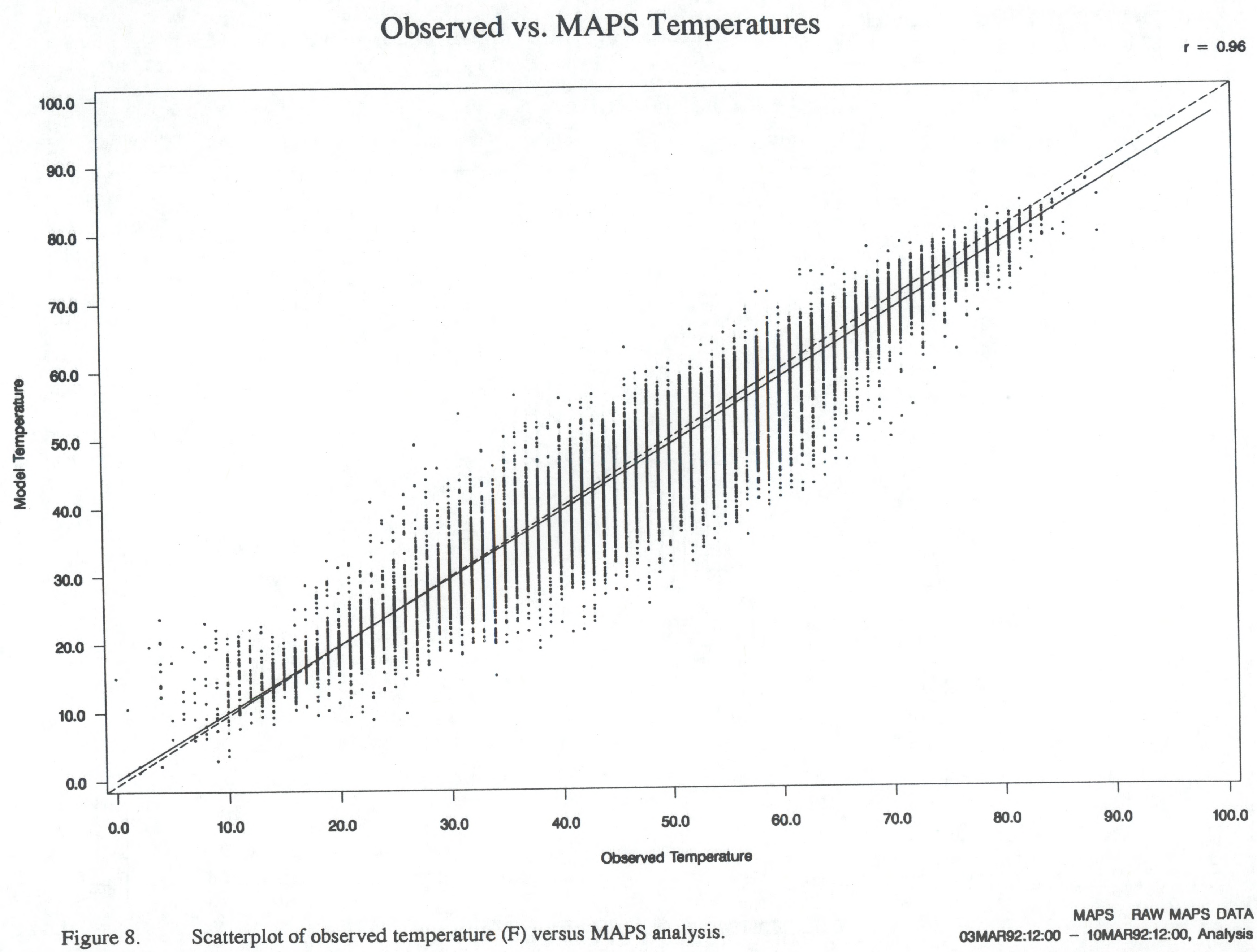
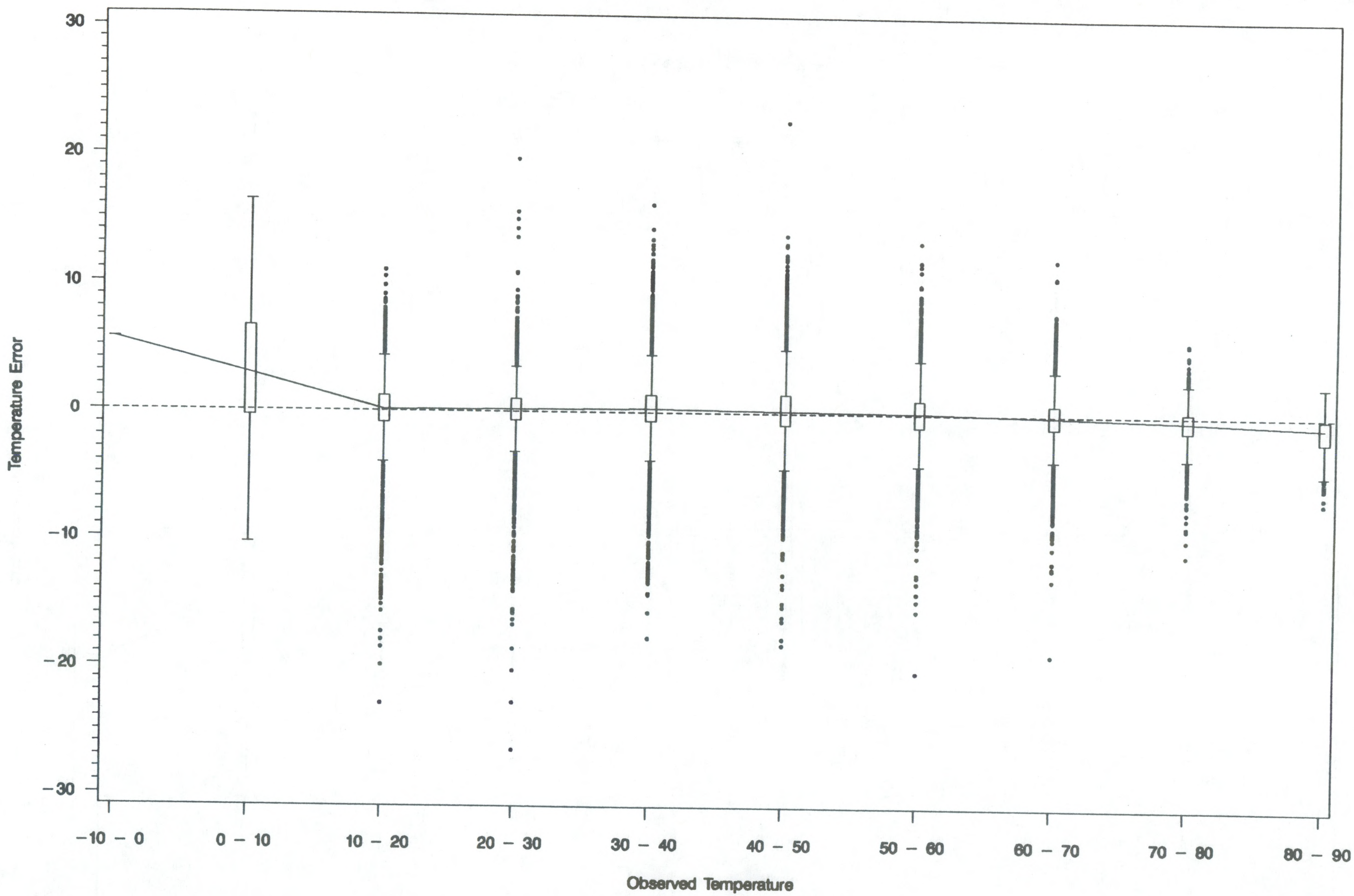


Figure 8. Scatterplot of observed temperature (F) versus MAPS analysis.

