NOAA Technical Memorandum ERL FSL-5

A PRELIMINARY EVALUATION OF AVIATION-IMPACT VARIABLES DERIVED FROM NUMERICAL MODELS

Mary M. Cairns
Ronald J. Miller
Steven C. Albers
Daniel L. Birkenheuer
Brian D. Jamison
Craig S. Hartsough
Jennifer L. Mahoney
Adrian Marroquin
Paula T. McCaslin
James E. Ramer
Jerome M. Schmidt

Forecast Systems Laboratory
Boulder, Colorado
April 1993

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

NOAA Technical Memorandum ERL FSL-5

A PRELIMINARY EVALUATION OF AVIATION-IMPACT VARIABLES DERIVED FROM NUMERICAL MODELS

Mary M. Cairns
Ronald J. Miller
Steven C. Albers
Daniel L. Birkenheuer
Brian D. Jamison
Craig S. Hartsough
Jennifer L. Mahoney
Adrian Marroquin
Paula T. McCaslin
James E. Ramer
Jerome M. Schmidt

Forecast Systems Laboratory
Boulder, Colorado
April 1993

UNITED STATES
DEPARTMENT OF COMMERCE

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION

Environmental Research
Laboratories

NOTICE

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

For sale by the National Technical Information Services
5285 Port Royal Road, Springfield, VA 22061

CONTENTS

ABSTRACT 1

1. INTRODUCTION 2

2. MODEL AND ANALYSIS DESCRIPTIONS 2

2.1. MAPS 3
2.2. ETA 3
2.3. LAPS 4
2.4. RAMS 4

3. VERIFICATION DATA 6

3.1. Surface Data 6
3.2. Rawinsonde Data 9
3.3. Profiler Data 9
3.4. Instrument Precision 9

4. FORECASTS AND ANALYSES 11

4.1. Model Forecasts and Analysis Types 11
4.2. Areas and Frequencies 12
4.3. Derivation of Aviation Impact Variables 14

5. STATISTICAL MEASURES 16

5.1. Summary of Statistical Measures 20

5.1.1. Bias or Mean (Algebraic) Error 21
5.1.2. Mean Absolute Error 21
5.1.3. Root Mean Square Error 21
5.1.4. Percent Correct 22
5.1.5. Probability of Detection 22
5.1.6. False Alarm Rate 22
5.1.7. Critical Success Index 22
5.1.8. True Skill Statistic 22
5.1.9. Heidke Skill Score 23

6. MODEL RESULTS 23

6.1. MAPS 23

6.1.1. SAVs 23
6.1.2. AIVs 42
6.1.3. Discussion 54

6.2. Eta 56

6.2.1. SAVs 57
6.2.2. AIVs 78
6.2.3. Discussion 83

6.3. LAPS 83

6.3.1. SAVs 83
6.3.2. AIVs 90
6.3.3. Discussion 95

6.4. RAMS 96

6.4.1. SAVs 97
6.4.2. AIVs 113
6.4.3. Discussion 116

7. CONCLUDING REMARKS 117

8. FUTURE PLANS 118

9. ACKNOWLEDGMENTS 118

10. REFERENCES 118

APPENDIX A WEATHER DISCUSSION FOR 1-10 APRIL 1991 122

APPENDIX B DIAGNOSING SAVS AND AIVS IN MAPS, Eta, LAPS, AND RAMS 127

APPENDIX C LIST OF MAPS STATISTICAL RESULTS 137

APPENDIX D LIST OF Eta STATISTICAL RESULTS 143

APPENDIX E LIST OF LAPS STATISTICAL RESULTS 149

APPENDIX F LIST OF RAMS STATISTICAL RESULTS 150

APPENDIX G LIST OF SYSTEM DEVELOPERS 162

GLOSSARY 164

FIGURES

Figure 1. Surface station locations, national domain, and four WFO domain forecast and analysis areas 5

Figure 2. Upper air and profiler station locations, national domain, and WFO domain forecast and analysis areas 10

Figure 3. Mean error of height, temperature, and dewpoint temperature between MAPS analyses and RAOBS 25

Figure 4. Same as Figure 3, except for MAPS 12-h forecast data 26

Figure 5. Root mean square error of height, temperature, and dewpoint temperature between MAPS analyses and RAOBS 27

Figure 6. Same as Figure 5, except for MAPS 12-h forecast data 28

Figure 7. Scatterplot of observed 850 mb dewpoint temperatures and MAPS analyses of dewpoint temperatures in °C 30

Figure 8. Same as Figure 7, except for MAPS 12-h forecast data 31

Figure 9. Mean error of wind speed and mean absolute error of wind direction between MAPS analyses and RAOBS 32

Figure 10. Same as Figure 9, except for MAPS 12-h forecast data 33

Figure 11. Mean error of wind speed and mean absolute error of wind direction between MAPS analyses and wind profilers 35

Figure 12. Same as Figure 11, except for MAPS 12-h forecast data 36

Figure 13. Map of mean absolute altimeter setting errors between MAPS analyses and surface observations 37

Figure 14. Scatterplot of observed surface temperatures versus MAPS analyses of surface temperatures in °F 38

Figure 15. Same as Figure 14, except for MAPS 12-h forecast data 39

Figure 16. Map of mean error of surface temperature between MAPS analyses and surface observations 40

Figure 17. Scatterplot of observed surface dewpoint temperatures versus MAPS analyses of surface dewpoint temperatures in °F 41

Figure 18. Map of mean absolute wind speed errors between MAPS analyses and surface observations 43

Figure 19. Distribution of mean absolute wind speed errors by wind direction between MAPS analyses and surface observations 44

Figure 20. Map of mean absolute wind direction error between MAPS analyses and surface observations 45

Figure 21. Distribution of observed and MAPS analyses of low cloud amounts. 46

Figure 22. Scatterplot of observed cloud-top heights versus MAPS analyses of cloud-top heights 48

Figure 23. Distribution of observed and MAPS analyses of ceiling and visibility by category 49

Figure 24. Same as Figure 23, except for MAPS 12-h forecast data 50

Figure 25. Distribution of observed obstructions to visibility versus MAPS analyses of obstructions to visibility 53

Figure 26. Scatterplot of observed 3-h precipitation amounts versus MAPS 3-h forecast accumulations of precipitation amounts 55

Figure 27. Mean error of height, temperature, and dewpoint temperature between Eta initial hour data and RAOBs 58

Figure 28. Same as Figure 27, except for Eta 12-h forecast data 59

Figure 29. Mean error of wind speed and mean absolute error of wind direction between Eta initial hour data and RAOBs 60

Figure 30. Same as Figure 29, except for Eta 12-h forecast data 61

Figure 31. Map of mean absolute altimeter setting errors between Eta initial hour data and surface observations 62

Figure 32. Distribution of observed surface temperatures and Eta initial hour surface temperatures 64

Figure 33. Scatterplot of observed surface temperatures versus Eta initial hour surface temperatures 65

Figure 34. Map of mean absolute surface temperature errors between Eta initial hour data and surface observations 66

Figure 35. Scatterplot of observed surface temperatures versus Eta 3-h forecast surface temperatures 67

Figure 36. Scatterplot of observed surface dewpoint temperatures versus Eta initial hour surface dewpoint temperatures 68

Figure 37. Same as Figure 36, except for Eta 12-h forecast data 69

Figure 38. Distribution of observed surface dewpoint temperatures and Eta initial hour surface dewpoint temperatures 71

Figure 39. Same as Figure 38, except for Eta 12-h forecast data 72

Figure 40. Map of mean errors of surface wind speed between Eta initial hour data and surface observations 73

Figure 41. Distribution of mean absolute wind speed errors by wind direction between Eta initial hour data and surface observations 74

Figure 42. Map of mean absolute wind direction error between Eta initial hour data and surface observations 75

Figure 43. Distribution of observed and Eta initial hour low cloud amounts . . . 79

Figure 44. Scatterplot of observed cloud-top heights versus Eta initial hour cloud-top heights 80

Figure 45. Scatterplot of observed 3-h precipitation amounts versus Eta 12-h forecast 3-h precipitation amounts 82

Figure 46. Mean error of height and temperature between LAPS and Denver RAOB 85

Figure 47. Root mean square error of wind speed between LAPS and Denver RAOB 86

Figure 48. Mean error of vertical wind speed between LAPS and Platteville wind profiler 88

Figure 49. Surface wind speed distribution of observed and LAPS winds 91

Figure 50. Total cloud amount distribution of observed and LAPS clouds 92

Figure 51. Scatterplot of observed cloud-top heights versus LAPS analyzed cloud-top heights 94

Figure 52. Mean error of height, temperature, and dewpoint temperature between RAMS initial hour data and RAOBs for the Colorado and Central Plains domains 98

Figure 53. Same as Figure 52, except for the Florida and Northeast U.S. domains 99

Figure 54. Mean error of wind speed and mean absolute error of wind direction between RAMS initial hour data and RAOBs for the Colorado and Central Plains domains 100

Figure 55. Same as Figure 54, except for the Florida and Northeast U.S. domains 101

Figure 56. Scatterplot of observed altimeter settings versus RAMS initial hour altimeter settings for the Florida and Northeast U.S. domains . . . 103

Figure 57. Surface observations for the Colorado domain at 0000 UTC 2 April 1991 104

Figure 58. Same as Figure 57, except for RAMS initial hour 105

Figure 59. Same as Figure 58, except for observations for the Colorado domain at 1800 UTC 1 April 1991 106

Figure 60. Same as Figure 59, except for MAPS analysis 106

Figure 61. Same as Figure 58, except for observations for the Colorado domain at 0600 UTC 2 April 1991 107

Figure 62. Same as Figure 61, except for RAMS 6-h forecast 107

Figure 63. Same as Figure 58, except for observations for the Northeast U.S. domain at 0000 UTC 2 April 1991 108

Figure 64. Same as Figure 63, except for RAMS initial hour 108

Figure 65. Same as Figure 58, except for observations for the Northeast U.S. domain at 0600 UTC 2 April 1991 109

Figure 66. Same as Figure 65, except for RAMS 6-h forecast 109

Figure 67. Same as Figure 58, except for observations for the Florida domain at 1200 UTC 5 April 1991 111

Figure 68. Same as Figure 67, except for RAMS initial hour 111

Figure 69. Same as Figure 58, except for observations for the Florida domain at 2200 UTC 5 April 1991 112

Figure 70. Same as Figure 69, except for RAMS 10-h forecast 112

Figure 71. Same as Figure 58, except for observations for the Central Plains domain at 1200 UTC 8 April 1991 114

Figure 72. Same as Figure 71, except for RAMS initial hour 114

Figure 73. Same as Figure 58, except for observations for the Central Plains domain at 2100 UTC 8 April 1991 115

Figure 74. Same as Figure 73, except for RAMS 9-h forecast 115

Figure 75. Map of mean observed 3-hr precipitation amounts for 1-10 April 1991 125

Figure 76. Map of observed frozen precipitation events for 1-10 April 1991 . 126

TABLES

Table 1. State-of-the-atmosphere and aviation-impact variables evaluated in Exercise 1 7

Table 2. Instrument root mean square differences 11

Table 3. SAVs and AIVs provided by models for Exercise 1 13

Table 4. Model run dates and times for RAMS 14

Table 5. AIV derivation methodology 15

Table 6. List of statistical measures by data source and variable 17

Table 7. Visibility and ceiling categories 20

Table 8. Contingency table of diagnosed low clouds by observed low clouds for MAPS analyses, 1-10 April 1991 47

Table 9. Contingency table of diagnosed ceilings by observed ceilings for MAPS analyses, 1-10 April 1991 51

Table 10. Contingency table of diagnosed visibilities by observed visibilities for MAPS analyses, 1-10 April 1991 52

Table 11. Contingency table of diagnosed precipitation phase by observed precipitation phase for MAPS analyses, 1-10 April 1991 54

Table 12. Contingency table of forecast wind direction by observed wind direction for Eta initial conditions, 1-10 April 1991 76

Table 13. Contingency table of forecast wind direction by observed wind direction for Eta 12-h forecasts, 1-10 April 1991 77

Table 14. Contingency table of forecast precipitation occurrence by observed precipitation occurrence for Eta 12-h forecasts, 1-10 April 1991 . . . 81

Table 15. Contingency table of diagnosed precipitation phase by observed precipitation phase for Eta 12-h forecasts, 1-10 April 1991 81

Table 16. LAPS absolute errors by station 89

Table 17. Contingency table of analyzed low clouds by observed low clouds for LAPS, 1-10 April 1991 93

Table 18. Contingency table of analyzed ceilings by observed ceilings for LAPS, 1-10 April 1991 95

Table 19. Extreme weather values observed at the surface stations used in Exercise 1 for 1-10 April 1991 124

A PRELIMINARY EVALUATION OF AVIATION-IMPACT VARIABLES DERIVED FROM NUMERICAL MODELS

Mary M. Cairns, Ronald J. Miller, Steven C. Albers, Daniel L. Birkenheuer, Brian D. Jamison, Craig S. Hartsough, Jennifer L. Mahoney, Adrian Marroquin, Paula T. McCaslin, James E. Ramer, Jerome M. Schmidt

ABSTRACT. This report describes the Aviation Division's Verification Program, located in the Forecast Systems Laboratory, and the results of the first evaluation of four analysis and forecast model systems. The impetus for the evaluation is to get baseline statistics of forecasts of aviation-impact variables (AIVs; e.g., clouds and visibility), which are commonly not forecast or verified in numerical weather prediction models. The results of this study are very preliminary. They are intended to serve as a baseline for future evaluations, not as judgments about the current model capabilities. Four systems were evaluated: Local Analysis and Prediction System (LAPS), Colorado State University Regional Atmospheric Modeling System (RAMS), Mesoscale Analysis and Prediction System (MAPS), and the Eta model. Analyses and forecasts were generated during a 10-day period, 1-10 April 1991, and verified at surface, upper-air, and profiler stations in the continental United States. This evaluation also included state-of-the-atmosphere variables (SAVs; e.g., temperature, winds, pressure). LAPS and RAMS were run in the Weather Forecast Office scale domain (approximately 500 nm by 500 nm); MAPS and Eta were run on a national-scale domain.

In general, it was found that all systems produced good analyses and forecasts of SAVs in all areas, except that the surface winds were too strong and wind directions were commonly 30° off. Also, all models were too dry. For AIVs, detection was good for ceiling, visibility, and cloud amount, but the ability to distinguish categories (e.g., scattered or broken clouds) was poor. This could be related to moisture problems in each system. Analyses and forecast skill of cloud top varied. For instance, Eta appears to predict too much cirrus cloud, whereas MAPS has too little, resulting in large cloud top errors. Only MAPS, Eta, and RAMS produced forecasts of precipitation amounts. Precipitation events were nearly always overforecast, and precipitation amounts were largely underforecast.

The sources of the errors noted here are difficult to determine. Contributing factors could include the model itself (e.g., model topography, moisture initialization), the derivation of the AIVs, and the verification data quality. One important result of this study is the recognition of the need for the AIV algorithms to reduce model data to the actual station elevation. Areas of improvement could be better model resolution, parameterizations, and initializations, as well as the AIV algorithms and quality control of the verification data. Overall, we are encouraged by the results of this first evaluation.

1. INTRODUCTION

The Federal Aviation Administration (FAA) has recently developed the Aviation Weather Development Program to improve aviation weather services. As part of that program, the Forecast Systems Laboratory's (FSL) Aviation Division is developing the Aviation Gridded Forecast System (AGFS). The AGFS is an interactive information-processing system that will generate analyses and gridded forecasts of state-of-the-atmosphere variables (SAVs) (e.g., temperature, winds, pressure) and aviation-impact variables (AIVs) (e.g., clouds and visibility). These analyses and forecasts will be used to generate products designed for the aviation community, and will help the FAA and the National Weather Service (NWS) provide better services.

The AGFS is designed to be a complete system [see Sherretz (1991) for a detailed description of the AGFS]. The AGFS ingests meteorological observations, generates gridded SAVs and AIVs with very high resolution in space and time, and allows forecasters to interact with forecast tools and with the gridded forecasts those tools generate.

The Aviation Division has created a verification program to assist in deriving AIVs from the output of meteorological models. The objective of this program is to provide atmospheric modelers with statistical measures of the accuracy of forecasts to help develop and improve aviation forecasts. This program will continue during the development of the AGFS and its implementation, planned for 1996.

A first evaluation (hereafter Exercise 1) was conducted to assess the current state of atmospheric models and to provide a baseline for future verifications of model AIV forecasts. Initially, the models were not able to provide forecasts for all requested AIVs at the resolutions listed in the AGFS plan. However, as the models are developed, the ongoing evaluation of each model should provide insight into how to generate and improve those forecasts. This exercise was exploratory in nature (Flueck and Brown, 1993), and was not designed to provide any statistical inference between models, but again is just to determine "where we stand" in the development of SAVs and AIVs.

Exercise 1 evaluated model forecasts over a 10-day period, from 1-10 April 1991; see Cairns (1992) for a short summary. The analyses and forecasts were made for select points across the contiguous United States. These 10 days included snow storms, wind storms, and convective events (see Appendix A). Not all models generated analyses and forecasts for all 10 days. This report describes the exercise in detail, and presents the results of the evaluation of model-generated SAVs and AIVs.

2. MODEL AND ANALYSIS DESCRIPTIONS

Four analysis and forecast models were involved in Exercise 1: 1) the Mesoscale Analysis and Prediction System (MAPS) Hybrid, FSL (Benjamin et al., 1991a,b); 2) the Eta

model, FSL and National Meteorological Center (NMC) (Mesinger et al., 1988); 3) the Local Analysis and Prediction System (LAPS), FSL (McGinley, 1989); 4) the Colorado State University (CSU) Regional Atmospheric Modeling System (RAMS), CSU (Tripoli and Cotton, 1982; Tripoli and Cotton, 1989).

2.1. MAPS

MAPS is a data assimilation system with objective analysis and forecast model components that has been run with a 3-h update frequency. Each 3-h analysis combines current observations (from rawinsondes, surface stations, profilers, and commercial aircraft) with the previous 3-h MAPS forecast valid at the current time. Another unique aspect of MAPS, besides the high-frequency updating, is the use of a hybrid isentropic-σ vertical coordinate. Isentropic coordinates allow improved resolution of fronts, the tropopause, and associated gradients of wind and moisture. All MAPS components (observation ingest, observation quality control, objective analysis, and forecast model) are configured in the hybrid coordinate system.

The version of MAPS evaluated for Exercise 1 was the operational version as of April 1991 (80-km horizontal resolution, 18 vertical levels, smooth terrain, convective and stable precipitation with no microphysics, no boundary layer effects, and no radiation). The current test version of MAPS has 60-km horizontal resolution, has 25 vertical levels using a new hybrid vertical coordinate and much more detailed terrain, and has added a boundary-layer parameterization, a much-improved boundary-layer resolution, surface fluxes of heat and moisture and momentum, and an initialization procedure that gives smoother short-range forecasts. Earlier versions of MAPS are described in Benjamin et al. (1991a,b).

It should be noted that the MAPS surface analyses and forecasts evaluated in Exercise 1 are derived from the MAPS hybrid model, and are separate from the MAPS surface assimilation system (Miller and Benjamin 1991, 1992).

2.2. Eta

The Eta model is a hydrostatic model that employs the so-called eta vertical coordinate. This coordinate treats the topography in a stair-step fashion so that the eta surfaces are quasi-horizontal. The purpose of this coordinate system is to reduce the error in the pressure gradient force term in the momentum equations. This term is likely to be large in steeply sloping coordinate surfaces in the sigma coordinate system. The model run on a national domain with a resolution of 30 km in the horizontal (horizontal grid 151 x 231) and 30 levels in the vertical. The top of the model is at about 19.5 km.

The initial conditions for the Eta runs come from the initial fields for the Nested Grid Model (NGM) interpolated to the Eta coordinates. The boundary conditions are taken from forecasts of NMC's global spectral medium-range forecast model. The Eta model includes the following physical packages: 1) convection: Betts and Miller (1986) convective

parameterization; 2) radiation: NMC's version of the Goddard Laboratory for Atmospheres (GLA) radiation scheme; and 3) boundary layer: Mellor-Yamada turbulent kinetic energy formulation level 2.0 for the surface layer and level 2.5 above.

The integration of the primitive equations is performed on a semi-staggered E grid using the time split-explicit approach: forward-backward time differencing for the adjustment stage and a modified Euler-backward time differencing for the advection stage. A time step of 70 seconds is used for the advection stage. An implicit backward time scheme is used to compute the vertical diffusion for reasons of model stability. For additional information and references, see Janjić (1990).

2.3. LAPS

At the time of Exercise 1, LAPS was an analysis system with no predictive capabilities. Its main goal is to assimilate many different data types into a fine horizontal-scale grid, which will be used as the initialization for a mesoscale predictive model (e.g., RAMS). The LAPS system for Exercise 1 comprised 57 x 57 grid points in the horizontal, and 17 vertical levels in pressure coordinates. The vertical resolution was 50 mb with the top level in LAPS at 100 mb. The horizontal grid spacing was 10 km, resulting in a local domain of about the size of Colorado.

LAPS produces analyses every hour, using its previous analysis as a first guess for the surface field. It also uses the MAPS model data as the first guess for its upper-air grid points. In addition to the traditional data sources [e.g., surface aviation observations (SAOs)], LAPS makes use of data from Doppler radar, a surface mesonet, profilers, and satellite data. The result is a detailed three-dimensional analysis of SAVs as well as some additional parameters infrequently found in other systems.

2.4. RAMS

A goal of the RAMS modeling component of this exercise was to provide short term (0-12 h) forecasts on a model grid exhibiting relatively fine horizontal grid spacing (2-10 km). Exercise 1 called for the RAMS to be run in four separate regions of the country, chosen for their differences in underlying terrain and expected weather conditions (see Fig. 1). Because of the preliminary nature of the exercise and constraints placed on the design due to data input, limitations resulting from the current implementation of physical parameterizations used in the model, or computer resources, the runs were somewhat limited in that the full model capabilities could not be utilized. This situation should be remedied by the acquisition of more powerful computers that will be on line in time for the next exercise.

The RAMS model design for Exercise 1 took the following form. The model equation set and the set of physical parameterizations used were identical for the four separate domains. The model was run on a terrain following the σ_z coordinate system using the fully compressible nonhydrostatic equations of motion. The vertical grid resolution was

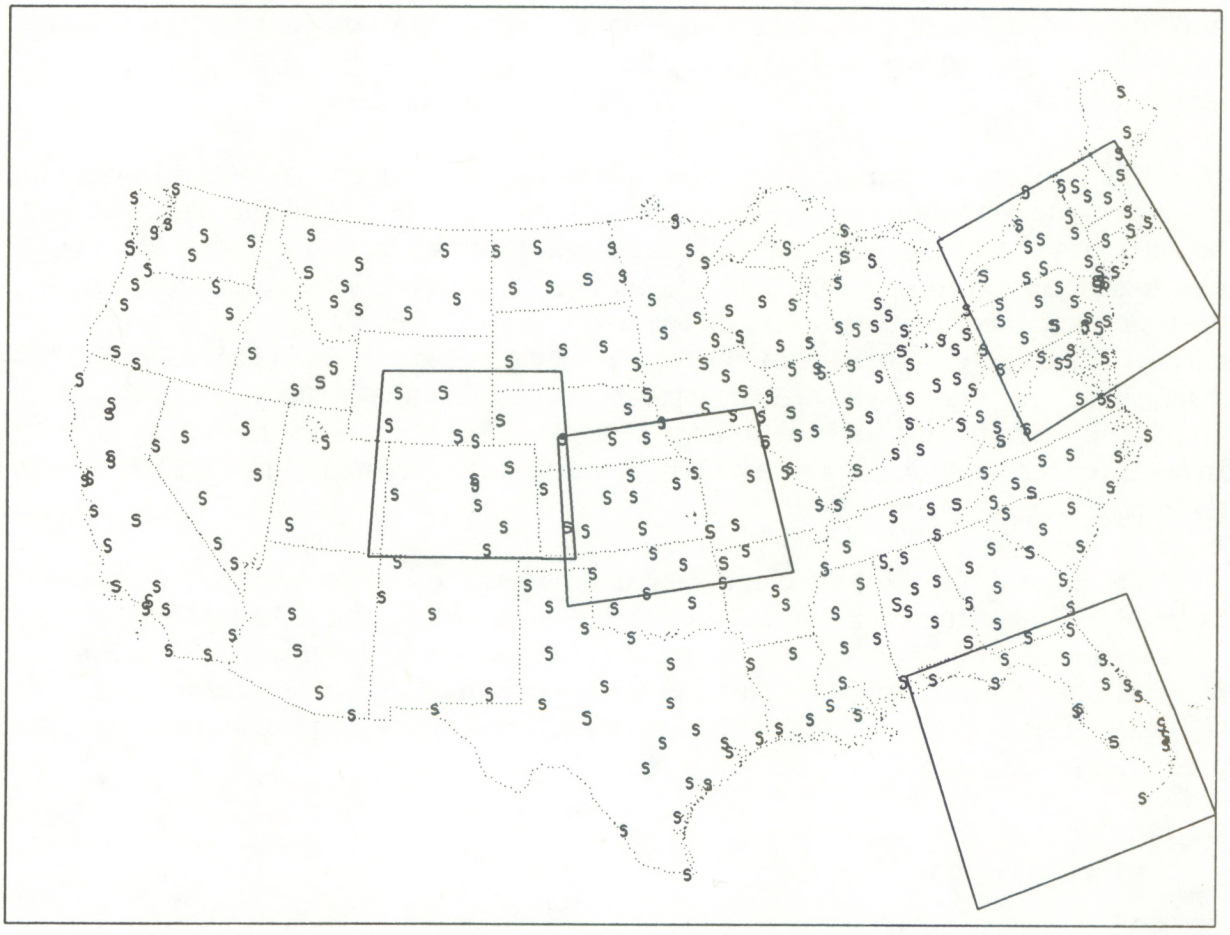

Fig. 1. Surface (S) station locations, national domain, and four WFO domain forecast and analysis areas.

stretched from 300 m near the surface to 750 m aloft. Thus the first model level above ground was approximately 150 m. The physical parameterizations include the radiation scheme of Chen and Cotton (1983); an 11-layer soil model described by Tremback and Kessler (1985), which differentiates the surface fluxes between land and water; a bulk ice/liquid microphysics package (Flatau et al., 1989); and an implicit cumulus parameterization scheme effective at a horizontal grid spacing exceeding 20 km. The lateral boundaries were treated with the Davies (1976) condition applied at the five outer grid points using MAPS analyses, and a solid wall served as the top boundary condition. The model was initialized 6 h before a given 12-h forecast period using MAPS analyses. Thus, any errors in the MAPS analyses at the off rawinsonde times were inherited in the initial RAMS input. This 6-h period was used to spin up the vertical motion in the model so that precipitation had sufficient time to form before a given forecast period.

The primary differences between the simulations were the grid resolution, underlying topography, and time step. For the two western domains, a nested grid approach was used. The outer grid consisted of an *x*-*y*-*z* 45x45x28 domain with a grid spacing of 70 km. The inner domains consisted of a 35x32x28 grid for both the Colorado and Kansas regions. A 3-to-1 ratio was used to generate an grid spacing of approximately 22 km on the inner domain. The underlying topography for these domains was identical to the MAPS topography. This was done so that comparisons between the MAPS analyses and the interpolated gridded fields could be made to determine if any bias in the fields may have resulted from the interpolation process. The forecasts were prepared using data from the fine grid in each case.

The basic strategy for the two eastern domains was to employ a single *x*-*y*-*z* grid of 45x45x28. The horizontal grid spacing for each domain was 30 km. The topography for these two domains was obtained from a standard 10-min data set (approximately 20 km horizontal resolution). A standard land use file was also used to differentiate areas of land and ocean. This distinction is critical for successfully simulating land/sea breezes along the Eastern Seashore.

3. VERIFICATION DATA

Exercise 1 verified the SAVs and AIVs listed in Table 1. Three primary observation data sources were selected for the verification data: SAO, upper-air rawinsonde (UPA), and vertical wind profiler (PRF). Raw observation data were obtained from the Facility Division at FSL. This section describes the verification data collection and quality control.

3.1. Surface Data

SAO stations conduct manual observations (i.e., no automated observations) 24 h a day. There were 318 such sites in the 48 conterminous United States (CONUS) at the time of the exercise (Fig. 1). The reasons for selecting were 1) to avoid any controversy over

Table 1. State-of-the-Atmosphere Variables (SAVs) and Aviation-Impact Variables (AIVs) Evaluated in Exercise 1

Variable Name	Units
SURFACE	
SAVs	
Altimeter Setting	in. of mercury
Temperature	°F
Dewpoint Temperature	°F
Wind Speed	kt
Wind Direction	degrees
AIVs	
Cloud Bases	hundreds of ft
Cloud-Top Height	hundreds of ft
Cloud Amount	coded (e.g., CLR, SCT)
Ceiling	hundreds of ft
Visibility	mi
Obstructions to Visibility	coded (e.g., H, K)
Precipitation Occurrence	yes/no
Precipitation Phase	coded (e.g., 1=liquid)
Precipitation Type	coded (e.g., RW, S)
Precipitation Amount	in./h or /3-h
RAWINSONDE	
SAVs	
Height	m
Temperature	°C
Dewpoint Temperature	°C
Wind Speed	kt
Wind Direction	degrees
Wind U,V	$m\ s^{-1}$
PROFILER	
SAVs	
Wind U,V	$m\ s^{-1}$
Wind Speed	kt
Wind Direction	degrees
Vertical Wind	$cm\ s^{-1}$

automated observation accuracy 2) because automated observations did not include all of the needed parameters such as cloud observations above 12,000 ft or present weather and 3) to avoid biasing day-time statistics by using stations that are closed at night. No special observations are used in the verification, only hourly observations.

Two quality control checks were performed on the observations. 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. Also, only the value was declared to be bad, not the entire observation. The first check was a simple check to make sure values were within acceptable ranges (e.g., $-90°F < Temp < +140°F$). For the second check, extreme values were printed and subjectively checked for time and space consistency. For example, if a station had an extreme temperature of 110°F, it was checked to see if observations taken before and after at that station had similar temperatures. Also, nearby stations were subjectively checked to determine if the extreme value was plausible. This second check revealed that 79 of approximately 72,000 observations had an error in winds, temperature, dewpoint, or altimeter setting.

Although the SAOs contained nearly all parameters to be verified, two parameters were needed in the verification that were unavailable from the SAO data: precipitation amount and cloud-top height.

Precipitation data were obtained from the National Climatic Data Center (NCDC) archives in Asheville, North Carolina. Many SAO sites had collocated precipitation gauges, but many did not. A program was written to select the closest gauges to the SAO site and let the user subjectively decide which gauge was most representative. This decision was easily made by looking at the descriptive name of the rain gauge location. In those cases, where a gauge was not selected, verification of precipitation amount was not performed for those sites. Also, some gauges listed as "WSFO AIRPORT" differed from the SAO location by as much as 0.2° latitude or longitude. Of the 318 SAO sites, 237 (74.5%) had precipitation data. A total of 196 (61.6% of 318) of the precipitation sites were collocated within 0.2° latitude or longitude of the SAO site. The remaining 41 sites more than 0.2° from the SAO site are still considered to be representative.

Cloud-top heights were derived from satellite infrared brightness temperatures (BT). The digital BT data were examined in a 0.2° latitude by 0.2° longitude area centered over the SAO site. The coldest pixel was chosen for the cloud-top temperature. This temperature was then compared to the nearest (in time and space) sounding to obtain a cloud-top height. Finally, the cloud-top height was compared to the observed cloud bases at that station to ensure that it was above the highest cloud base. The cloud was required to have a default thickness of at least one reportable value (i.e., low clouds, 500 ft; middle and high clouds, 1000 ft). If the cloud-top height could not be determined using the above method, the height was considered missing and was not used in the verification. It should be noted that this method of determining cloud-top heights is highly subject to errors, some of which could be due to misnavigation of the satellite imagery and incorrect BT due to cloud transmissivity.

3.2. Rawinsonde Data

Approximately 75 rawinsonde stations were available in the CONUS (Fig. 2). Only balloon launches at 0000 and 1200 UTC are used. Although all mandatory and significant level data were retrieved from the archives, only mandatory data were used in the Exercise 1 verification.

For quality control of rawinsonde data, a hydrostatic check was performed on each sounding. Temperature, dewpoint, or height values found to be in error were changed (when possible) to make them hydrostatically correct. This resulted in only two changes to the approximately 1400 rawinsonde observations (RAOBs). No corrections for radiative effects or rawinsonde type were applied to the RAOBS. (Rawinsonde and radiosonde are interchangeable terminology. Rawinsonde refers to the entire upper-air sounding data [pressure, temperature, moisture, and winds], and radiosonde typically refers to the instrument attached to the weather balloon that houses the temperature, pressure, and humidity sensors. The NWS prefers the term rawinsonde.) The only check applied to the winds was to ensure that the wind speed was less than 250 kt. No horizontal (buddy) check was performed on the RAOB data. No account was made for the drift of the balloon.

There has been much discussion about the accuracy of RAOB data, and in particular data from the moisture sensor at high levels. For a more complete review of the complications involved in using RAOB data for analysis and forecast verification, see Schwartz and Doswell (1991).

3.3. Profiler Data

The wind profiler is a rather new technology (Hogg et al., 1983; Weber et al., 1990). Only six wind profiler stations had been installed (Fig. 2) when the Exercise 1 data set was collected. Winds are reported every 250 m at 6-min intervals, and the observations are combined into a 1-h average. The hourly averaged data set was used for the verification data. Since profiler wind data are quality controlled before they are released, no additional checks were performed (Brewster and Schlatter, 1986; Brewster, 1989). The quality control method used does occasionally allow erroneous data.

3.4. 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

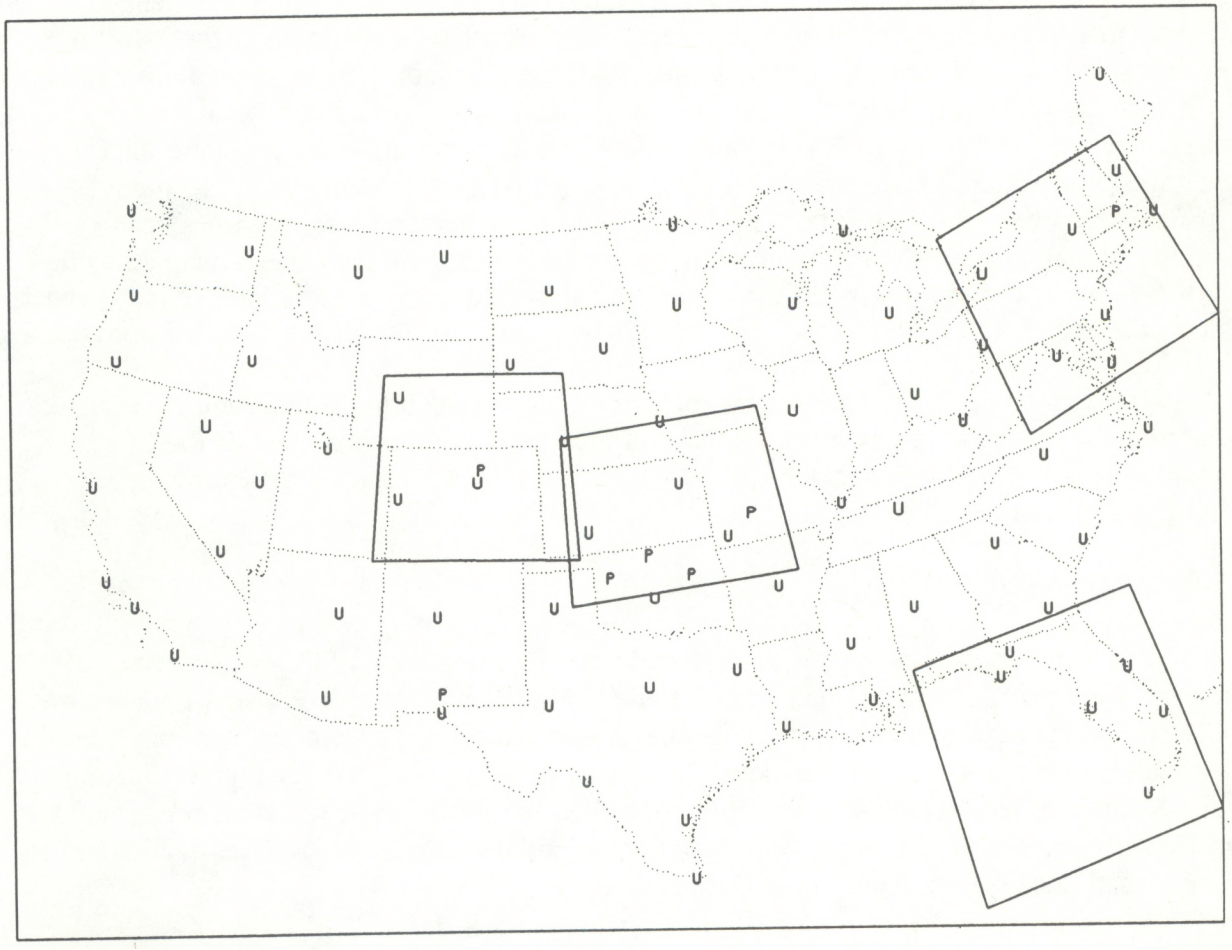

Fig. 2. Upper air (U) and profiler (P) station locations, national domain, and four WFO domain forecast and analysis areas.

example, two rawinsondes are attached to the same balloon, and the differences in their measurements are computed. Table 2 lists the surface and upper-air instrument RMS differences (Hoehne, 1980; NWS, 1991a,b; NOAA, 1991). These numbers should be kept in mind when evaluating the relative accuracy of the numerical models. Even though a perfect 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 (see Daley, 1991) due to the spacing of the observational data can affect the results.