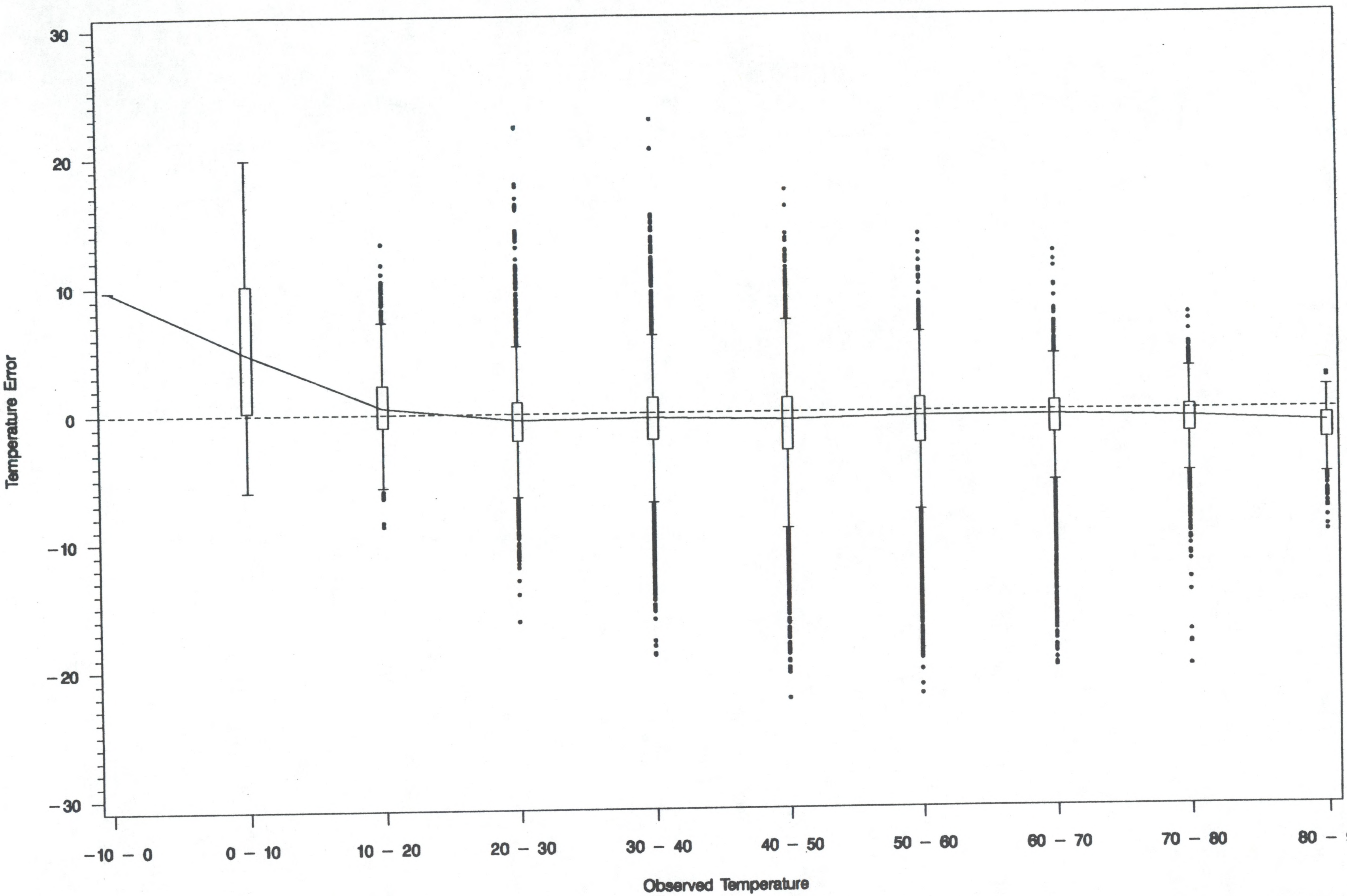
Box plots of MSAS Temperature Errors



MAPS NATL MSAS USING TD
03MAR92:12:00 − 10MAR92:09:00, Analysis

Figure 9. Box plots of MSAS temperature errors.

Box Plots of MAPS Temperature Errors



MAPS RAW MAPS DATA
03MAR92:12:00 – 10MAR92:12:00, Analysis

Figure 10. Box plots of MAPS temperature errors.

MAPS Temperature Bias

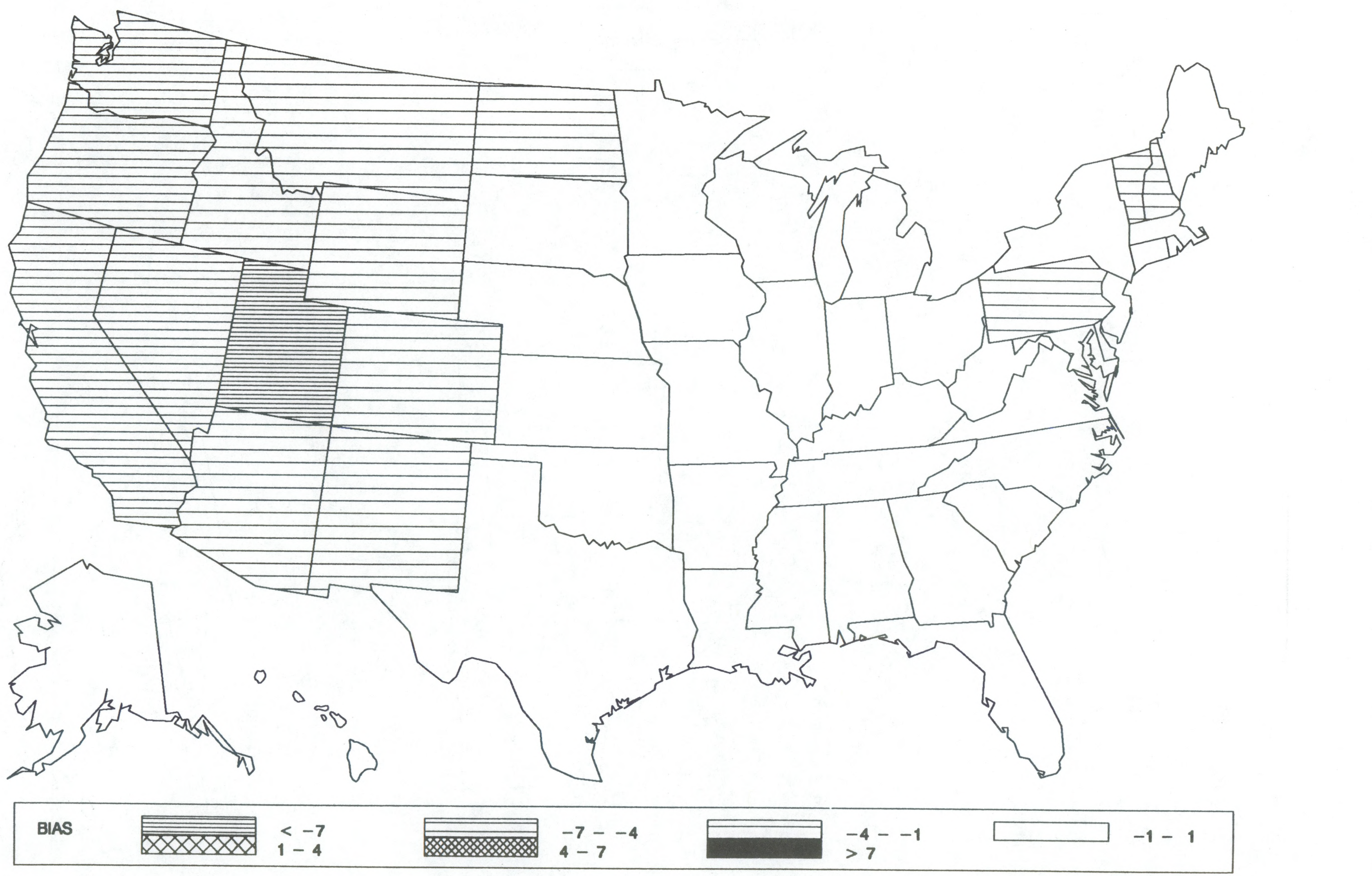


Figure 11. Geographic plot of MAPS temperature bias.

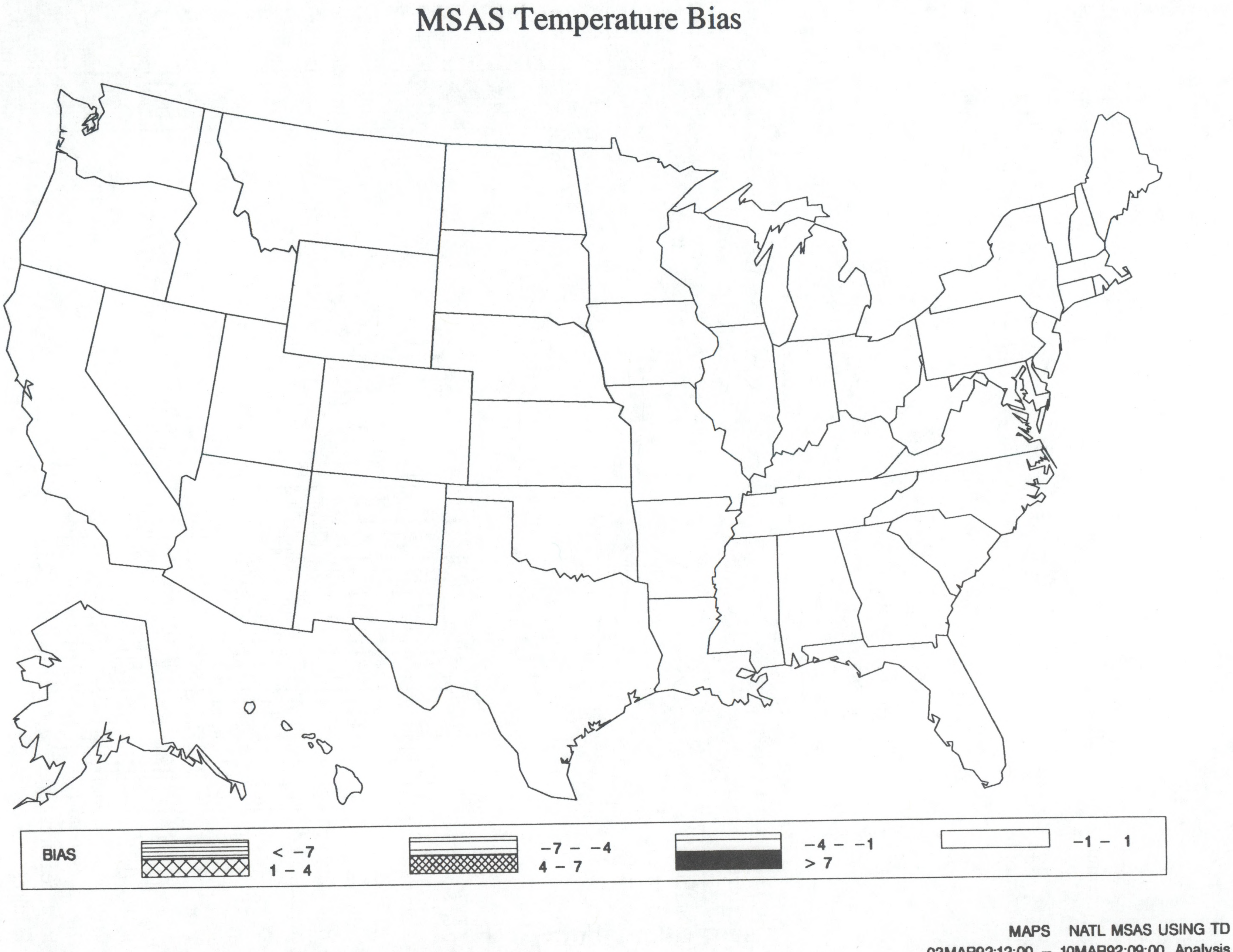


Figure 12. Geographic plot of MSAS temperature bias.

and the model pressure which is defined by the envelope topography used by MAPS. This envelope topography field produces a higher model topography that rides over the true terrain. This method of depicting the terrain is needed to sustain the predictive part of MAPS. MSAS, on the other hand, is able to more closely match the temperatures over the Intermoutain West (Fig. 12) where the differences between MSAS and MAPS are greatest, since the system incorporates nearly all available surface temperatures into its analysis techniques. These techniques were developed with an emphasis on extracting consistent information from surface data collected in mountainous regions. MSAS uses a minimum topography-minimum elevation method to represent the minimum elevation at a grid point. These analysis techniques take into account physical blocking and channeling by mountainous terrain (Miller and Benjamin 1992).