Table 2. Instrument Root Mean Square (RMS) Differences

UPPER AIR	
Pressure:	2.0 mb
Temperature:	0.54°C - 0.68°C
Relative Humidity:	2.2%
Dewpoint Depression:	2.7°C - 3.4°C
Height:	15.3 m - 16.3 m
Wind Speed:	6.0 kt
Wind Direction:	14° at 10 kt, 2° at 120 kt
SURFACE	
Altimeter:	0.02 in. Hg
Temperature:	1°C (1.8°F)
Dewpoint Temperature:	from 1.1°F for Td > 32°F & T-Td < 11°F to 7.9°F for Td < 32°F & T-Td > 54°F
Wind Speed:	2 kt or 5% (whichever is greater)
Wind Direction:	5° for wind speed > 5 kt
Ceiling:	50 ft up to 12,000 ft
Precipitation:	0.3 mm for 8-in. stick gauge 0.02 in. for weighing gauge
Visibility:	Subjective
Cloud Amount:	Subjective

4. FORECASTS AND ANALYSES

4.1. Model Forecasts and Analysis Types

The numerical model data were interpolated from the usual gridded form to the observation locations described in Section 3. The model data were constructed to mirror the observational format. For example, if there is a low overcast cloud deck, observers cannot report any clouds above that deck because their vision is blocked by the low clouds.

Models, on the other hand, can "see" cloud decks above a lower overcast deck. However, for the purposes of this exercise, these data were not verified.

In addition to the SAVs that were directly available from the model output, the AIVs had to be derived from the model data. In some cases, a model did not yet have the capability to produce a particular AIV (e.g., visibility) and thus was not evaluated for that parameter. In other cases, a model could not provide the AIV in the same categories as observed. For example, precipitation can be observed as rain, drizzle, snow, or hail. Some models could indicate only whether or not precipitation was occurring, and others could discern liquid, freezing, and frozen precipitation. Thus, some of the AIV verifications of one model will not be as complete as those of another model. Table 3 lists the SAVs and AIVs that each model could diagnose. It is anticipated that for Exercise 2, nearly all parameters will be produced by the models, including some additional variables such as icing and turbulence.

4.2. Areas and Frequencies

The overall validation area was defined by approximately 25°-50°N and 65°-130°W, which encompasses the CONUS. National models (MAPS, Eta) produced forecasts within their grid region. For example, the MAPS hybrid model ran between the area defined by 22.84°N, 120.5°W (lower left corner) to 45.99°N, 60.83°W (upper right corner). Within these areas, MAPS and Eta were required to interpolate point forecasts for the surface, upper air, and profiler stations. Figure 1 illustrates the MAPS model domain, which is a subset of the Eta model domain.

For WFO domain analysis and forecast systems (LAPS, RAMS), four subareas were defined: 1) Colorado, 2) Central Plains, 3) Florida, and 4) the Northeast (Fig. 1). Point forecasts were interpolated to those surface, upper-air, and profiler stations within the four subareas. LAPS analyses were generated for the Colorado region only.

The frequencies of the model forecasts and system analyses were similar to those stated for the AGFS (Sherretz, 1991), but were modified for Exercise 1. LAPS analyses were generated hourly throughout the 10-day period. RAMS forecasts were valid at 1-h intervals out to 12 hours. Only selected cases were run. These are listed in Table 4.

For Eta and MAPS, the valid times varied. The MAPS hybrid model runs on a 3-h cycle. MAPS produced 3-h and 6-h forecasts at asynoptic times (e.g., 0600 UTC), and forecasts out to 12 h at 0000 and 1200 UTC. The valid times for the Eta model forecasts were for hours 1 through 6 at 1-h intervals, and for hours 9 through 24 at 3-h intervals. Eta model forecasts were generated for the 24-h period once a day beginning at 0000 UTC.

Table 3. SAVs and AIVs by Model for Exercise 1

	MAPS	LAPS	Eta	RAMS
Surface				
Wind Direction	yes	yes	yes	yes
Wind Speed	yes	yes	yes	yes
Temperature	yes	yes	yes	yes
Dewpoint	yes	yes	yes	yes
Altimeter	yes	yes	yes	yes
Precip. Occurrence	yes	no	yes	yes
Precip. Phase	yes	no	yes	yes
Precip. Type	no	no	no	no
Precip. Amount	yes	no	yes	yes
Visibility	yes	no	no	no
Obstruction to Vis.	yes	no	no	no
Ceiling	yes	yes	yes	yes
Cloud Heights	yes	yes	yes	yes
Cloud Amount	yes	yes	yes	no
Cloud Top	yes	yes	yes	yes
Rawinsonde				
Height	yes	yes	yes	yes
Temperature	yes	yes	yes	yes
Dewpoint	yes	yes	yes	yes
Wind Speed	yes	yes	yes	yes
Wind Direction	yes	yes	yes	yes
Profiler				
U-Component	yes	yes	yes	yes
V-Component	yes	yes	yes	yes
W-Component	no	yes	yes	yes

Table 4. Model Run Dates and Times for RAMS. (The numbers in parentheses are missing data times that were not available for verification.)

Date	Times (UTC)
Colorado	
2 April 1991	0000 - 1200 (0000,1100,1200)
8 April 1991	0000 - 1200 (0600,1200)
8 April 1991	1200 - 2400
Central Plains	
2 April 1991	1200 - 2400 (1200)
3 April 1991	0000 - 1200 (0000)
4 April 1991	0000 - 1200 (0700,0800)
7 April 1991	0000 - 1200
8 April 1991	1200 - 2400
Florida	
4 April 1991	0000 - 1200
5 April 1991	1200 - 2400
Northeast	
2 April 1991	0000 - 1200
4 April 1991	0000 - 1200
5 April 1991	1200 - 2400
10 April 1991	1200 - 2400

4.3. Derivation of Aviation-Impact Variables

Numerical models analyze and forecast SAVs, such as winds, temperature, and moisture. Weather variables that affect aviation procedures are typically not forecast by the model and must be derived from the SAVs. Many methods can be used to derive these AIVs. One method currently used at NMC is the model output statistics (MOS) approach.

One of the main objectives of the verification exercise was to evaluate the accuracy not only of the numerical models but also of the AIV derivation methods. For Exercise 1, each modeling group developed its own methods of AIV derivation. Table 5 gives a brief description of the methods used, and Appendix B provides an expanded description of the AIV derivations.

Table 5. AIV Derivation Methodology

Cloud Base/Ceiling/Top

LAPS -- Explicitly analyzed by LAPS.
RAMS -- Look for liquid cloud water, which is an explicit variable in RAMS.
MAPS -- Look for relative humidities higher than RH_0 (see list in Cloud Amount section that follows in this table).
Eta -- Look for relative humidities of 66%, 50%, and 40% for low, middle, and high clouds respectively.

Cloud Amount

LAPS -- Explicitly analyzed by LAPS.
RAMS -- Not produced.
MAPS -- $F_c = ((RH - RH_0) / (1.0 - RH_0))^{**} p_c$, where F_c is the cloud fraction (0 to 1), RH is the relative humidity, and rh0 and pRH are given by:

height agl (m)	RH_0	p_c
20000	.675	1
4500	.675	1
3000	.725	1
500	.725	1
200	.900	2
0	.975	3

Eta -- $CA = ((RH - RH_c) / (100 - RH_c))^{**} 2$, where RH_c is the critical RH value (66%, 50%, or 40% for low, middle, and high clouds). CA varies from 0 to 1.

Precipitation Phase/Type

LAPS -- Not produced.
RAMS -- Check surface temperature above or below 32°F to determine rain or snow.
MAPS -- Check vertical sounding of wet bulb temperature and relative humidity. Use a decision tree to determine precipitation phase.
Eta -- Check surface temperature above or below 32°F to determine rain or snow.

Table 5. (continued)

Precipitation Amount

LAPS -- Not produced.
RAMS -- Explicitly modeled.
MAPS -- Explicitly modeled.
Eta -- Explicitly modeled.

Visibility/Obstructions to Visibility

LAPS -- Not produced.
RAMS -- Not produced.
MAPS -- Regression equation using precipitation rate, relative humidity, wind speed, dewpoint, and snow fraction.
Eta -- Not produced.

5. STATISTICAL MEASURES

The interpolated point forecasts for the variables were verified at the previously mentioned stations and at valid time intervals. Table 6 lists the statistical measures generated for each variable for surface, upper air, and profiler. The measures chosen are in line with those typically used in current model evaluations performed by the meteorological community. The statistical package SPSS Inc. was used to generate most of results. Additional programs calculated meteorological scores not available from SPSS Inc. Table 1 lists the verified variables and their corresponding units. Unless otherwise stated or displayed on a figure, these are the units used throughout this memorandum. These units were chosen on the basis of common usage, and the units that will be used in the AGFS.

In general, a statistic was computed whenever there was a matching observation and model analysis or forecast for a given time and station. If either piece of data was missing, no calculation was performed. The definition of missing data is straightforward for variables such as temperature and wind speed. But for other variables such as ceiling height or precipitation type and phase, the value is missing if no event occurred. For example, this can be true for scattered clouds (less than 0.6), which does not fit the definition of a ceiling, even though a cloud base is present. The ceiling is then declared to be missing. The important ramification of this is that such variables are verified only when both the observation and the model indicate that a ceiling existed. Thus, if the observed ceiling was 800 ft, and yet the model had only clear or scattered clouds, no verification took place. This has the advantage of verifying the actual ceiling height error without biasing the results due to good or bad cloud amount analyses and forecasts by the model. To evaluate model

Table 6. List of Statistical Measures by Data Source and Variable

SURFACE

Altimeter
- Arithmetic, Absolute, and Square Errors (by station for WFO domains)
- RMSE
- Boxplots Observation by Forecast
- Boxplots Forecast by Observation
- Scatterplot
- Distribution Bar Graph
- Map Distribution of Errors

Clouds
- Arithmetic, Absolute, and Square Errors of Cloud Top (by station for WFO domains)
- RMSE of Cloud Top
- Scatterplot of Cloud Top
- Contingency Tables of Low, Middle, High, and Total by Amount
- Contingency Tables of Low, Middle, High, and Total by Occurrence
- Distribution Bar Graphs of Low, Middle, High, and Total
- Skill scores (POD, FAR, CSI, TSS, HSS, etc.)

Obstructions to Visibility
- Contingency Tables of Obstruction Type (e.g., Haze, Smoke) 6 tables
- Distribution Bar Graphs of Obstruction Type
- Skill Scores (POD, FAR, CSI, TSS, HSS, etc.)

Precipitation
- Contingency Table of Precipitation Occurrence
- Contingency Table of Precipitation Phase
- Contingency Tables of Precipitation Types (e.g., Rain, Snow) 12 tables
- Arithmetic, Absolute, and Square Error of Precipitation Amount (by station for WFO domains)
- RMSE of Precipitation Amount
- Boxplots of Precipitation Amount Observation by Forecast
- Boxplots of Precipitation Amount Forecast by Observation
- 2 Scatterplot of Precipitation Amount
- Map Distribution of Amount Errors
- Distribution Bar Graph of Precipitation Type
- Distribution Bar Graph of Precipitation Phase
- Distribution Bar Graph of Precipitation Amount
- Skill Scores (POD, FAR, CSI, TSS, HSS, etc.) for Type, Phase, Occurrence

Table 6. (continued)

Temperatures

- Arithmetic, Absolute, and Square Error of Temperature (by station for WFO domain)
- RMSE of Temperature
- Boxplots of Temperature Observation by Forecast
- Boxplots of Temperature Forecast by Observation
- Scatterplot of Temperature
- Arithmetic, Absolute, and Square Error of Dewpoint Temperature (by station for WFO domain)
- RMSE of Dewpoint Temperature
- Boxplots of Dewpoint Temperature Observation by Forecast
- Boxplots of Dewpoint Temperature Forecast by Observation
- Scatterplot of Dewpoint Temperature
- Arithmetic, Absolute, and Square Error of Dewpoint Depression (by station for WFO domain)
- Distribution Bar Graphs of Temperature and Dewpoint Temperature
- Map Distribution of Temperature and Dewpoint Temperature Errors

Winds

- Arithmetic, Absolute, and Square Error of u and v (by station for WFO domain)
- RMSE of u and v and RMSVE
- Boxplots of u Observation by Forecast
- Boxplots of u Forecast by Observation
- Boxplots of v Observation by Forecast
- Boxplots of v Forecast by Observation
- 2 Scatterplots of u and v
- Absolute Direction Error (by station for WFO domains)
- Contingency Table of Direction
- Arithmetic, Absolute, and Square Speed Error (by station for WFO domains)
- RMSE of Speed
- Boxplots of Speed Observation by Forecast
- Boxplots of Speed Forecast by Observation
- Scatterplot of Speed
- Pie Chart of Speed by Direction
- Distribution Bar Graph of Direction
- Distribution Bar Graph of Speed
- Map Distribution of Wind Direction and Wind Speed Errors
- Skill Scores (POD, FAR, CSI, TSS, HSS, etc.) for Wind Direction & Speed

Table 6. (continued)

Visibility/Ceiling Height
- Contingency Table of Visibility
- Contingency Table of Ceiling Height
- Contingency Table of Visibility/Ceiling Height
- Mean, Standard Deviation, and Median of Visibility
- Mean, Standard Deviation, and Median of Ceiling Height
- Distribution Bar Graph of Visibility
- Distribution Bar Graph of Ceiling Height
- Distribution Bar Graph of Visibility/Ceiling Height
- Skill Scores of Visibility
- Skill Scores of Ceiling Height
- Skill Scores of Visibility/Ceiling Height

RAWINSONDE

Height
- Arithmetic, Absolute, and Square Errors (by station for WFO domains)
- RMSE
- Scatterplot

Temperatures
See Surface Data

Winds
See Surface Data

Vertical Plots of Mean Errors and Root Mean Square Error
- T, T_d, and Height
- u, v
- Speed, Direction

PROFILER

Winds
See Surface Data

Vertical Plots of Mean Errors and Root Mean Square Error
- u, v, w
- Speed, Direction

performance on ceiling occurrence, statistics were computed for cloud amount. Thus, in the above case, the model would be penalized for not having a ceiling when one existed in the observation.

Another exception to the rule of verifying all data is the stratification placed on some variables. Throughout this exercise, a wind direction was verified only if the *observed* wind speed was 10 kt or greater. Thus, even if the model wind speed was light (i.e., 0-9 kt), the direction was verified if the observed wind speed was at least 10 kt. This prevents a model from showing good direction results by underforecasting the wind speed and limiting the number of cases. This strategy was also used for observed visibilities and ceilings where the criteria were <10 mi and $<8{,}000$ ft, respectively.

For this report, low clouds are defined as those clouds with a base below 2 km (6500 ft), middle clouds bases are 2-6 km (6500 - 20,000 ft), and high clouds have bases above 6 km (20,000 ft). Visibility and ceiling are verified according to the standard categories used in aircraft operations (Table 7). For cloud amounts: 1) clear (CLR) < 0.1 cloud cover; 2) scattered (SCT) $= 0.1 - 0.5$; 3) broken (BKN) $= 0.6 - 0.9$; 4) overcast (OVC) > 1.0; and 5) obscured (X or OBSCD) is when the sky is completely hidden from surface-based phenomena (e.g., fog, rain, snow).

Table 7. Visibility (Vis) and Ceiling (Cig) Categories

Flight Rules	Vis (mi)	Cig (ft)
Visual Flight Rules (VFR)	$5 <$ vis	$3000 \leq$ cig
Marginal Visual Flight Rules (MVFR)	$3 \leq$ vis ≤ 5	$1000 \leq$ cig < 3000
Instrument Flight Rules (IFR)	$1 \leq$ vis < 3	$500 \leq$ cig < 1000
Low Instrument Flight Rules (LIFR)	$1 >$ vis	$500 >$ cig

5.1. Summary of Statistical Measures

The bias, mean absolute error, and root mean square error are easily defined using F_i as the forecasts, O_i as the observations, and n as the number of forecast/observation pairs. The rest of the measures discussed here are formulated on the 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). Other scores typically used in meteorological verification involve climatology. Long-term climatology was not collected, and it was felt that the 10-day sample climatology was insufficient to encourage the use of those statistical measures.

5.1.1. Bias or mean error

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

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

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.

5.1.2. 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

$$MAE = (1/n)\sum_{i=1}^{n}|F_i - O_i| .$$

The MAE range is from 0 to ∞, a MAE of 0 is desired.

5.1.3. 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

$$RMSE = [(1/n)\sum_{i=1}^{n}(F_i - O_i)^2]^{1/2} .$$

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 ∞, and values close to the MAE (or zero) indicate a desired forecast.

The root mean square vector error (RMSVE) is similar, and is designed for evlauting magnitude and directional eerors of the wind. RMSVE is defined as

$$RMSVE = [(1/n)\sum_{i=1}^{n}((u_{fi} + u_{oi})^2 + (v_{fi} + v_{oi})^2)]^{1/2} ,$$

where u and v are the components of the wind, and as before f and o denote forecst and observed, respectively.

5.1.4. Percent correct

Percent correct is just that -- the percentage of correctly forecast events. In a contingency table, this refers to the numbers added along the diagonal (or correct forecasts) divided by the total number of events. The values range from 0% to 100%; 100% represents desired forecasting.

5.1.5. Probability of detection

The probability of detection (POD), or prefigurance, is calculated from a contingency table. It is defined as the number of correct forecasts divided by the number observed in each category. In other words, it is the percentage of events that occurred that were correctly forecast. It is a measure of the ability to correctly forecast a particular category. The POD ranges from 0.0 to 1.0, and 1.0 represents a desired forecast of that category.

5.1.6. False alarm rate

The false alarm rate (FAR), or postagreement, is defined as the number of incorrect forecasts divided by the total number of forecasts for each category. That is, it is the percent of forecasts that did not verify for each category. The FAR ranges from 0.0 to 1.0; a FAR of 0.0 is desirable.

5.1.7. Critical success index

The critical success index (CSI), or threat score, is defined as the number of correct forecasts of a given category divided by the number of cases forecast and/or observed for that category. The range is from 0.0 to 1.0, and values close to 1.0 are desired. The CSI is sensitive to both false alarms and missed events, and therefore is more sensitive in situations where rare events are involved. However, it gives no credit for correct forecasts of the null event.

5.1.8. True skill statistic

The true skill statistic (TSS) compares the number of correct forecasts, minus those attributable to random guessing, to a hypothetical set of perfect forecasts. It ranges from -1.0 to 1.0, and a score of 1.0 is desired. The TSS attempts to measure the skill of a forecast against what one would obtain if the forecast were merely a random guess. It also

gives credit for the correct forecast of a null event, unlike the CSI. However, in cases where the null event dominates (e.g., tornadoes, thunderstorms), the TSS approaches the POD. This means that the TSS is vulnerable to hedging in rare event forecasts; if uncertain, the forecaster will score better by forecasting the null event. Another disadvantage of the TSS is that it is a complicated calculation when there are more than two categories.

5.1.9. Heidke skill score

Similar to the TSS, the Heidke skill score (HSS) attempts to remove any artificial skill due to pure chance, while still giving credit for correctly forecasting the null event. However, the HSS does not have the same problems as the TSS, because it does not encourage hedging and is easily calculated for a multiple category situation. It also ranges from -1.0 to 1.0, and 1.0 represents a perfect score. The HSS is always equal to or greater than the CSI, implying that it is giving credit for correct null event forecasts. For these reasons, the HSS is considered to be the best skill score.

6. MODEL RESULTS

6.1. MAPS

Appendix C contains the summaries of MAPS statistics for the SAVs and AIVs at the surface and at the rawinsonde levels of 850, 500, and 250 mb. Further details of the MAPS evaluation grouped by SAVs or AIVs for the surface and upper air are discussed here.

6.1.1. SAVs

Upper air

When analyzing the MAPS upper-air results, keep in mind the analysis and forecast interval and length of MAPS, as described in Section 4. MAPS produces an analysis and 3-h and 6-h forecasts every 3 h. At 0000 and 1200 UTC, the model is run longer for a 9-h and 12-h forecast. Additionally, the MAPS analysis uses the previous 3-h forecast for its first guess field.

Since RAOB data are available only at 0000 and 1200 UTC, only upper-air forecasts valid at those two times can be verified. In other words, all 12-h forecasts are verified, and only the 3-h forecasts beginning at 0900 and 2100 UTC can be verified. Likewise, the 6-h forecasts from 0600 and 1800 UTC are the only ones that can be verified.

These important facts are mentioned because some of the results for MAPS are better at 12 h or 6 h than at 3 h. This does not indicate that the MAPS model is better at longer forecasts. Rather, it is because of the assimilation cycle, initial times of the model runs, and the observation data that went into that initialization. For example, at 0000 and 1200 UTC,

a full suite of observational data is available, whereas at asynoptic time, only the Aeronautical Radio INCorporated (ARINC) Communications Addressing and Reporting System (ACARS) reports, SAOs, and profiler data are available for the update cycle. This means that asynoptic observations, assimilated by MAPS, does not help and might even hurt forecasts of certain fields.

Height. The analysis height errors show a pronounced negative bias in the middle and upper troposphere (Fig. 3), reaching a maximum of nearly -25 m at 300 mb. This error profile is virtually the same through the 12-h forecasts, except that it grows slightly more negative (Appendix C). The absolute errors also increase with the forecast time, and especially at 850 mb, where the error nearly triples over 12 h.

Note that although the 12-h MAE for all three levels is larger than the 0-h, 3-h, or 6-h forecast, the 6-h forecast is actually slightly better than the 3-h forecast. As pointed out earlier, this is a result of the initialization times and not the model performance. What this implies is that the 6-h forecast (from 0600 or 1800 UTC) is better than two consecutive 3-h forecasts from the same initial time. Larger errors at the shorter time periods (i.e., 3 h) may also be a symptom of gravity wave imbalances that occur early in the model forecast simulation.

Temperature and dewpoint temperature. The bias profile of analysis temperature errors (Fig. 3) shows that MAPS has virtually no bias at any level, and the largest error occurs at 1000 mb, where it grows to -2.5° at 12 h (Fig. 4). The profiles of RMSE (Figs. 5 and 6) also show that the 1000 mb level has the largest error. Note that by 12 h, the RMSE has nearly doubled at all levels (also see Appendix C).

Looking at the temperature errors in Appendix C, it is evident that the analysis errors are nearly doubled by the 3-h forecast, but then remain nearly constant through 12 h, although the 12-h forecast error is slightly worse than the 3-h forecast error. Again, the explanation lies in the necessity of verifying model forecasts valid at 0000 and 1200 UTC only. The 3-h forecast errors here are from runs initialized at 0900 and 2100 UTC. Since the analysis fields with which these runs were initialized did not have RAOB data, the analysis errors at these times were probably larger than those that did have RAOB data. Thus, the "jump" in the error from 0 to 3 h is not directly attributable to model behavior.

Similar to that of the temperature, the MAPS analysis of dewpoint temperature shows a near zero bias, and only slightly negative (dry) up to 400 mb (Fig. 3). However, a large dry bias exists at 300 mb. Cirrus clouds are usually found between 400 and 300 mb, so the dry bias at those levels could directly affect MAPS ability to diagnose cirrus clouds.

By 12 h, the MAPS model has corrected this dry bias up to 300 mb (Fig. 4). However, the RMSE of dewpoint temperatures (Figs. 5 and 6) shows that the errors are larger at the 12-h forecast than in the analysis. Thus, the improvement in the bias is somewhat deceiving. Also note that lower levels (below 600 mb) indicate a drying bias at

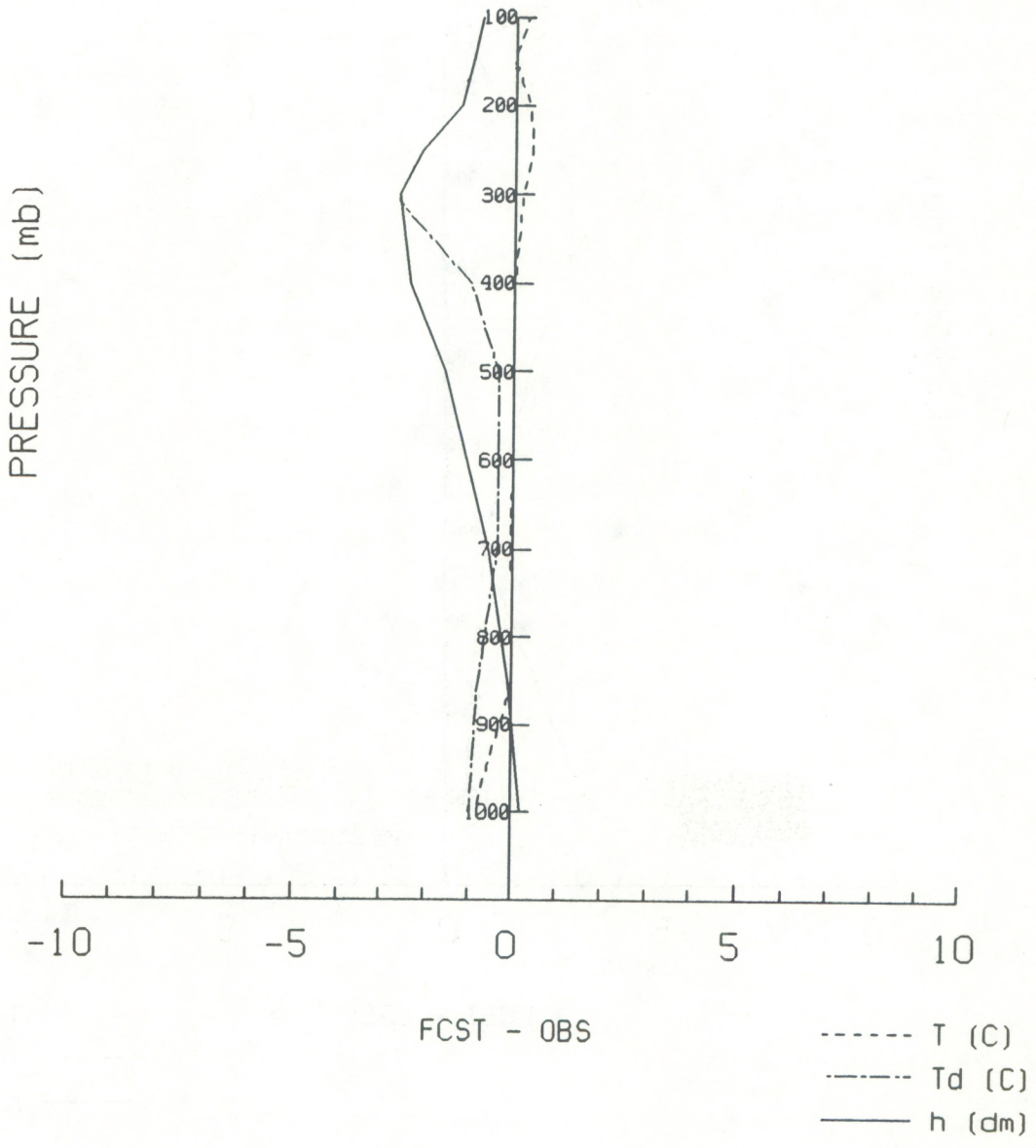

Fig. 3. Mean error (bias) of height (h), temperature (T), and dewpoint temperature (T_d) between MAPS analyses and RAOBS.

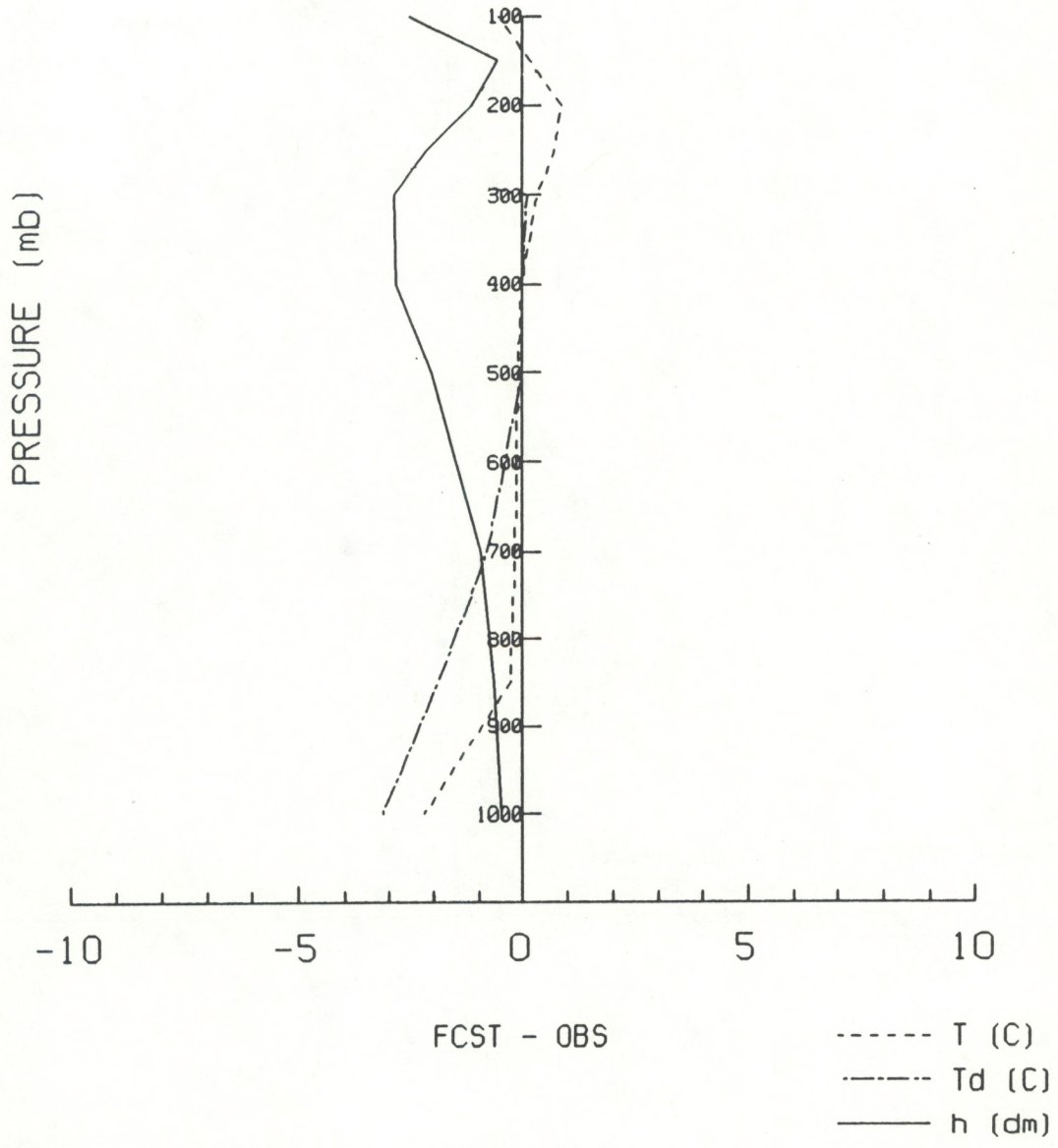

Fig. 4. Same as Fig. 3, except for MAPS 12-h forecast data.

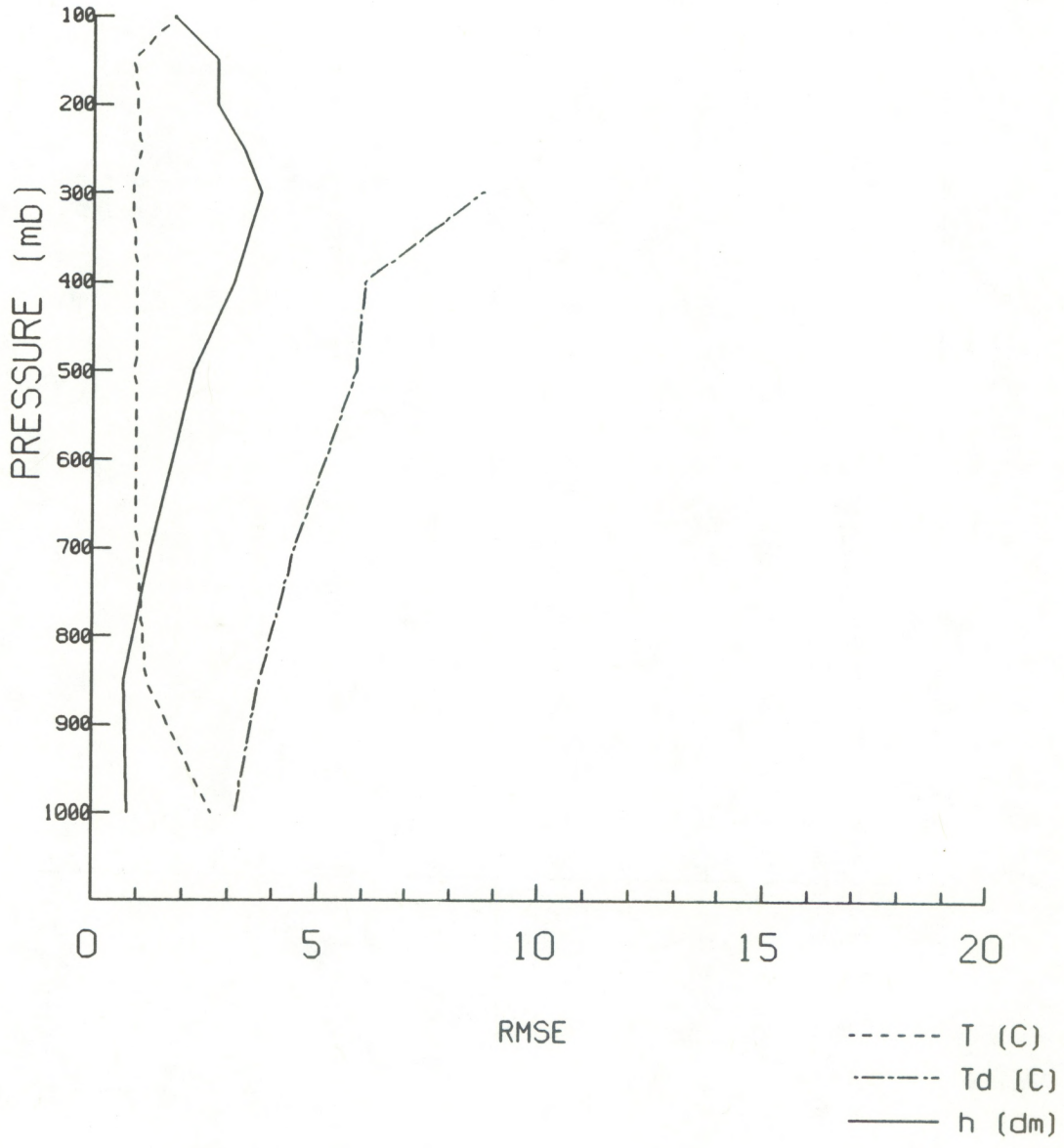

Fig. 5. Root mean square error (RMSE) of height (h), temperature (T), and dewpoint temperature (T_d) between MAPS analyses and RAOBS.