5.3. Moisture.

The dewpoint errors are similar in size to the temperature errors, even though dewpoints are analyzed directly and temperatures are derived. Scatterplots of dewpoint (Figs. 13 and 14) indicate that MSAS analyzes dewpoints that are too moist for dewpoints less than 45 ^{o}F, while MAPS dewpoints are too dry for all observed dewpoints. The dry bias in MAPS is addressed in Cairns et al. (1994a). Geographically, MSAS analyzes a near 0 bias over most of the country, expect for South Dakota and Arkansas where the bias is between 1 and 4 ^{o}F (Fig. 15). The dry bias in MAPS is evident mainly over the western U.S. (Fig. 16). Overall, the RMS error for MSAS is 2.60 ^{o}F and for MAPS its nearly double. Again, the differences lie in the number of surface observations that get ingested into the MAPS system as compared to MSAS.

A slight dry bias in the relative humidity field was evident in MSAS and MAPS (Appendix B). The warm surface temperature bias evident in MSAS may in part be the cause of the negative relative humidity bias; warmer temperatures result in lower relative humidity. The distribution of relative humidity errors for MSAS and MAPS are shown by the box plots (Figs. 17 and 18). The relative humidity error values exhibit increasing variability with increasing relative humidity for both MSAS and MAPS. The variability in the MAPS surface relative humidity is larger than MSAS, possibly due to the absence of many surface observations in the model. Overall, the RMSE errors for MSAS are 7.55 % and 9.94 % for MAPS.

5.4. Wind.

The surface wind speed errors are roughly the same for MSAS and MAPS. These type of errors were also evident for Eta (Cairns et al. 1994b) and LAPS (Mahoney et al. 1995). As shown by the box plots (Figs. 19 and 20), the winds are underanalyzed for wind speeds greater than 5 kts with the largest errors occurring at the greatest wind speeds. Strong winds tend to be local features, thus the 60 km gridded data of MSAS and MAPS are not expected to catch these features. Geographically, MAPS has a lower bias when analyzing the speed of the wind than MSAS (Figs. 21 and 22). Errors in wind direction for speeds greater than 10 kts (Figs. 23 and 24) indicate that both MSAS and MAPS have difficulty analyzing the direction over the

Observed vs. MSAS Dewpoint

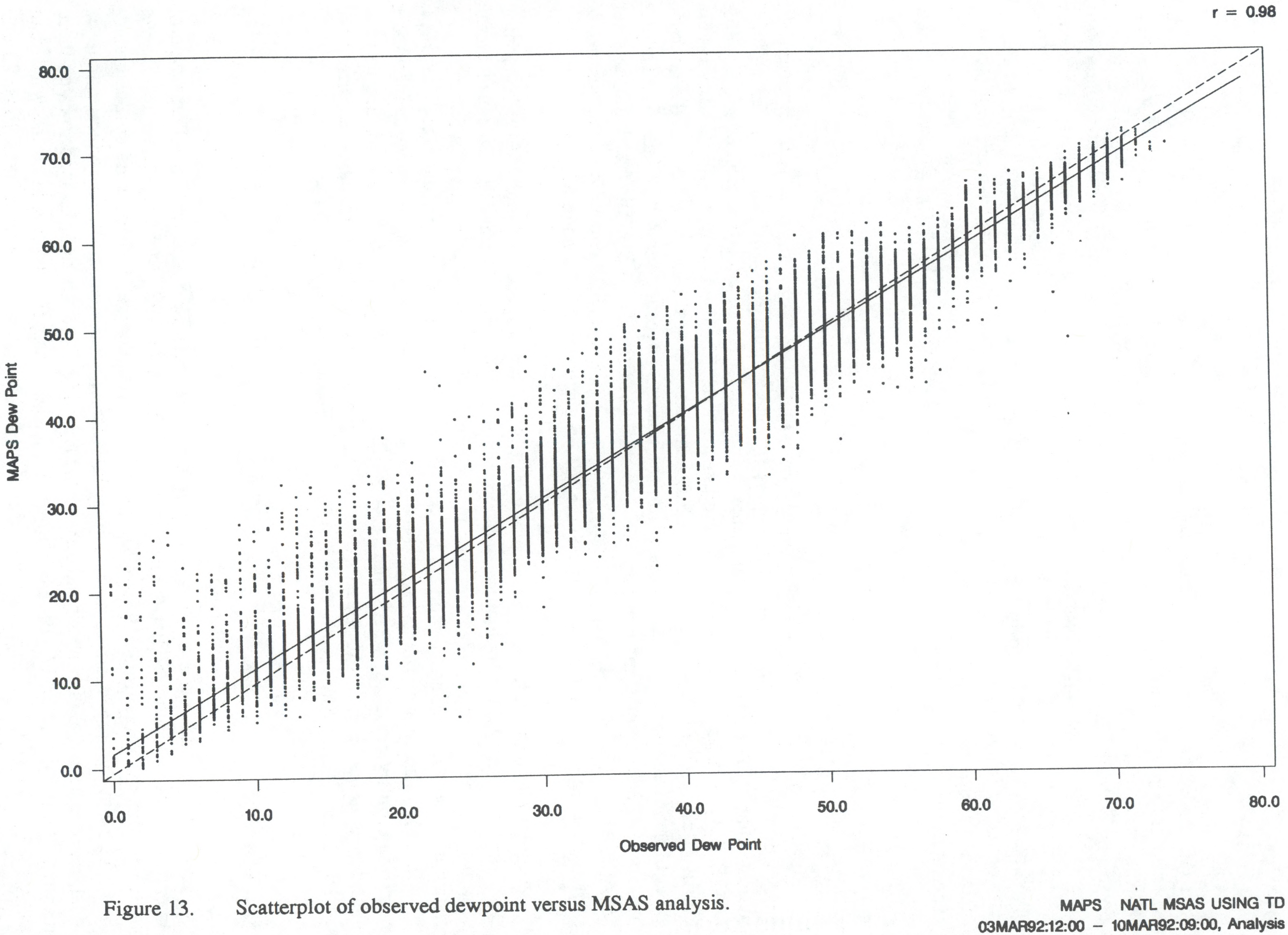


Figure 13. Scatterplot of observed dewpoint versus MSAS analysis.

Observed vs. MAPS Dewpoint

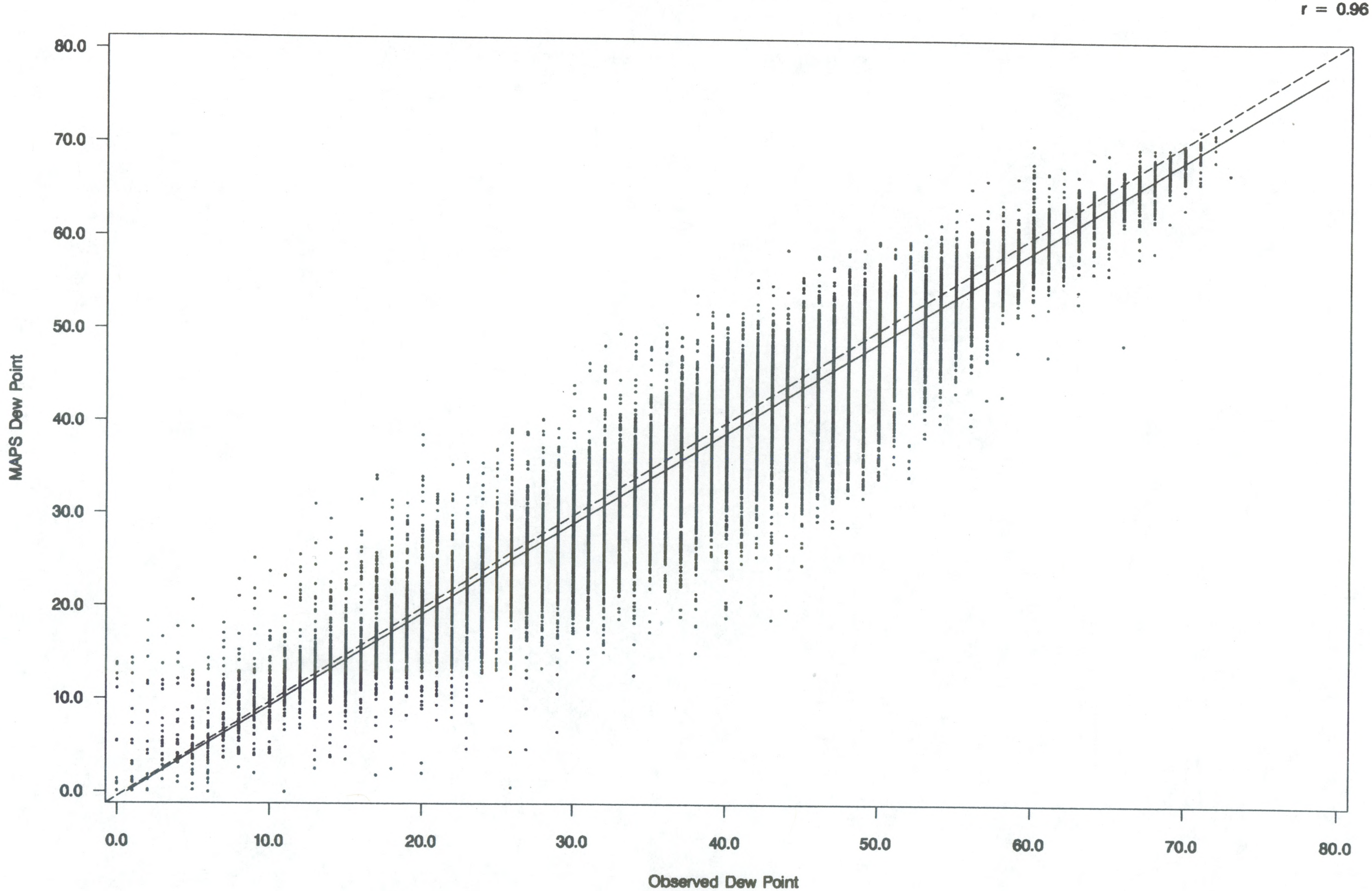


Figure 14. Scatterplot of observed dewpoint versus MAPS analysis.