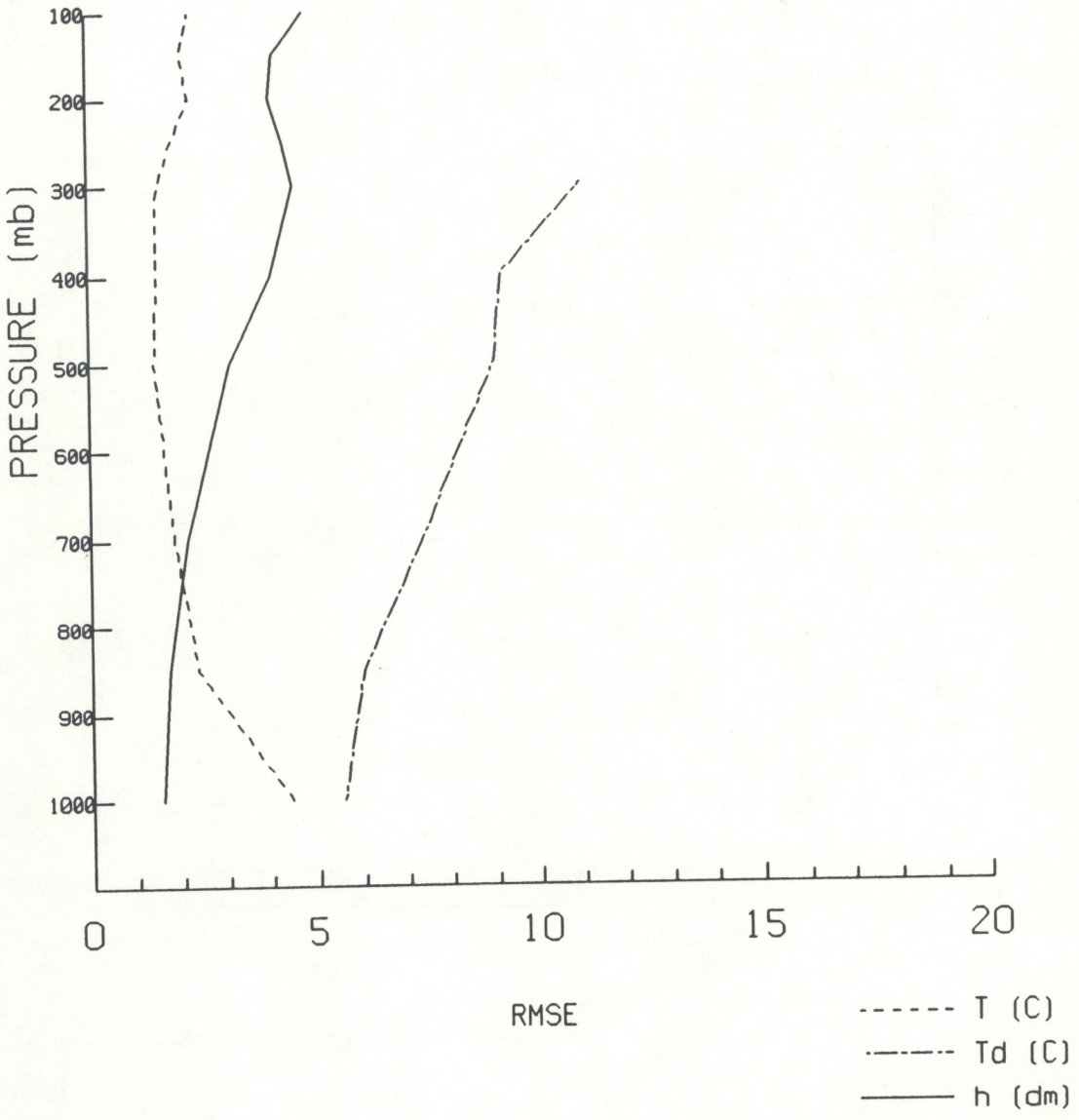

Fig. 6. Same as Fig. 5, except for MAPS 12-h forecast data.

12 h (Fig. 4). Although a portion of this bias can be attributed to the cold temperature bias at 1000 mb, the remaining levels show little or no temperature bias. Thus, it is believed that the model is drying at low levels as well. A comparison of the analysis and forecast scatterplots of 850 mb dewpoint (Figs. 7 and 8) also supports this conclusion. Note that while the analysis data were fairly well correlated, most of the forecast dewpoints are below the diagonal correlation line. (The diagonal line represents a perfect match between model and observed data.)

An examination of the 850 mb dewpoint errors (Appendix C) shows results similar to the temperature errors. The dramatic increase in error at 3 h is not real. However, note that the 12-h forecast is actually better than the 3-h forecast valid at the same time. (Although not shown, the 1000 and 700 mb errors at 12 h are relatively unchanged from 3 h.) This again is due to the difference in initialization times. The 12-h forecast was initialized with an analysis which had RAOB data; the 3-h forecast was not. In fact, aside from the RAOB data, there is no other upper-air source of moisture observations used by MAPS. Thus, the assimilation cycle is basically using the previous 3-h forecast of moisture as its analysis.

The 850 mb dewpoint errors indicates that a 12-h forecast initialized at synoptic time when RAOB data are available is better than four consecutive 3-h forecasts. As can be seen from the data in Appendix C, an analogous conclusion can be drawn for the 6-h forecast at 850 mb. It is worse than the 12-h forecast, but better than the 3-h forecast. This is very important since it is often assumed that a 3-h forecast is better than a 12-h forecast, regardless of the initialization time. According to the results presented here, it appears that it is better to use the 12-h forecast than the fourth consecutive 3-h forecast when considering moisture.

Winds. MAPS does very well at analyzing the upper-air winds, as shown in Fig. 9. The small absolute errors in the analysis are approximately doubled by 12 h (Appendix C).

As would be expected, the wind speed and direction errors are also quite low (Appendix C); the analyzed mean absolute speed errors are less than 4 kt and the mean absolute direction errors are under 10°. The analysis data show a near zero bias in wind speed (Fig. 9). But by 12 h, the 1000 mb winds appear to be too strong, and the winds near the tropopause are too light (Fig. 10).

As with the height and temperature data, the error at 3 h is nearly the same as the error at 12 h. This indicates that little is gained by using a 3-h forecast from 0900 or 2100 UTC as opposed to using the previous 12-h forecast. This does not mean that the 3-hourly assimilation cycle is of no value, since the analysis undoubtedly benefits from the assimilation cycle. However, the intermediate forecasts during the cycle appear to be only marginally better than the full 12-h forecasts at 0000 and 1200 UTC, except for the 250 mb level. This is expected since most of the asynoptic ACARS data are near this level.

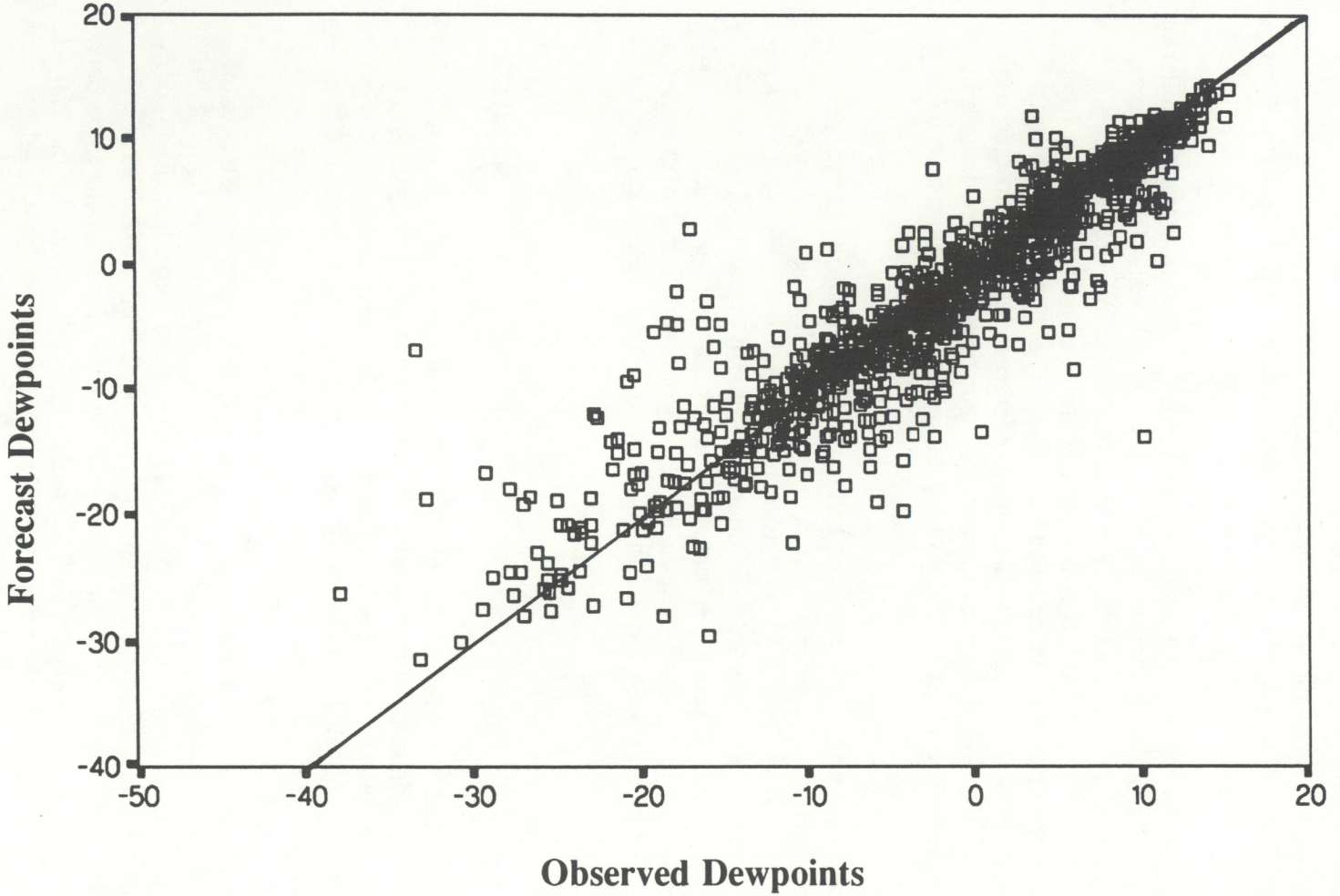

Fig. 7. Scatterplot of observed 850 mb dewpoint temperatures and MAPS analyses of dewpoint temperatures in °C.

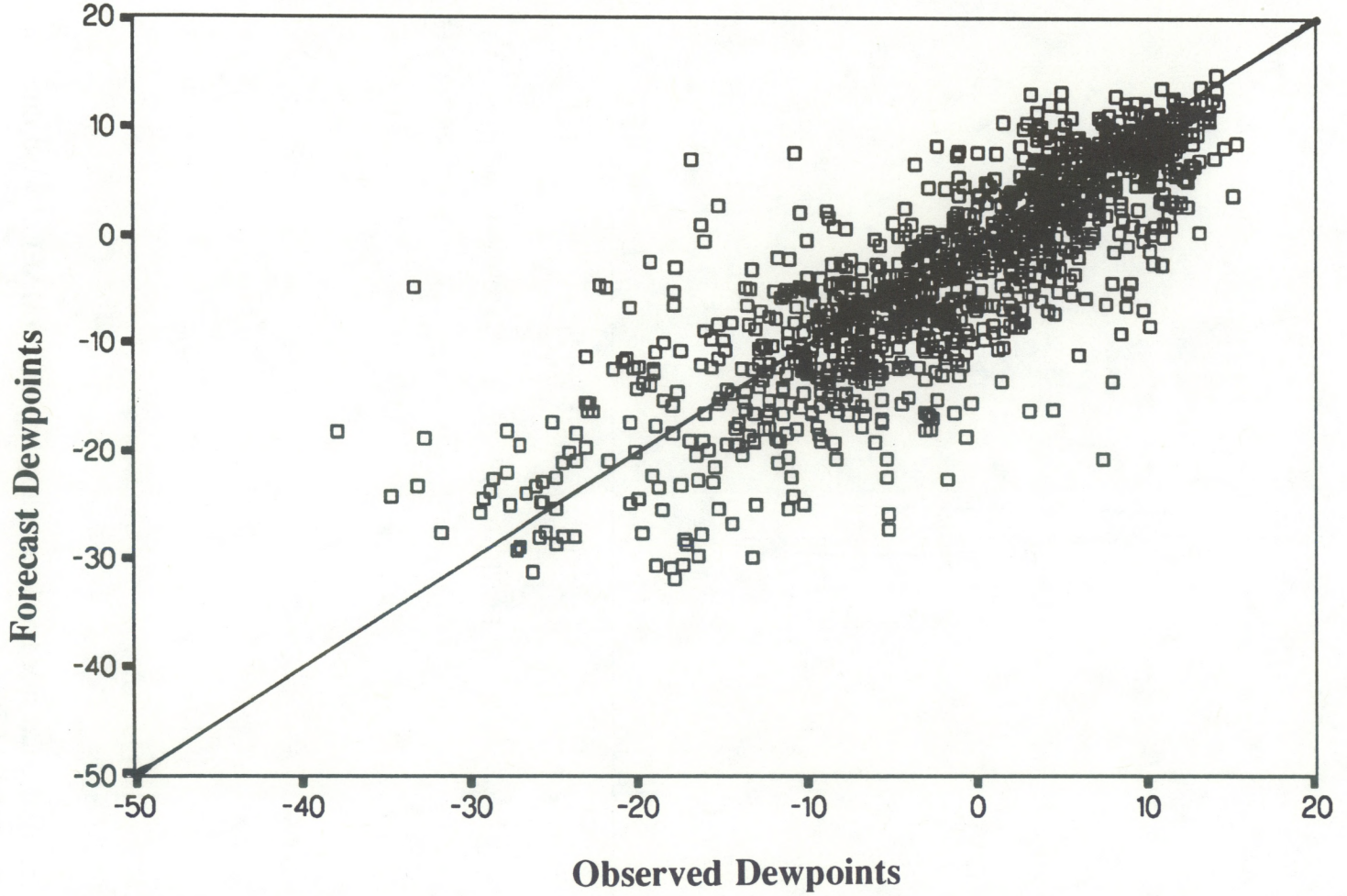

Fig. 8. Same as Fig. 7, except for MAPS 12-h forecast data.

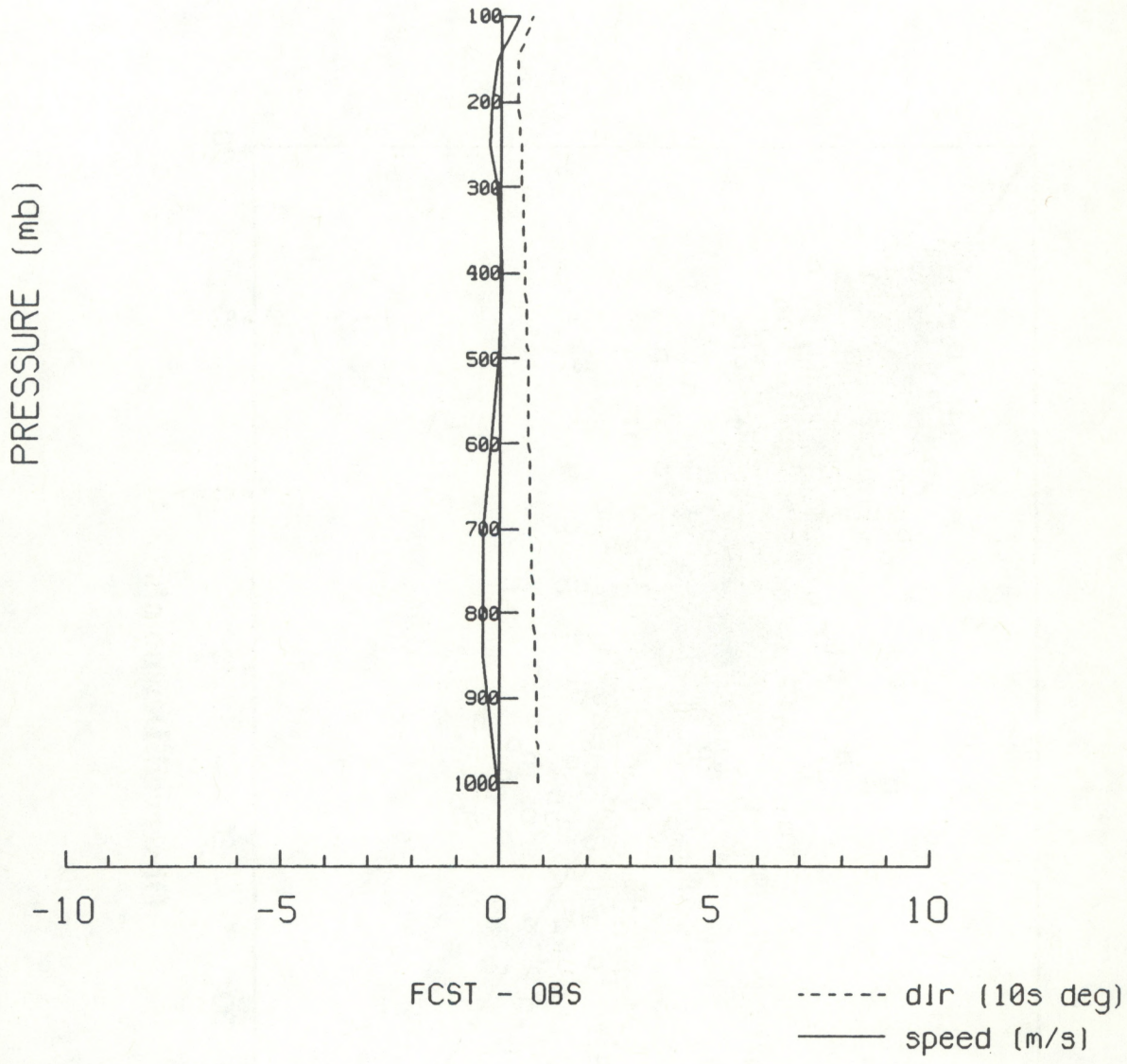

Fig. 9. Mean error (bias) of wind speed and mean absolute error (MAE) of wind direction between MAPS analyses and RAOBS.

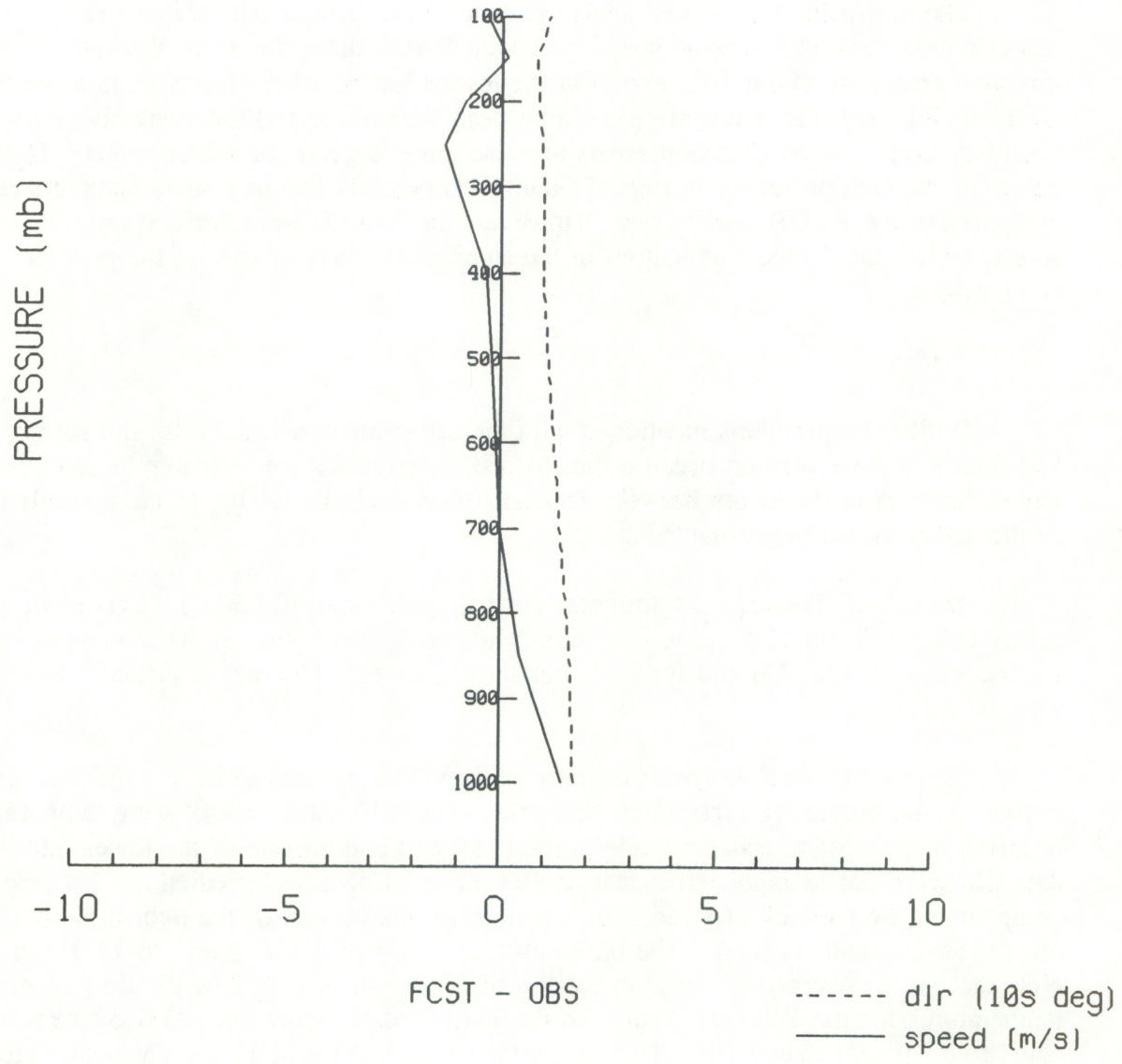

Fig. 10. Same as Fig. 9, except for MAPS 12-h forecast data.

The verification of MAPS analyzed winds with profiler observations (Fig. 11) also shows a near-zero bias of wind speed. As with RAOB data, the mean absolute wind direction errors are about 10°, except at the lowest levels, where the error is more than 20°. By 12 h (Fig. 12), the winds are too strong near the surface and too weak above the boundary layer. Wind direction errors are also quite large in the lower levels. The noisy nature of the data presented in Figs. 11 and 12 is possibly due to a small sample size as compared to the RAOB sample size. However, the RAOBs were verified only at mandatory levels, which could miss fluctuations at the significant levels as seen in the profiler comparisons.

Surface

Unlike the problems mentioned with the upper-air verification, all the surface analyses and forecasts were verified because the surface observations are available hourly. Therefore, any differences in the errors between forecast times could be related to the assimilation cycle *or* the actual model performance.

Altimeter. The analysis altimeter error is quite small (0.03 in.). This error gradually grows to 0.07 in. by 12 h. The largest absolute errors are found in the mountainous states for the analysis (Fig. 13) and forecast (not shown) times. The bias is virtually zero at all forecast times.

Temperature and dewpoint temperature. MAPS appears to have a difficult time analyzing and predicting surface temperatures. The difficulties in analyzing temperature are probably a result of the coarse model terrain (80 km) and the use of the lowest model level data that were not extrapolated to station elevations. For model prediction, this error is compounded by the lack of a radiation equation in this version of the hybrid model (i.e., no diurnal heating and cooling). The initial absolute error of 6.5°F grows to 11.3°F by 12 h. Note that the largest errors occur in the 9-h forecast, which would be the time of maximum temperatures for the 1200 UTC run. In the analysis data, more than 23% of the errors are equal to or greater than 10°F. This percentage rises to 53% at 12 h. The scatterplot of the analyzed temperatures (Fig. 14) shows the data clustered around the diagonal, but by 12 h (Fig. 15) there is more scatter in the data, which would indicate less correlation.

The overall bias is negative (cold) because of the lack of surface heating and the reduction to station elevation from the lowest model level. The scatterplots at 0 and 12 h (Figs. 14 and 15) illustrate the lack of radiation in the model and a pronounced cold bias for observed warm temperatures, i.e., greater than 60°F. Figure 16 clearly shows a cold bias both in the southereastern United States and in the intermountain West.

The absolute dewpoint errors are actually smaller than the temperature errors. Even the scatterplot (Fig. 17) shows better agreement than the ambient temperature (Fig. 14). This is somewhat expected because dewpoint temperature is less affected by the diurnal cycle.

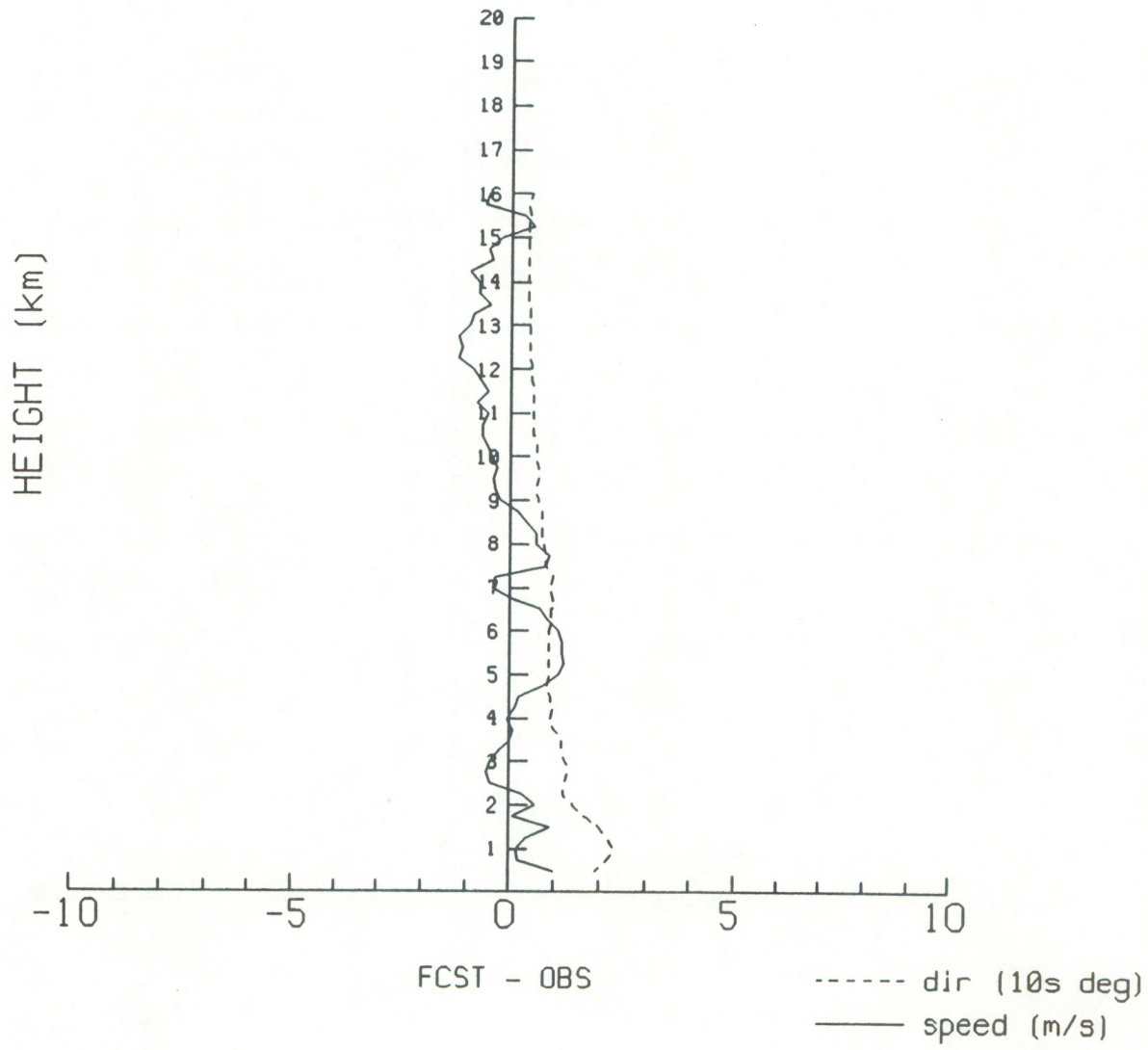

Fig. 11. Mean error (bias) of wind speed and mean absolute error (MAE) of wind direction between MAPS analyses and wind profilers.

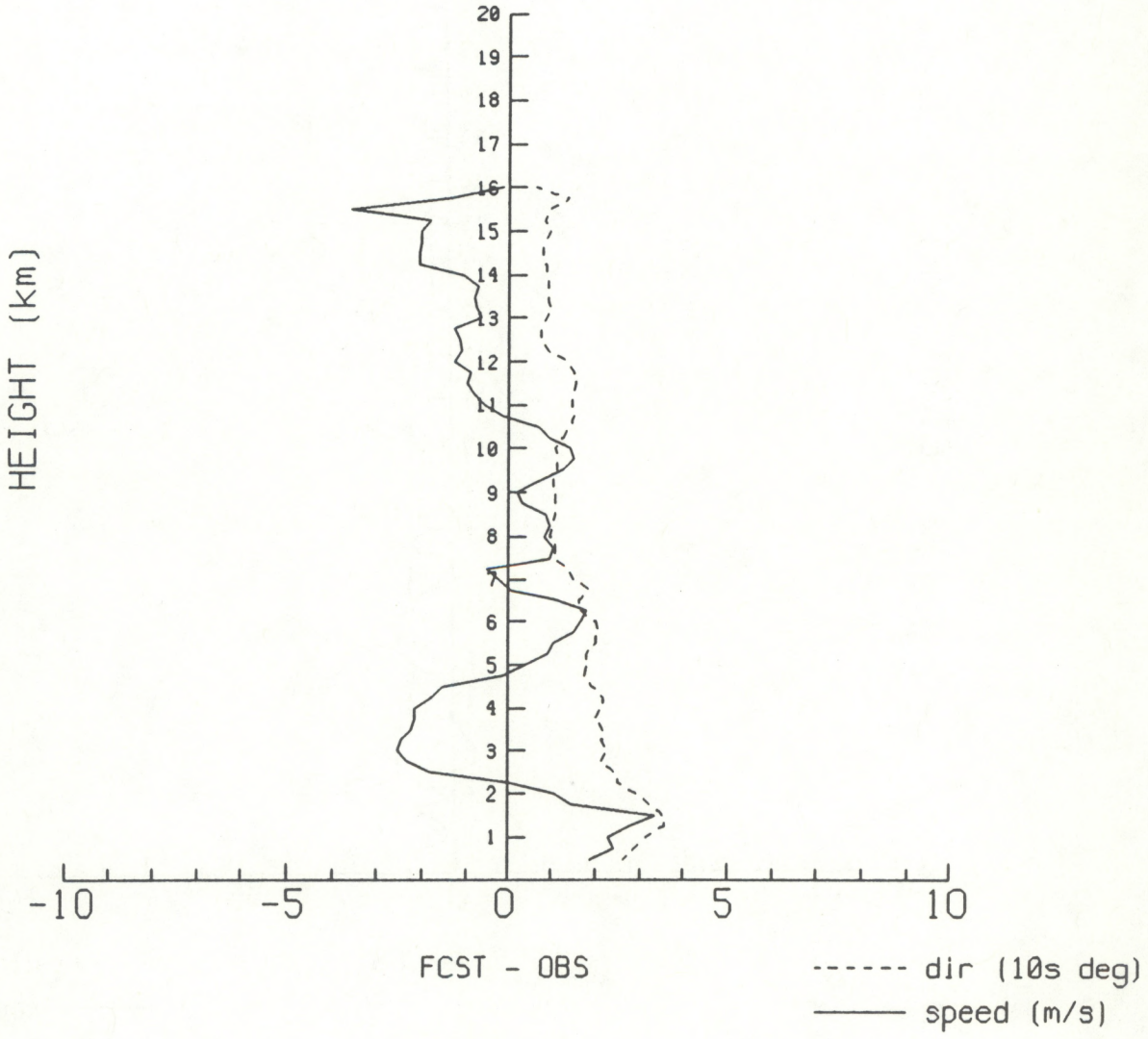

Fig. 12. Same as Fig. 11, except for MAPS 12-h forecast data.

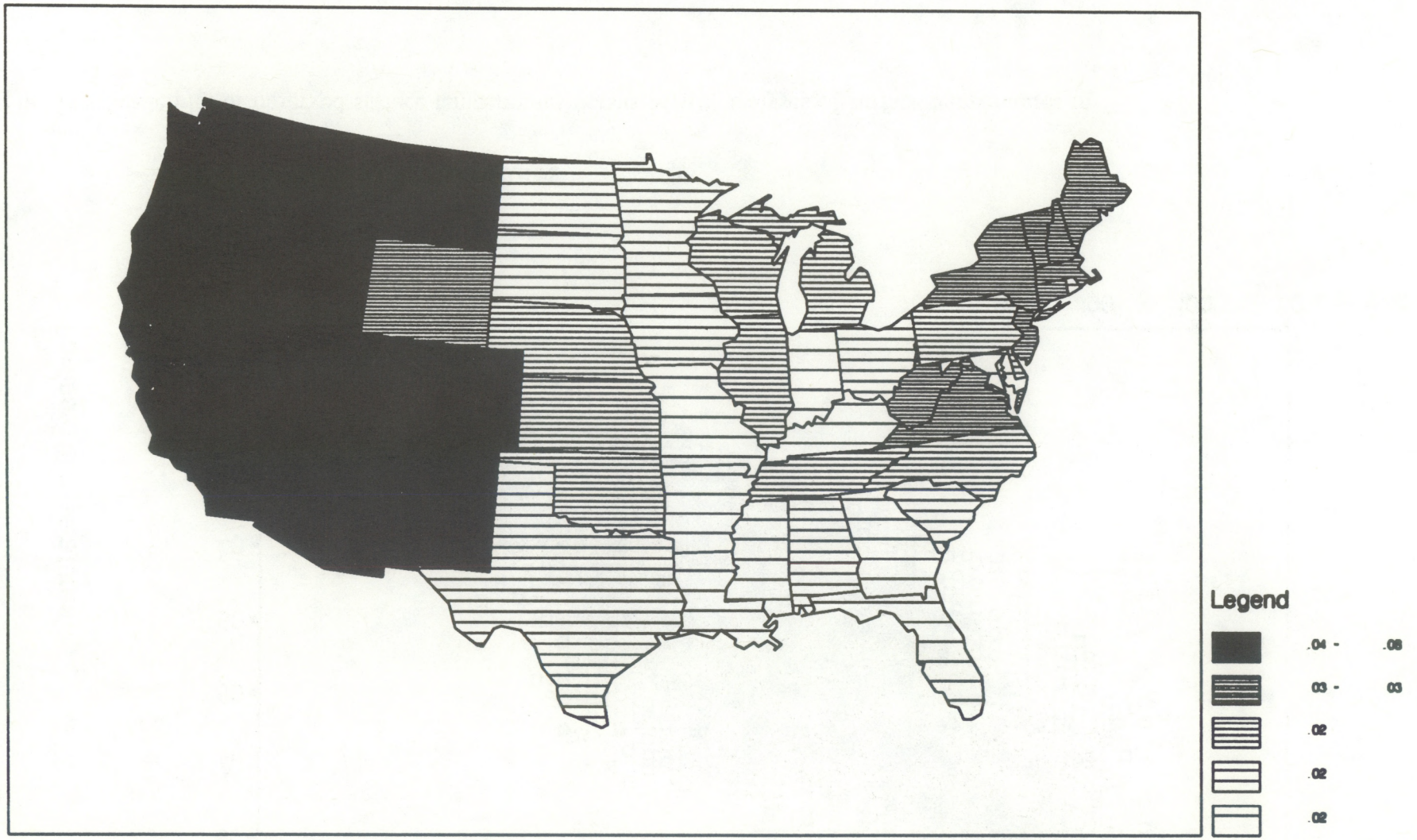

Fig. 13. Map of mean absolute altimeter setting errors (in. Hg) between MAPS analyses and surface observations.