MSAS Dewpoint Bias

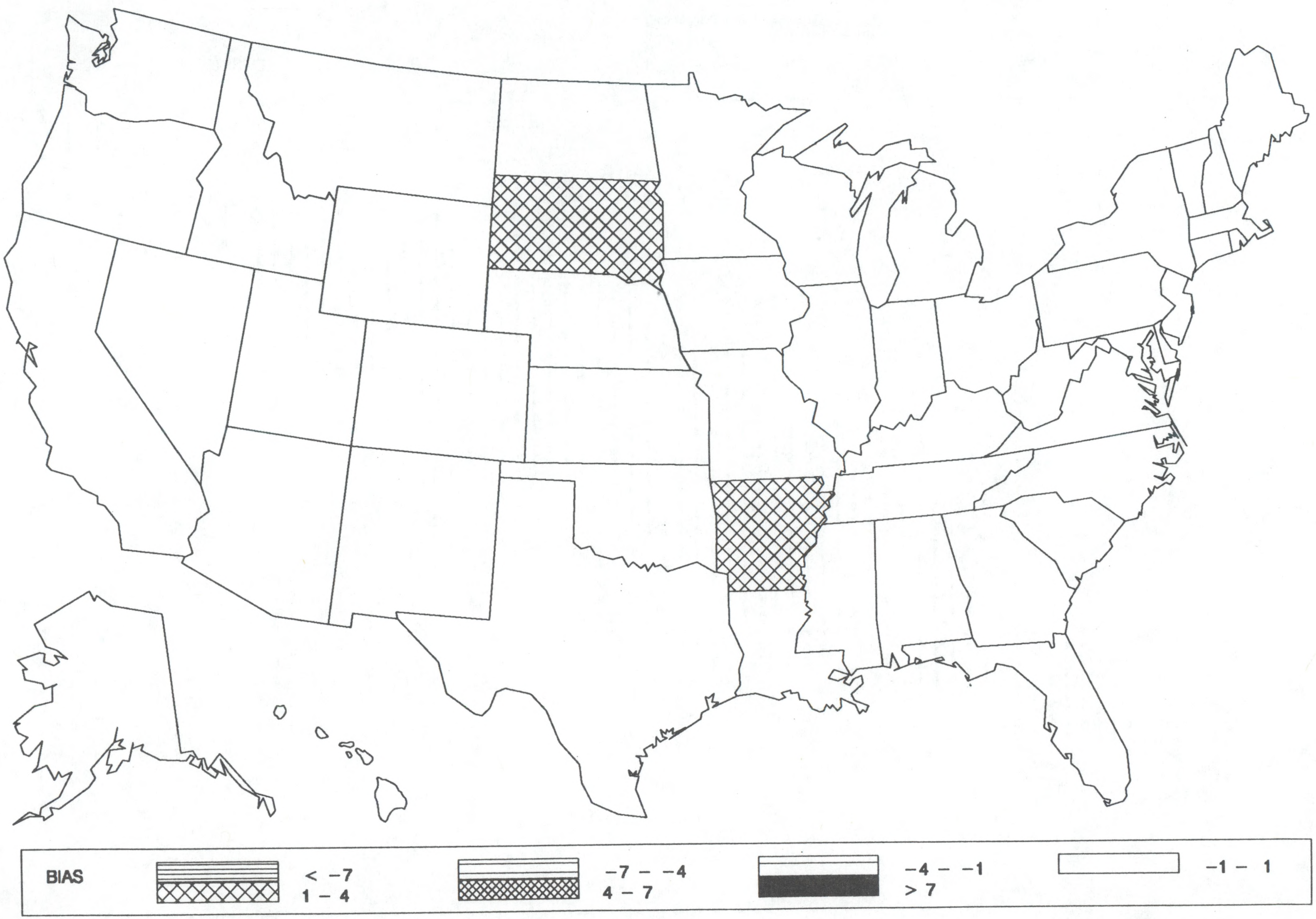


Figure 15. Geographical plot of MSAS dewpoint bias.

MAPS Dewpoint Bias

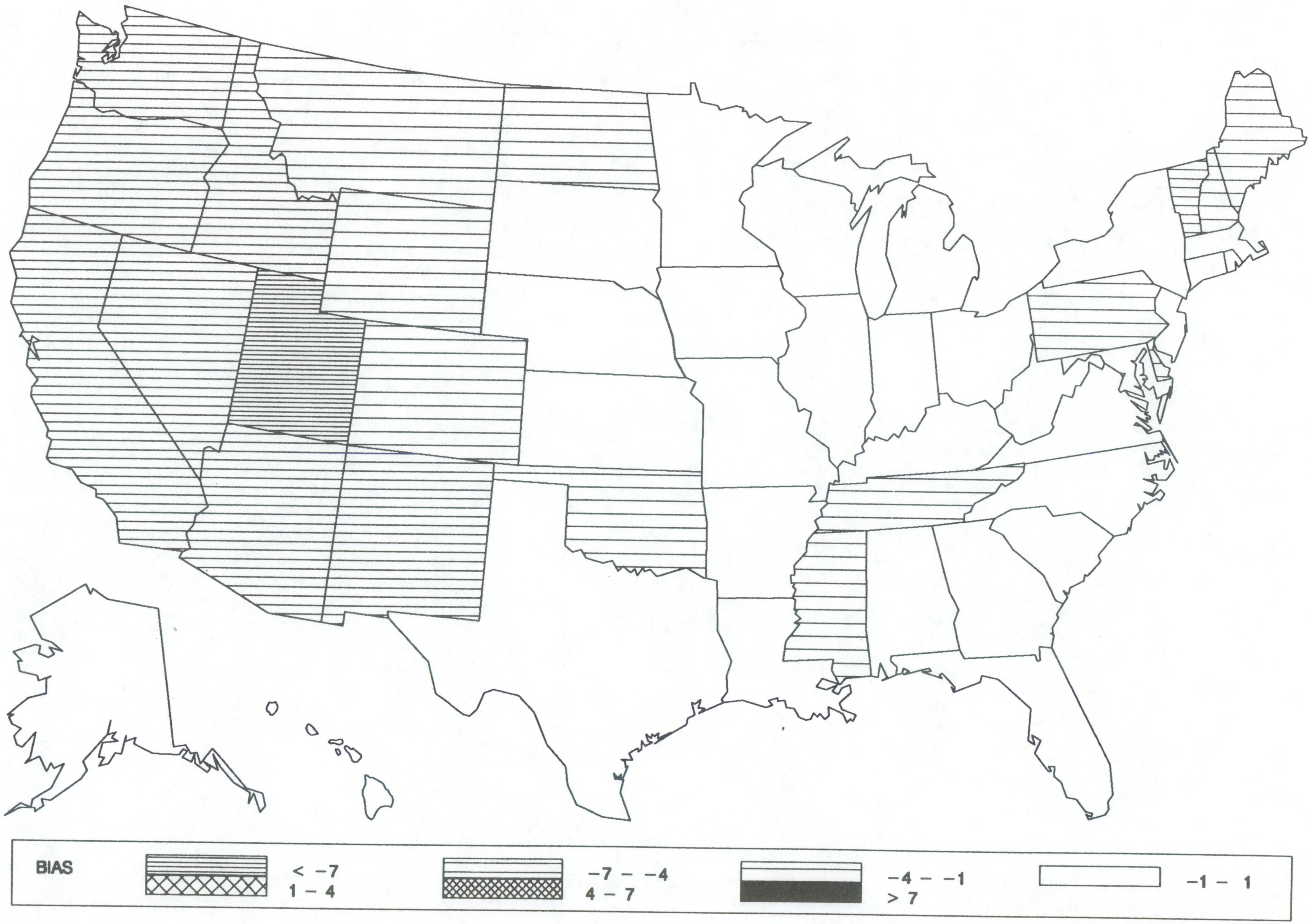


Figure 16. Geographical plot of MAPS dewpoint bias.