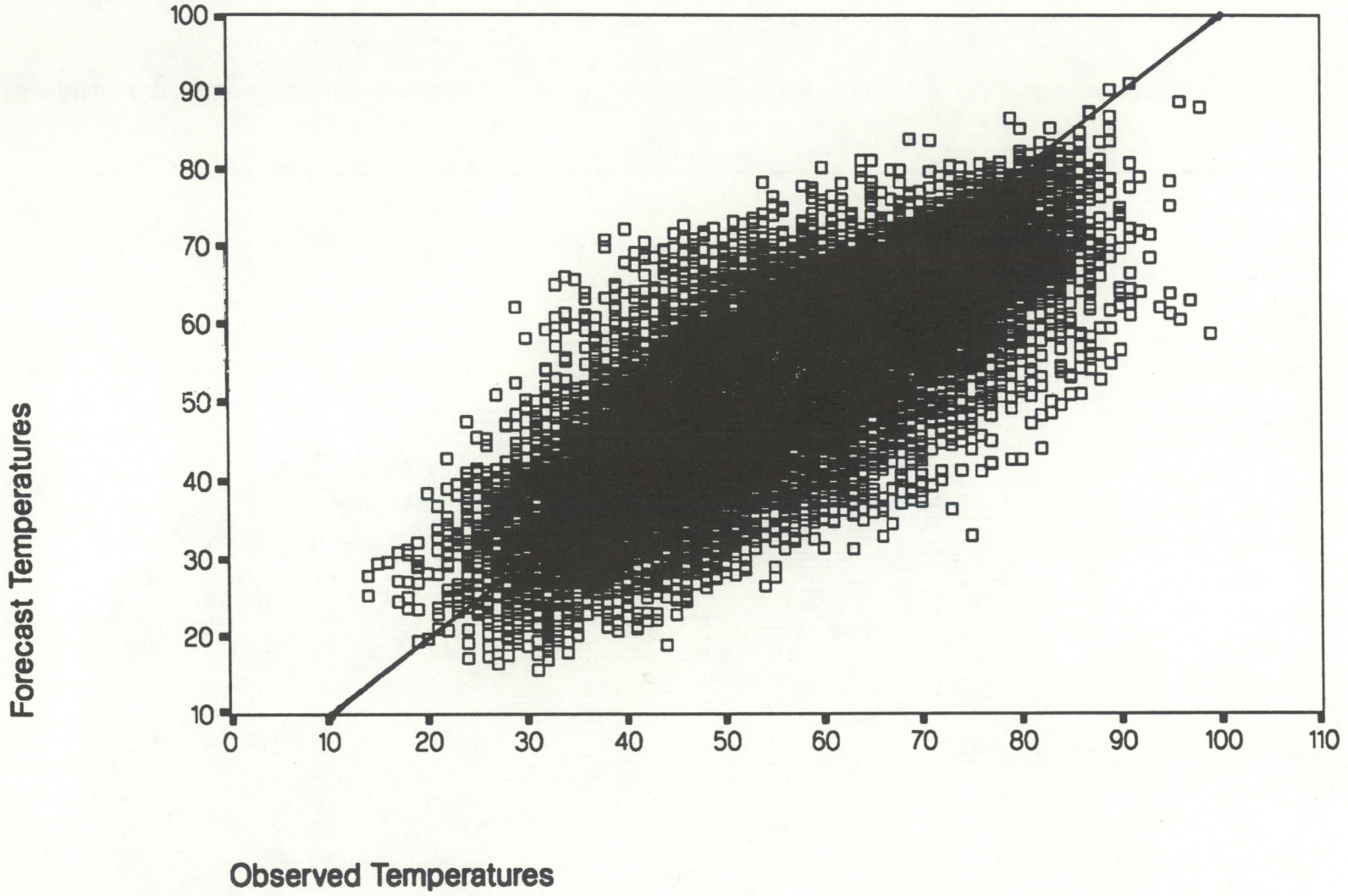

Fig. 14. Scatterplot of observed surface temperatures versus MAPS analyses of surface temperatures in °F.

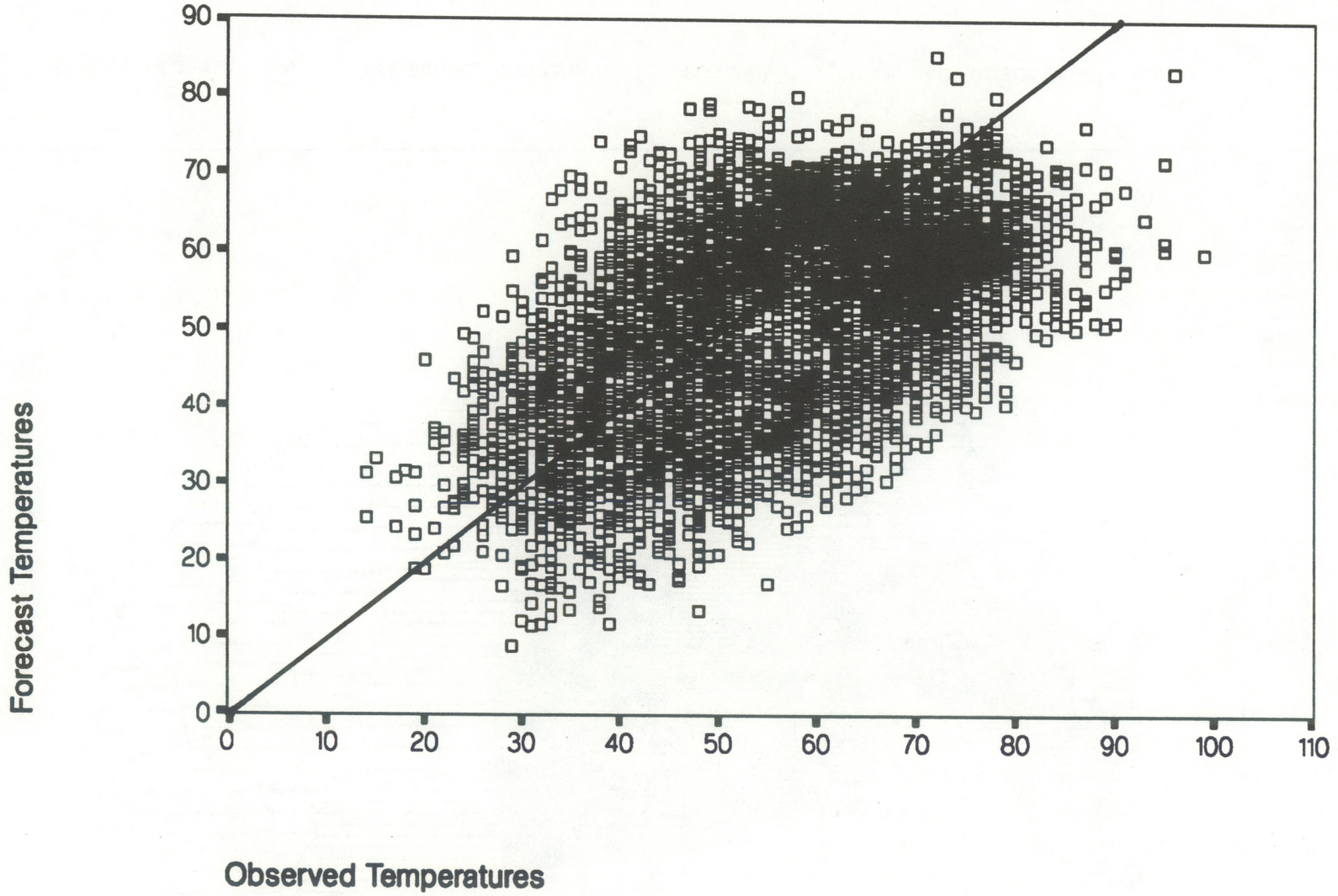

Fig. 15. Same as Fig. 14, except for MAPS 12-h forecast data.

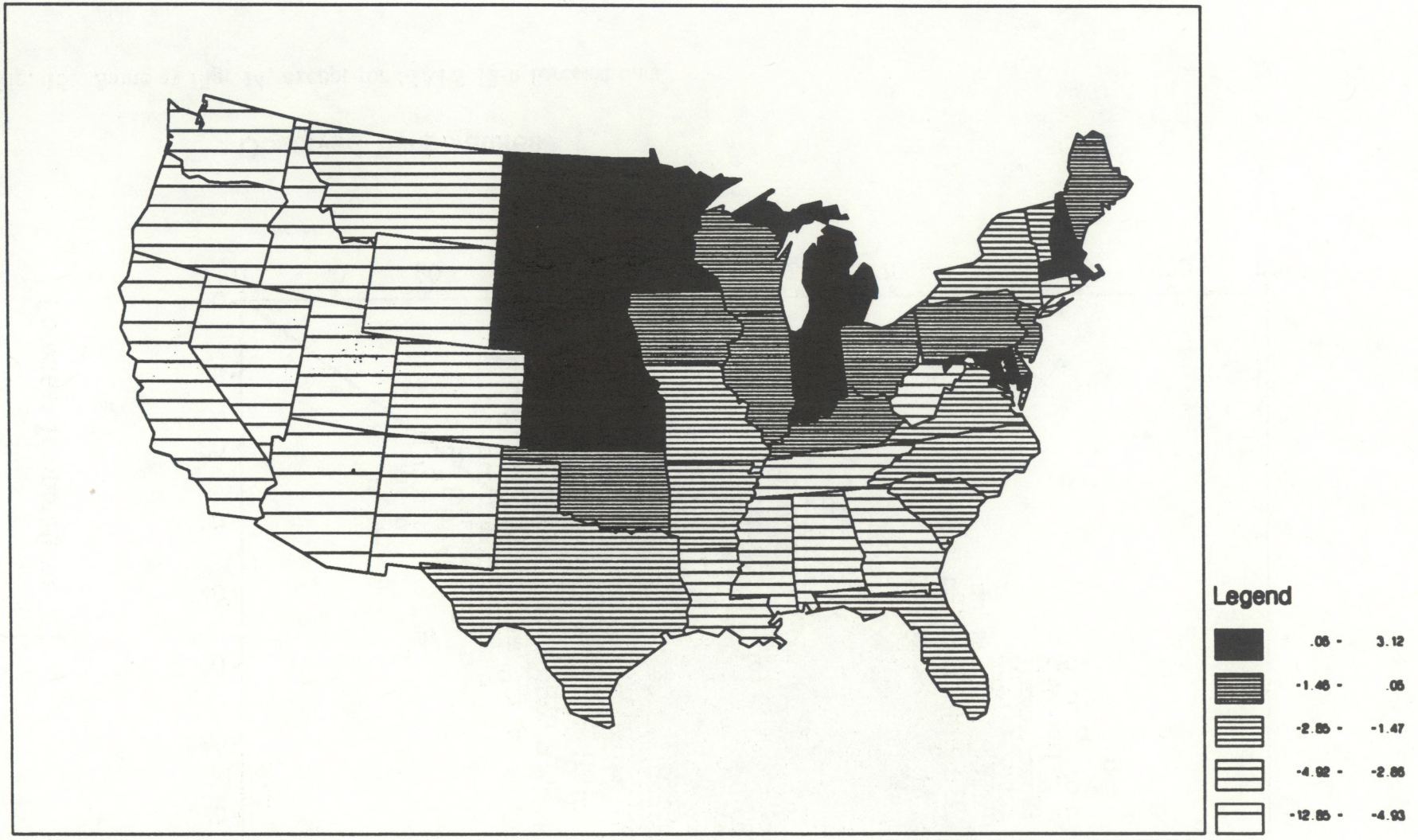

Fig. 16. Map of mean error (bias) of surface temperature between MAPS analyses and surface observations.

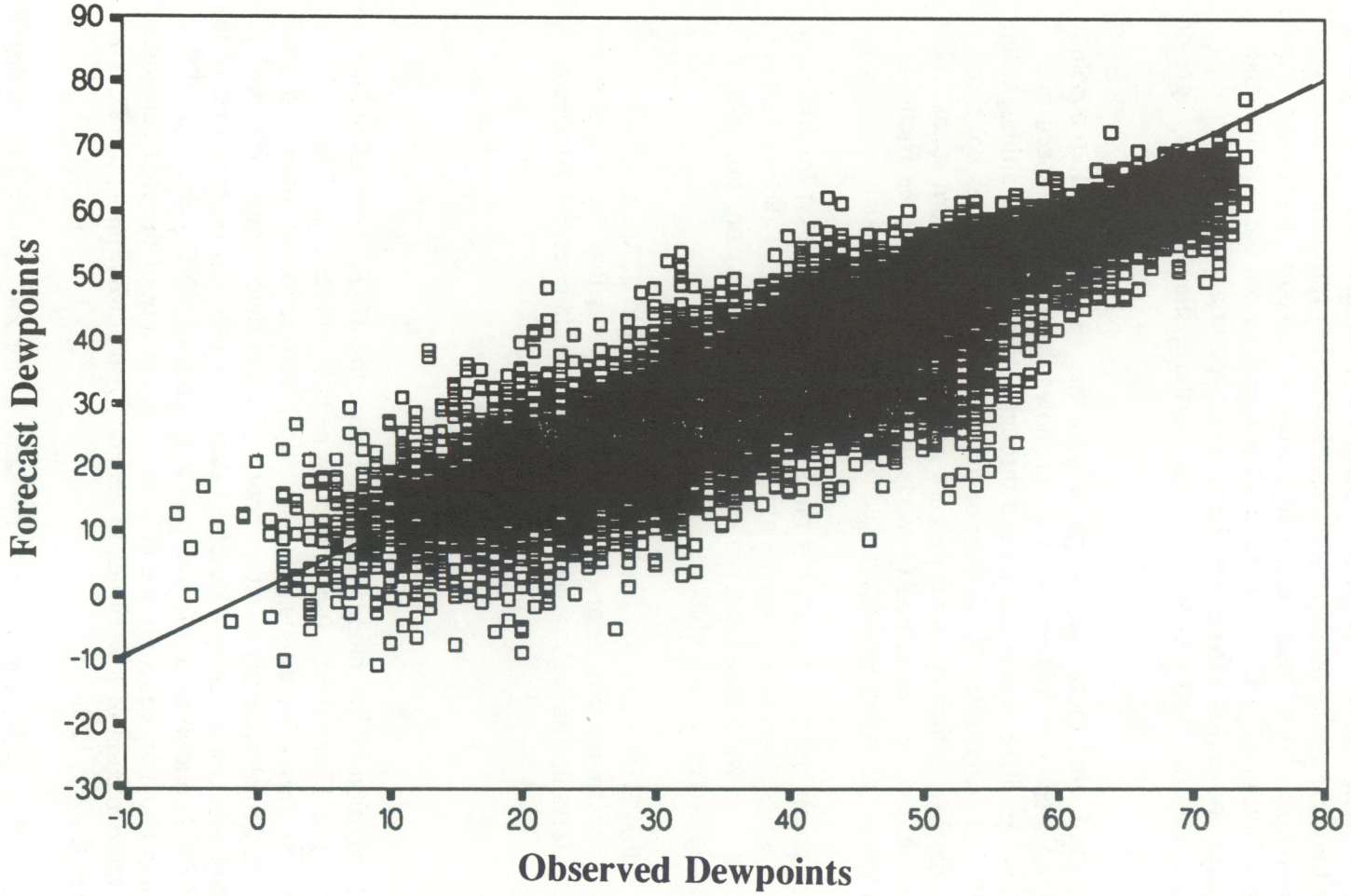

Fig. 17. Scatterplot of observed surface dewpoint temperatures versus MAPS analyses of surface dewpoint temperatures in °F.

As with the upper-air dewpoint errors, the bias of surface dewpoint indicates that the model is somewhat dry in the low levels. Although some of this negative bias could be due to the cold temperature bias, the dewpoint depression bias is also dry (positive), even out to 12 h. Note though, from Appendix C, that the absolute error of dewpoint depression is quite large, especially at 9 and 12 h (12.3°F). Thus, although it appears from the low bias and scatter (Fig. 17) mentioned above that MAPS does better at dewpoint analysis and forecasting, the model actually has difficulty with surface moisture, showing only a slight dry bias.

Winds. The analyzed wind speed has a large positive bias (+7.5 kt) and an absolute error of 8.2 kt. The wind speed distribution (not shown) shows the high wind speed bias well. This pattern of large positive bias is maintained throughout the forecast, although the errors do increase somewhat (Appendix C). A map of the errors (Fig. 18) does not show any definite pattern except that, generally, the errors are greater in the northern states. Note that the MAPS assimilation does not include SAO winds in the analysis, and the frictional affect of the terrain on the winds is not well accounted for in the MAPS model.

Figure 19 clearly points out that MAPS has much larger errors (8-9 kt) for southerly winds. A distribution of wind directions (not shown) indicates that most of the winds have a southerly direction. Thus, this wind speed error is especially critical. Again, this pattern is similar at 12 h, but slightly larger in magnitude.