MAPS RAW MAPS DATA
03MAR92:12:00 − 10MAR92:12:00, Analysis

Box plots of MSAS Relative Humidity Errors

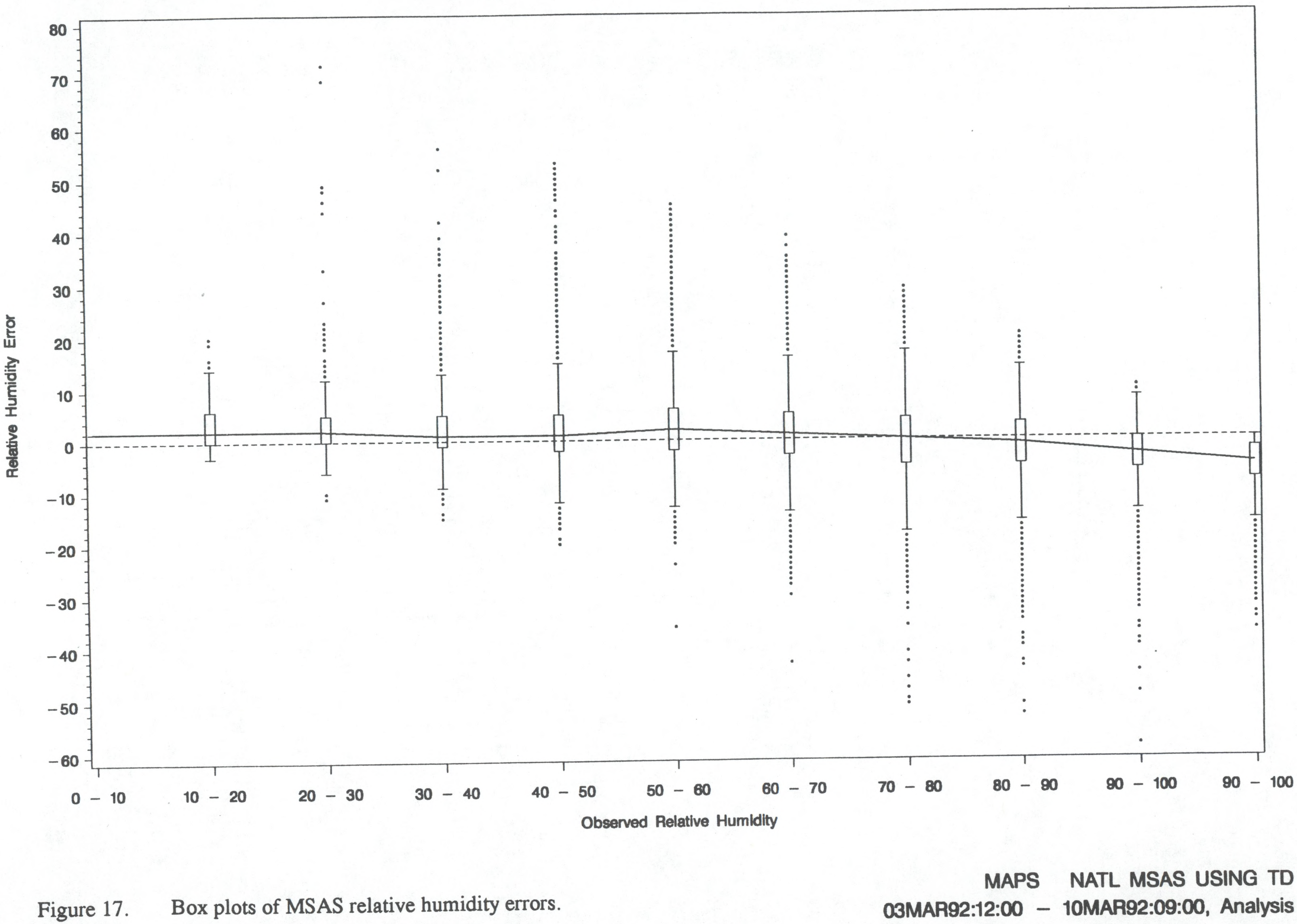


Figure 17. Box plots of MSAS relative humidity errors.

Box Plots of MAPS Relative Humidity Errors

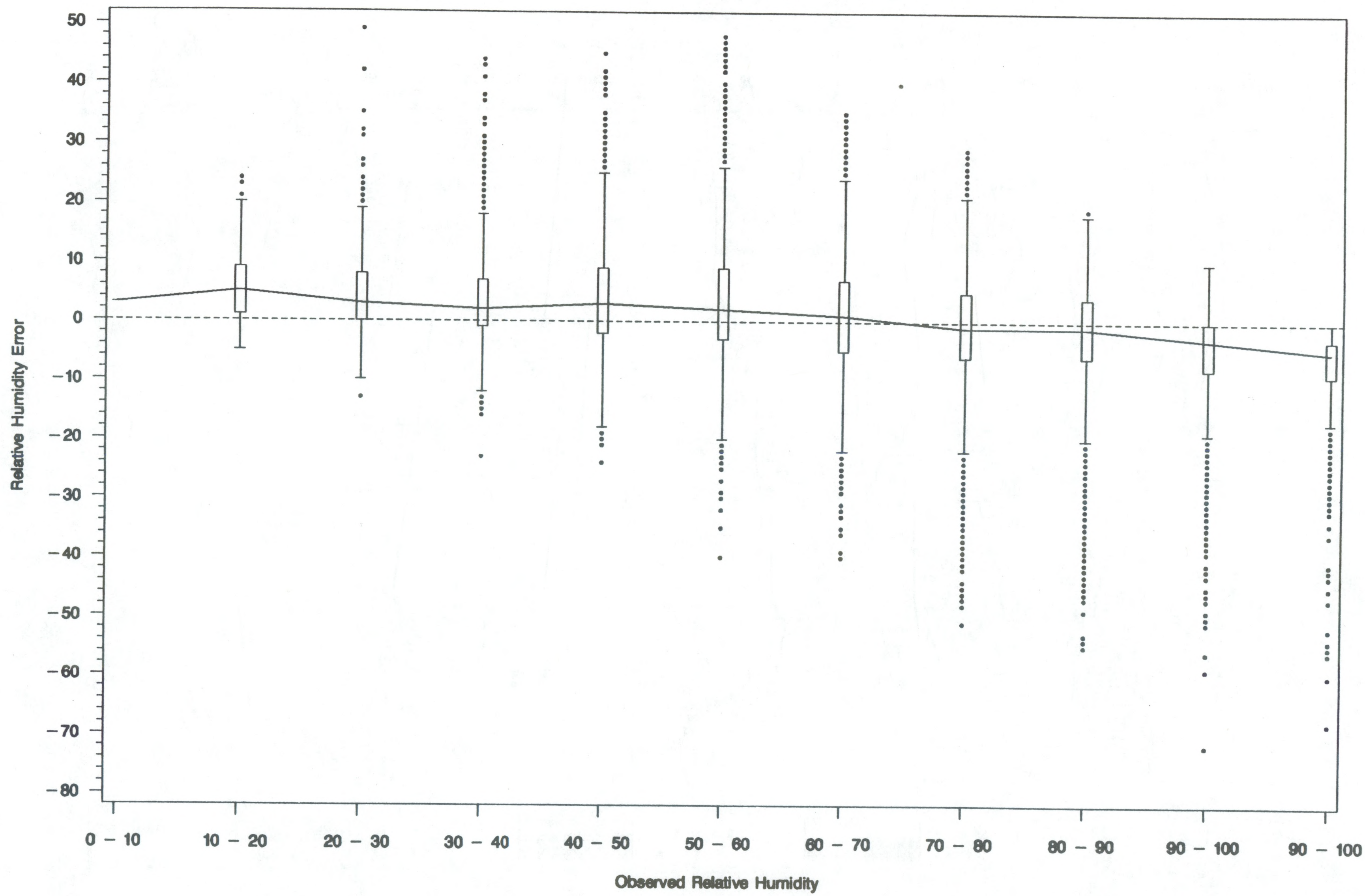


Figure 18. Box plots of MAPS relative humidity errors.

Box Plots of MSAS Wind Speed Errors

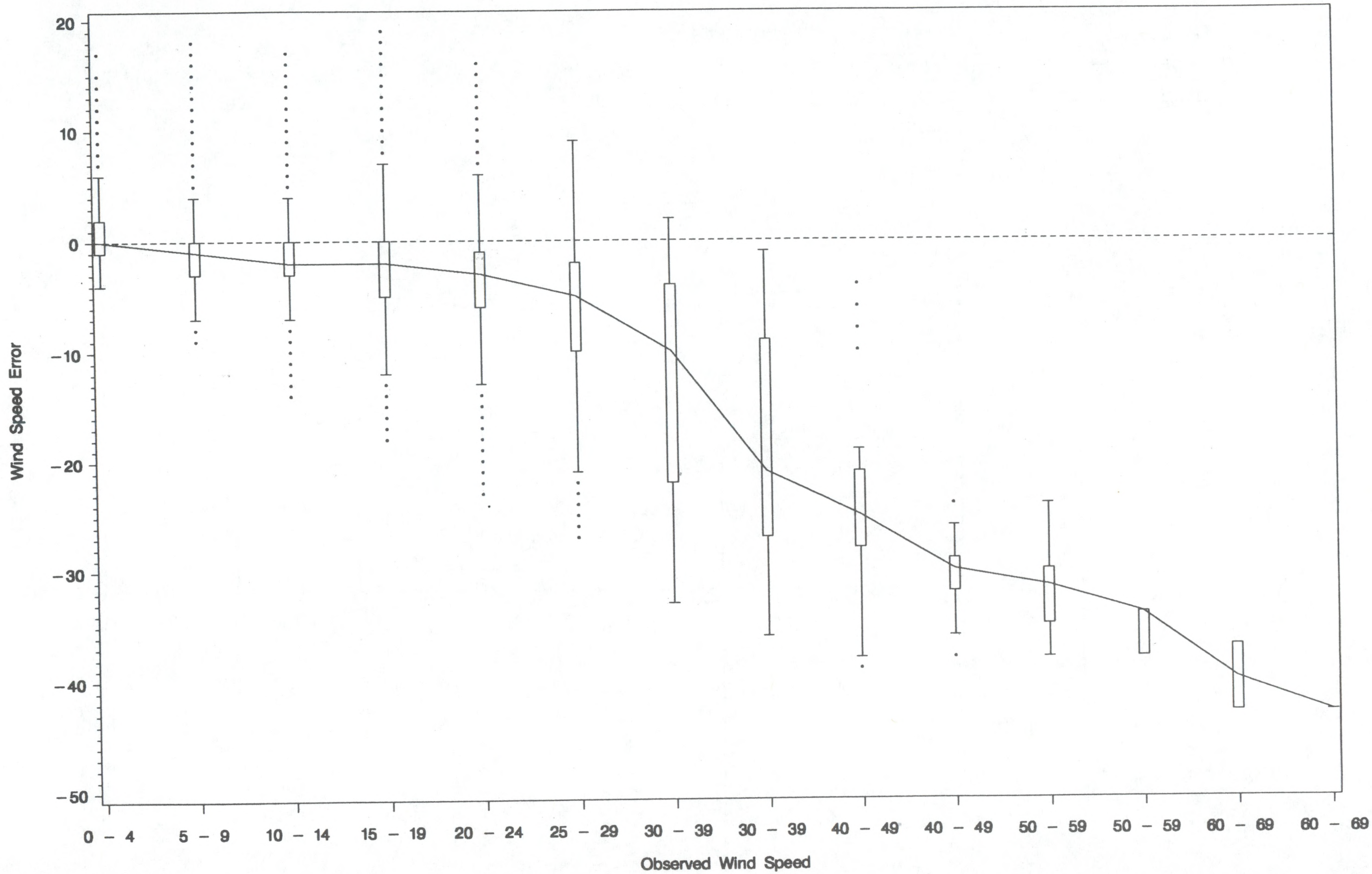


Figure 19. Box plots of observed wind speed error (kt) versus MSAS analysis. The box represents the 25th percentile, median, and 75th percentile, the whiskers the 10th and 90th percentiles and the dots the outliers.

MAPS NATL MSAS USING TD
03MAR92:12:00 − 10MAR92:09:00, Analysis

Box Plots of MAPS Wind Speed Errors

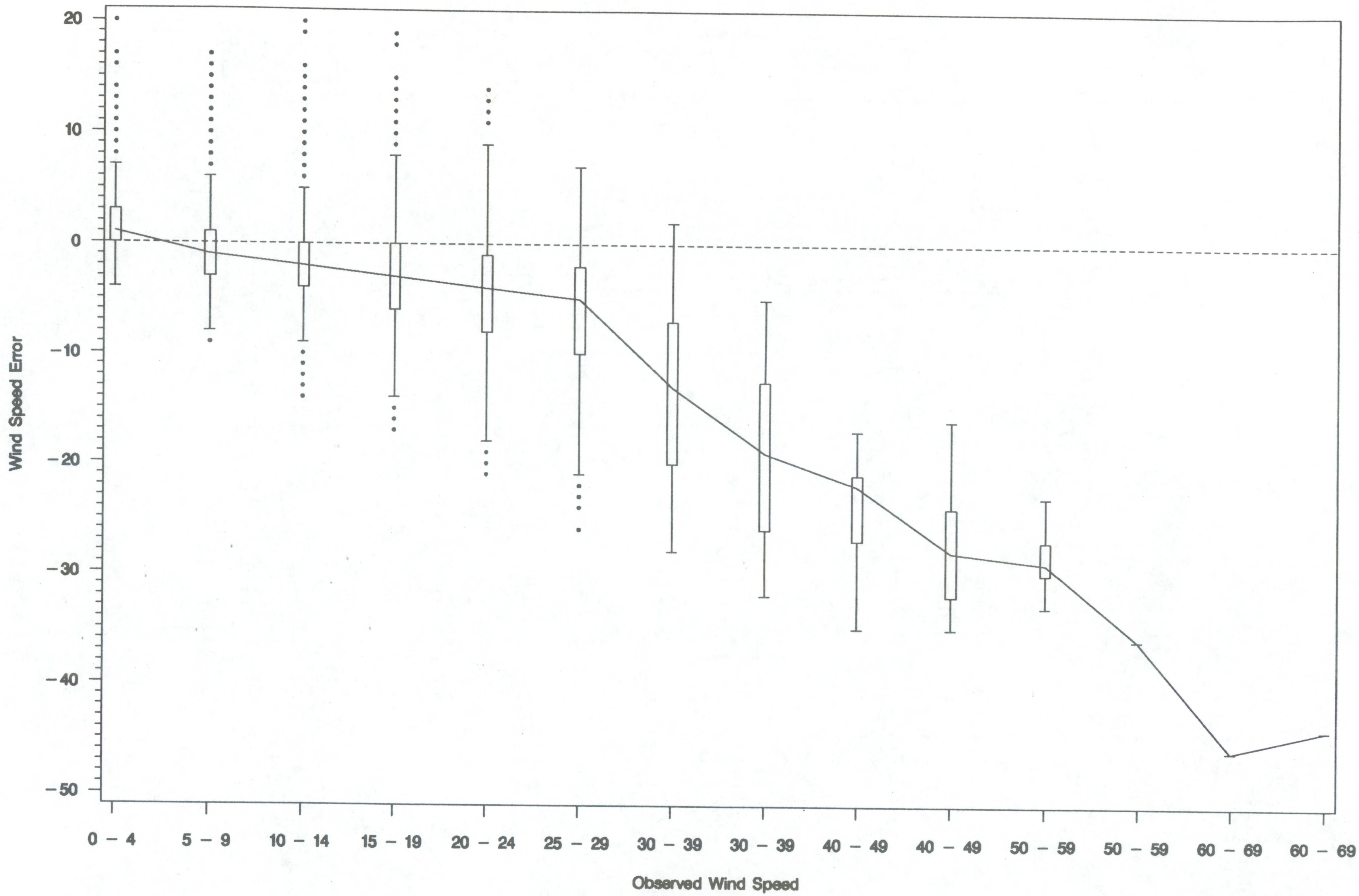


Figure 20. Same as Fig. 19, except for MAPS analysis.

MAPS RAW MAPS DATA
03MAR92:12:00 − 10MAR92:12:00, Analysis

MAPS Wind Speed Bias

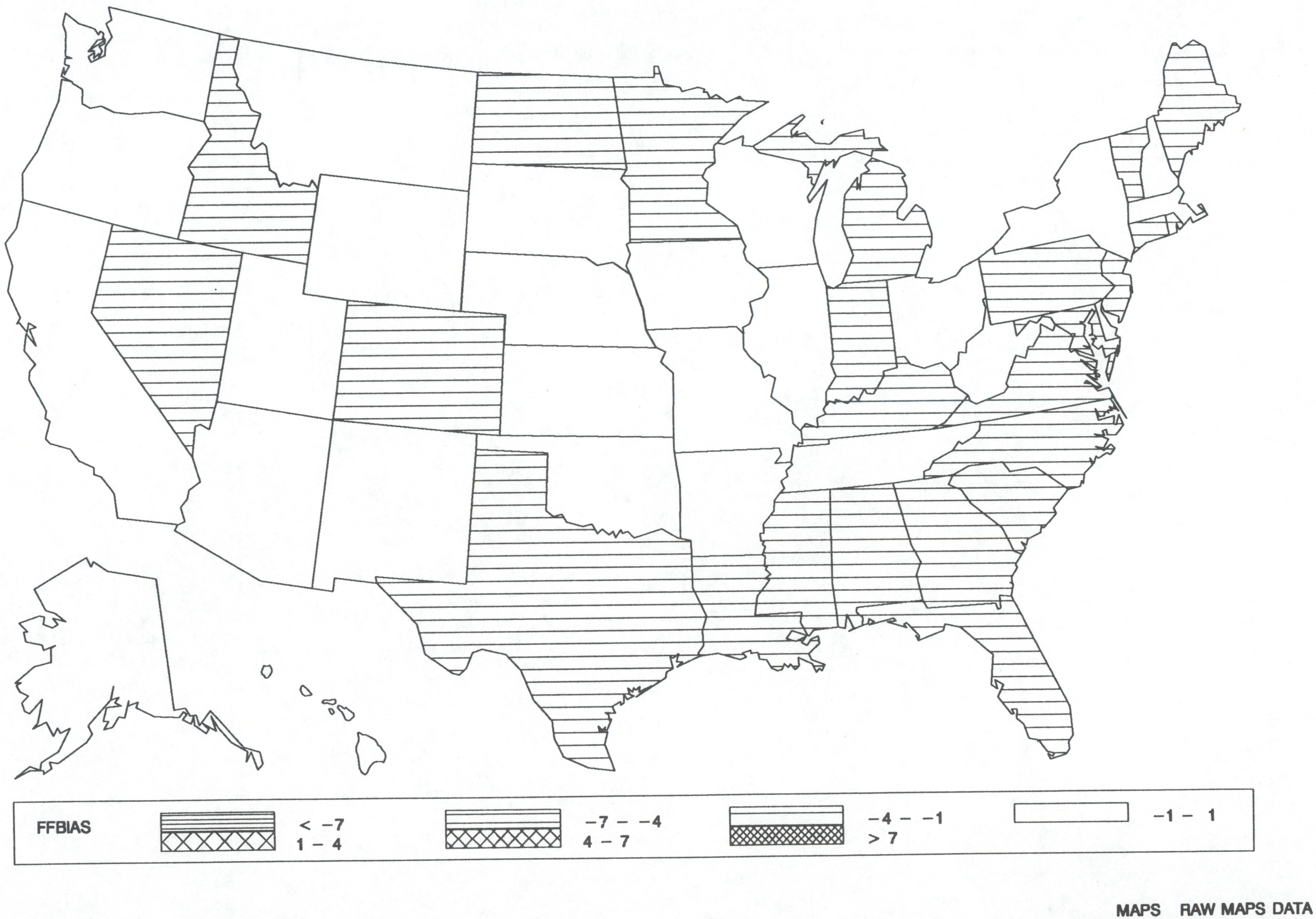


Figure 21. Geographical plot of MAPS wind speed bias.

MSAS Wind Speed Bias

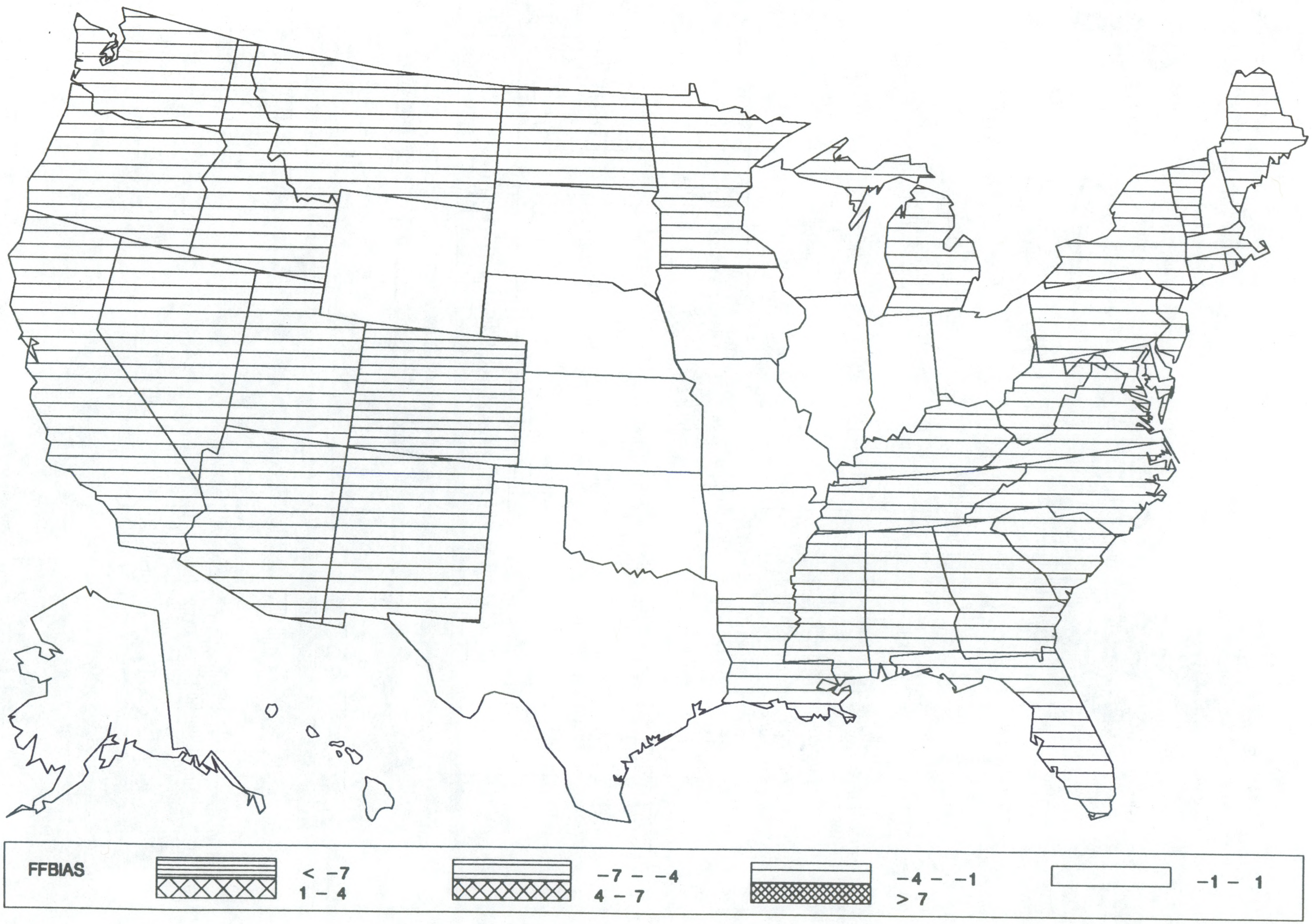


Figure 22. Geographical plot MSAS wind speed bias.

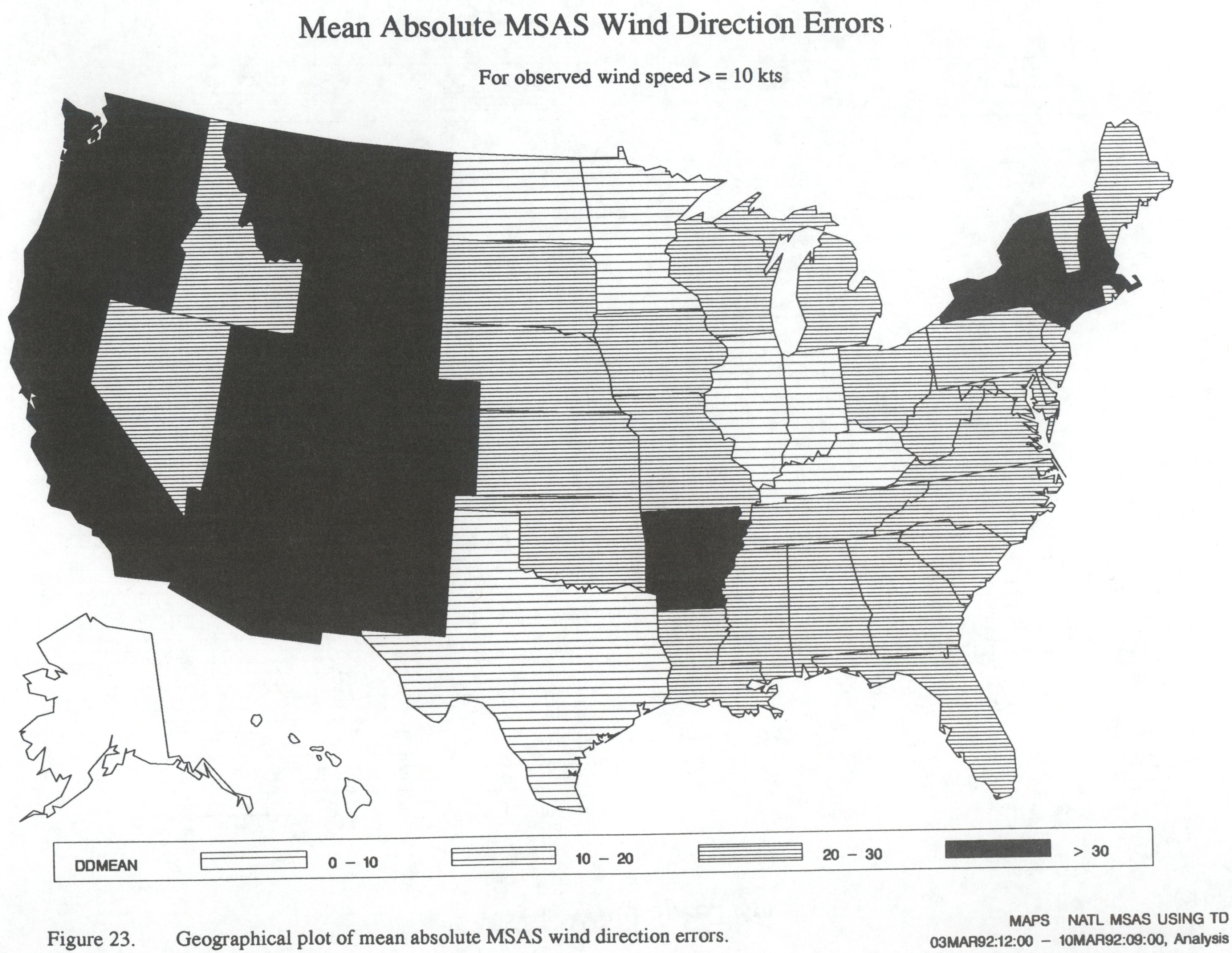


Figure 23. Geographical plot of mean absolute MSAS wind direction errors.

Mean Absolute MAPS Wind Direction Errors

For observed wind speeds >= 10 kts

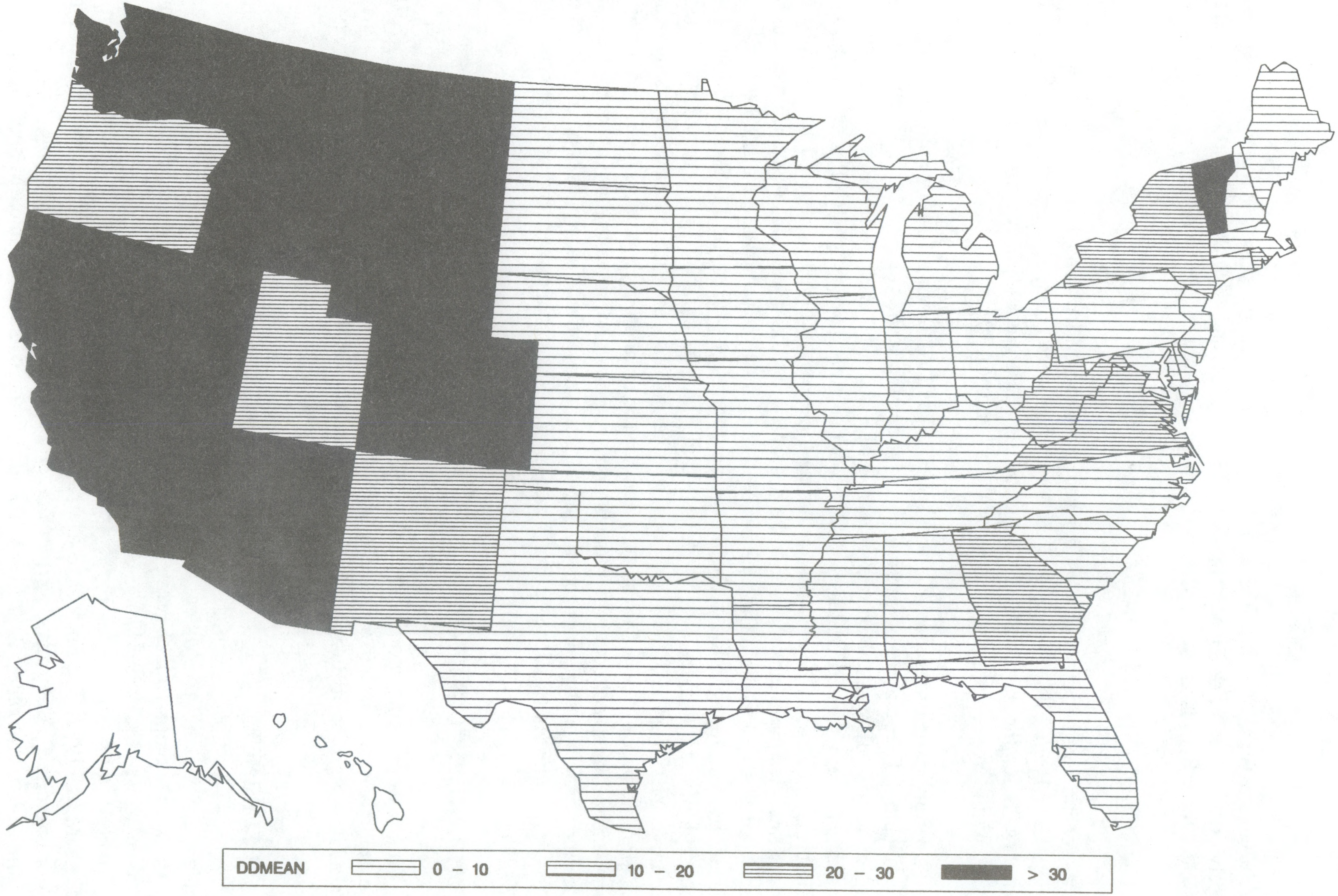


Figure 24. Geographical plot of mean absolute MAPS wind direction errors.

Intermountain West. Overall, the RMSE errors for speed and direction for the two systems are closely comparable. For speed, MSAS RMS is 3.75 kt and MAPS is 3.90 kt, and for direction, the RMS for MSAS is 25.81 deg and 30.14 deg for MAPS.

6. CONCLUDING REMARKS

A second evaluation of SAVs derived from the MSAS analyses was conducted from 3 through 10 March 1992. In general, better results were obtained from the MSAS system than from MAPS for the surface fields; altimeter, temperature and dewpoint. The analyses of the surface winds were roughly the same for MSAS and MAPS.

Two important differences between MSAS and MAPS analysis techniques may explain these results. The number of surface observations ingested into the MAPS system is roughly a third of nearly all available observations. The majority of these missing observations occur over the western United States. MSAS, on the other hand, ingests all surface observations available at the time of its analysis, including those over the western U.S. Therefore, the MSAS surface analyses more closely follow the observations. The inclusion or exclusion of the surface observations is tied to the difference between the actual surface pressure and the model pressure defined by the topography. The envelope topography used in MAPS produces a higher topography field that rides over the true terrain. This topography field is needed to sustain the predictive part of MAPS. MSAS, on the other hand, uses a minimum topography-minimum elevation method. These MSAS techniques were developed to extract consistent information from surface data collected in mountainous regions and to take into account the physical blocking and channeling induced by the moutainous terrain, as well as, distinguish between different air masses (Miller and Benjamin 1992). Future improvements are expected for both systems.

7. ACKNOWLEDGMENTS

This work is being sponsored by the FAA's Aviation Weather Research Program (ARD-80). The authors would like to thank the MAPS developers who provided the data for this study, Craig Hartsough for his insightful comments, and Nita Fullerton for her editorial review.

8. REFERENCES

Benjamin, S.G. and P.A. Miller, 1990: An alternative sea level pressure reduction and a statistical comparison of geostrophic wind estimates with observed surface winds. *Mon. Wea. Rev.*, **118**, 2099-2116.

Cairns, M.M., 1992: The Forecast Systems Laboratory's Aviation Verification Program Plan for Exercise 1. Preprints, *Twelfth Conf. on Probability and Statistics in the Atmospheric Sciences*, 22-26 June, Toronto, Ontario, Canada, Amer. Meteor. Soc., Boston, MA, 150-152.

Cairns, M.M., R.J. Miller, S.C. Albers, D.L. Birkenheuer, B.D. Jaminson, C.S. Hartsough, J.L. Mahoney, A. Marroquin, P.T. McCaslin, J.E. Ramer, and J.M. Schmidt, 1993: A preliminary evaluation of aviation-impact variables derived from numerical models. NOAA Tech. Memo. ERL FSL-5, NOAA Forecast Systems Laboratory, Boulder, CO, 165 pp.

Cairns, M.M., R.J. Miller, J. Chen, J.L. Mahoney, S.G. Benjamin, J.E. Ramer, T.L. Smith, 1994a: A second evaluation of aviation-impact variables generated by MAPS. NOAA Tech. Memo. ERL FSL-13, NOAA Forecast Systems Laboratory, Boulder, CO 96 pp.

Cairns, M.M., R.J. Miller, J. Chen, A. Marroquin, J.L. Mahoney, 1994b: A second evaluation of aviation-impact variables generated by Eta. NOAA Tech. Memo. ERL FSL-14, NOAA Forecast Systems Laboratory, Boulder, CO 94 pp.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, MA, 457 pp.

Doswell, C.A.III, and J. A. Flueck, 1989: Forecasting and verifying in a field research Project: DOPLIGHT '87. *Wea. Forecasting*, **4**, 97-109.

Doswell, C.A.III, R. Davies-Jones, and D.L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting,* **5**, 576-585.

Hoehne, W.E., 1980: Precision of National Weather Service upper-air measurements. NOAA Technical Memorandum NWS T&ED-16, Office of Systems Development, Test and Evaluation Division, Sterling, VA, 22 pp.

Kraus, M.J. 1993: The Forecast Systems Laboratory's role in the FAA's aviation weather development program. Preprints, *5th International Conference on Aviation Weather Systems,* 2-5 August, Vienna, Virginia, Amer. Meteor. Soc., Boston, MA, p5.

Mahoney, J. Luppens, C. Hartsough, P.A. Miller, M.M. Cairns, R.J. Miller, 1995: A second evaluation of aviation-impact variables generated by LAPS. NOAA Tech. Memo. ERL FSL-16, NOAA Forecast Systems Laboratory, Boulder, CO 89pp.

Miller, P.A. and S.G. Benjamin, 1992: A system for the hourly assimilation of surface observations in mountainous and flat terrain. *Mon. Wea. Rev.*, **120**, 2342-2359.

Miller, R.J. and M.M. Cairns, 1993: Preliminary evaluation of numerical analysis and prediction of aviation-impact variables. Preprints, *5th Int. Conf. on Aviation Weather Systems,* 2-6 August, Vienna, VA, Amer. Meteor. Soc., Boston, MA, 241-244.

Murphy, A.H., B.G. Brown and Y. Chen, 1989: Diagnostic verification of temperature forecasts. *Wea. Forecasting*, **4**, 485-501.

NOAA, 1991: ASOS Product Specification, Contract No. 50-SANW-1-00050, ASOS Program Office, NOAA/SPO2, 8455 Colesville, Suite 705, Silver Springs, MD 20910, 95 pp.

NWS, 1991a: Functional Precision of National Weather Service Upper-Air Measurements using VIZ Manufacturing Co. "B" Radiosonde (Model 1492-520). NOAA Technical Report NWS 45, Office of Systems Operations, Engineering Division, Test and Evaluation Branch, Sterling, VA, 39 pp.

NWS, 1991b: Functional Precision of National Weather Service Upper-Air Measurements using Space Data Division Radiosonde (Model 909-10-01). NOAA Technical Report NWS 46, Office of Systems Operations, Engineering Division, Test and Evaluation Branch, Sterling, VA, 41 pp.

Panofsky, H.A. and G.W. Brier, 1963: *Some Applications of Statistics to Meteorology*. The Pennsylvania State University, University Park, Pennsylvania, 224 pp.

Stanski, H.R., L.J. Wilson, and W.R. Burrows, 1989: Survey of common verification methods in meteorology. Atmospheric Environment Service Research Report (MSRB) 89-5, Downsview, Ontario, Canada, 114 pp.

APPENDIX A

SUMMARY OF STATISTICAL MEASURES

A multitude of statistical measures are available for the description of a set of data. Among the most commonly used measures for meteorological verification are the bias or mean error, mean absolute error, root mean square error, percent correct, probability of detection, false alarm rate, critical success index, true skill statistic, and the Heidke skill score.

The bias, mean absolute error, and root mean square error are easily defined using Fi as the forecasts, Oi as the observations, and n as the number of forecast/observation pairs. The rest of the measures discussed here are formulated based on a contingency table. For complete descriptions, see Panofsky and Brier (1963), Stanski et al. (1989), Murphy et al. (1989), Doswell and Flueck (1989), and Doswell et al. (1990). Table 4 provides a quick reference for interpreting the verification measures.

Table 4: Quick Reference for Interpreting Verification Measures

VERIFICATION MEASURE	RANGE	"BEST" SCORE
Bias or Mean Error (ME)	-00 to +00	0.0
Mean Absolute Error (MAE)	0.0 to +00	0.0
Root Mean Square Error (RMSE)	MAE to +00	0.0

Bias or Mean Error

The bias, or mean (algebraic) error (ME), indicates the average direction of the deviation of the forecasts from the observed values. The bias is defined as

$$\text{Bias} = \text{ME} = (1/n) \sum_{i=1}^{n} (F_i - O_i)$$

A positive bias indicates that the forecast exceeds the observed value on the average (overforecasting), and a negative bias corresponds to a forecast below the observed value on the average (underforecasting). Also, for all arithmetic errors in this memorandum, the subtraction is always performed as model minus observation. For example, a positive arithmetic temperature error means that the model was too warm. The bias range is from $-\infty$ to $+\infty$; a value of zero is desired.

Mean Absolute Error

The mean absolute error (MAE) is a linear score that calculates the average magnitude of the error. The MAE is defined as

$$\mathrm{MAE} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left| F_i - O_i \right|$$

The MAE range is from 0 to ì; a MAE of 0 is desired.

Root Mean Square Error

The root mean square error (RMSE) is commonly used in meteorology. The RMSE is a quadratic score that gives the average magnitude of the errors, and is defined as

$$\mathrm{RMSE} = \left[\left(\frac{1}{n}\right) \sum_{i=1}^{n} (F_i - O_i)^2 \right]^{\frac{1}{2}}$$

The RMSE gives more weight to large errors than to small errors in the average, and is useful when large errors are undesirable. Values for RMSE range from the MAE to 1, and values close to the MAE (or zero) are desired.

The similar root mean square vector error (RMSVE) is designed for evaluating magnitude and directional errors of the wind. RMSVE is defined as where u and v are the components of the wind, and, as before, f and o denote forecast and observed, respectively.

$$\mathrm{RMSVE} = \left(\left(\frac{1}{n}\right) \sum_{i=1}^{n} (u_{fi} - u_{oi})^2 + (v_{fi} - v_{oi})^2 \right)^{1/2}$$

APPENDIX B

LIST OF STATISTICAL RESULTS FOR SURFACE SAVs

	MSAS	MAPS
Altimeter (in. Hg)		
ME	.001	-.001
MAE	.011	.020
RMSE	.019	.029
N	47897	23204
Temperature (F)		
ME	.054	-.939
MAE	1.486	2.582
RMS	2.225	3.910
N	56817	27504
Dewpoint (F)		
ME	.139	-1.181
MAE	1.718	2.847
RMS	2.601	4.047
N	56179	27182
Relative Humidity (%)		
ME	-0.591	-1.038
MAE	5.362	7.155
RMS	7.557	9.940
N	54275	27156

	MSAS	MAPS
Wind Speed (kt)		
ME	-2.031	-.7
MAE	2.492	2.9
RMS	3.756	4.0
N	56359	27287
Wind Direction (deg) Errors. Observed Wind Speed >= 10 kt		
ME	-2.031	-1.298
MAE	14.069	18.895
RMS	25.812	30.143
N	17552	6654

* U.S. GOVERNMENT PTINTING OFFICE 1996-0-774-018/40011