MAPS wind direction error increases from 27.3° at analysis to 32.0° at 12 h (Appendix C). A map of wind direction error (Fig. 20) clearly shows that the largest errors occur in the mountainous terrain, although large errors are also found over the Southeast.

6.1.2. AIVs

Surface

Clouds. The distribution of low clouds diagnosed from the MAPS analysis appears to be very good (Fig. 21). The 12-h forecast (not shown) is equally impressive. The contingency table (Table 8) shows that although the statistical distribution of diagnosed low cloud amount is good, the actual accuracy is rather poor. (Bold numbers on the diagonal represent perfect matches between model and observed data.) MAPS was correct 79% of the time when it diagnosed CLR, and 49% of the time when it indicated OVC. But if MAPS indicated BKN, there was about an equal chance of it verifying as any of the four categories, including CLR. Even the detection of low clouds in general is somewhat poor; the POD is 0.64 for the analysis and 0.67 at 12 h.

For middle clouds, the results are generally the same as for low clouds, except that the PODs are even lower (0.37) and the FARs are higher (0.32) for the analysis data. At 12 h, the POD is essentially the same, but the FAR drops to 0.20, which is largely a result of fewer forecasts of middle clouds. As with low clouds, the cloud amount discernment of

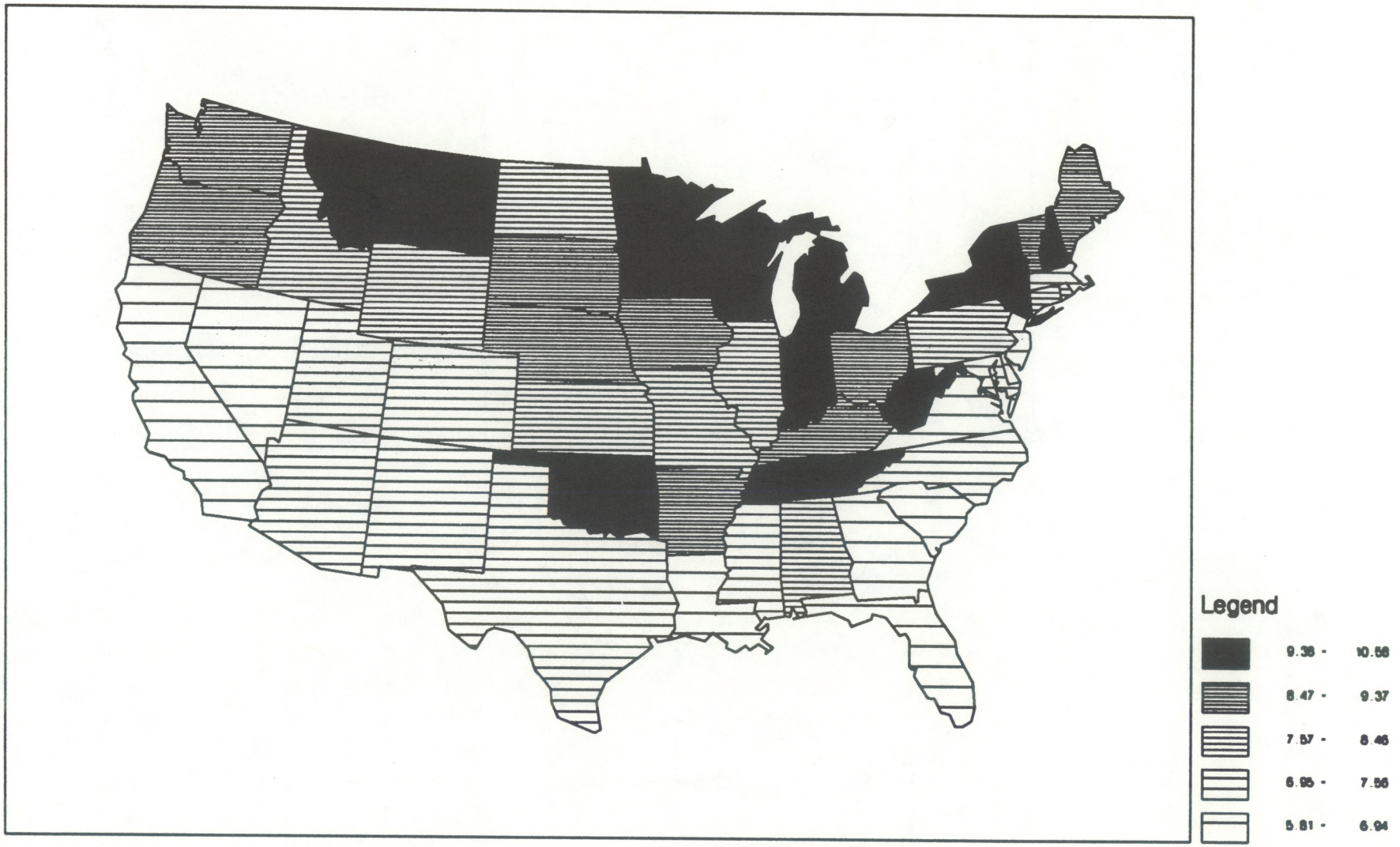

Fig. 18. Map of mean absolute wind speed errors (kt) between MAPS analyses and surface observations.

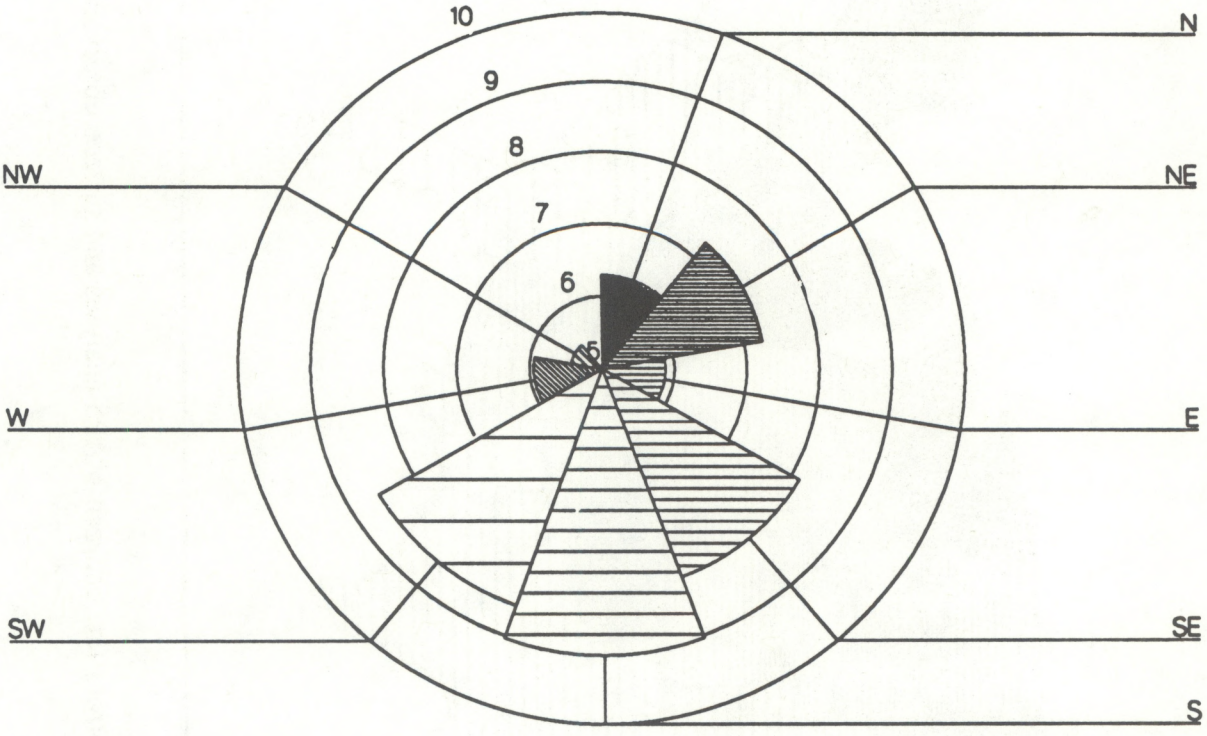

Fig. 19. Distribution of mean absolute wind speed errors (kt) by wind direction between MAPS analyses and surface observations. To avoid light and variable wind directions, only cases with *observed* wind speeds ≥ 10 kt are verified.

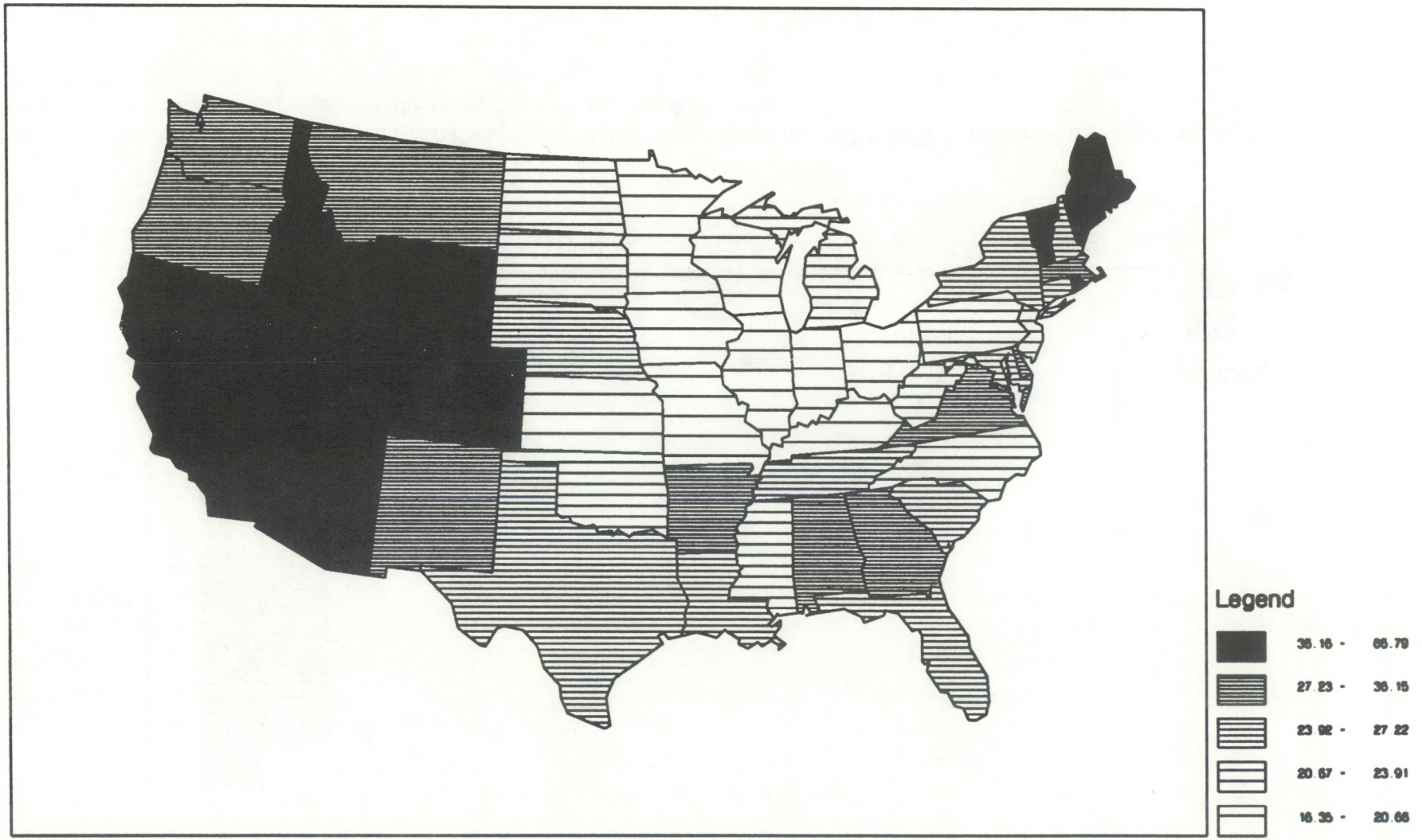

Fig. 20. Map of mean absolute wind direction error (degrees) between MAPS analyses and surface observations. To avoid light and variable wind directions, only cases with *observed* wind speeds $\geq$10 kt are verified.

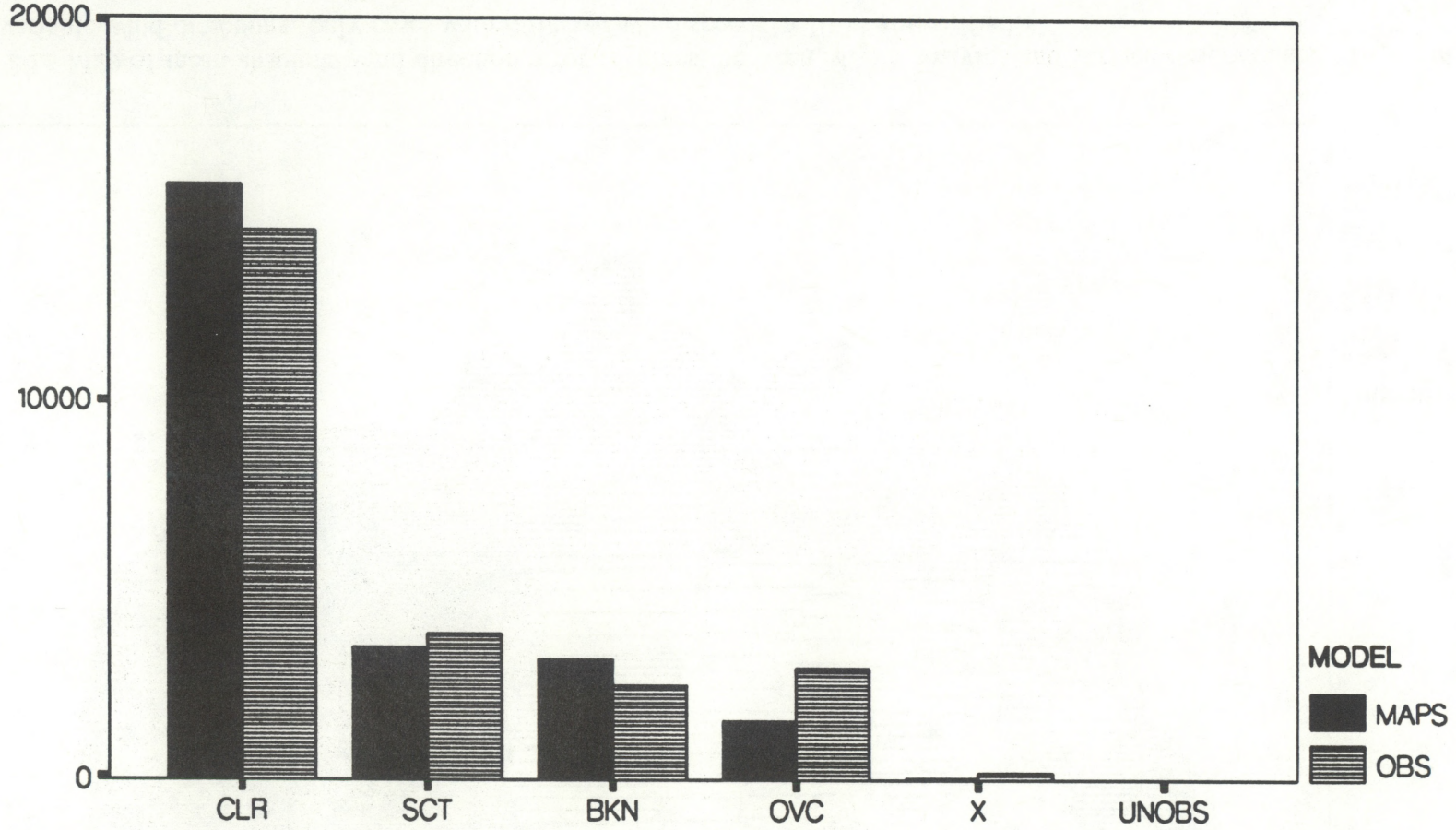

Fig. 21. Distribution of observed and MAPS analyses of low cloud amounts. UNOBSD (unobservable) is defined as those cases when a lower overcast deck prevents the observation of any (higher) clouds.

Table 8. Contingency Table of Diagnosed Low Clouds by Observed Low Clouds for MAPS Analyses, 1-10 April 1991

	Observed Low Clouds					
Diagnosed Low Clouds	CLR	SCT	BKN	OVC	X	Total
CLR	**12241**	1882	830	550	55	15558
SCT	1181	**864**	635	676	37	3393
BKN	725	730	**690**	891	51	3087
OVC	212	264	259	**733**	24	1492
X	7	10	5	9	**5**	36
Total	14366	3750	2419	2859	172	23566

middle clouds is poor. The dry bias aloft in the MAPS model (see upper air temperature and dewpoint temperature discussion in section 6.1.1. and Fig. 3) translates into a lack of cirrus clouds.

Cloud tops. As is apparent from the statistics in Appendix C, MAPS is very poor at diagnosing cloud-top height, and average errors are about 17,000 ft. Note from the bias that the MAPS heights are almost exclusively too low. This is not surprising because of the lack of cirrus clouds diagnosed from the MAPS model. As seen from the scatterplot in Fig. 22, many of the MAPS cloud tops are below 10,000 ft.

Ceiling. Similar to that of the low clouds, the distribution of ceiling diagnosed from MAPS analyses and forecasts appears very good (Figs. 23 and 24). But as seen from the data in Appendix C, the MAE at 12 h is worse than the analysis, although the bias has improved. The HSS drops from 0.23 for the analysis to 0.05 for the 12-h forecast. Note that the bias is largely positive, implying that the diagnosed ceilings are too high.

The analysis of the upper-air dewpoint errors noted that the 12-h forecast was better than the 3-h forecast at 850 mb. The ceiling errors in Appendix C somewhat agree with this. Note that all error measures indicate better performance at 9 h than at 3 h. Since all 3-h surface forecasts can be verified for the surface data, it is difficult to draw the same conclusions as those of the upper air. Still, the errors show that the usual assumption that the shorter forecast is better than the longer one may not always be correct. This is because of the assimilation cycle and the observation data available to initialize each run.

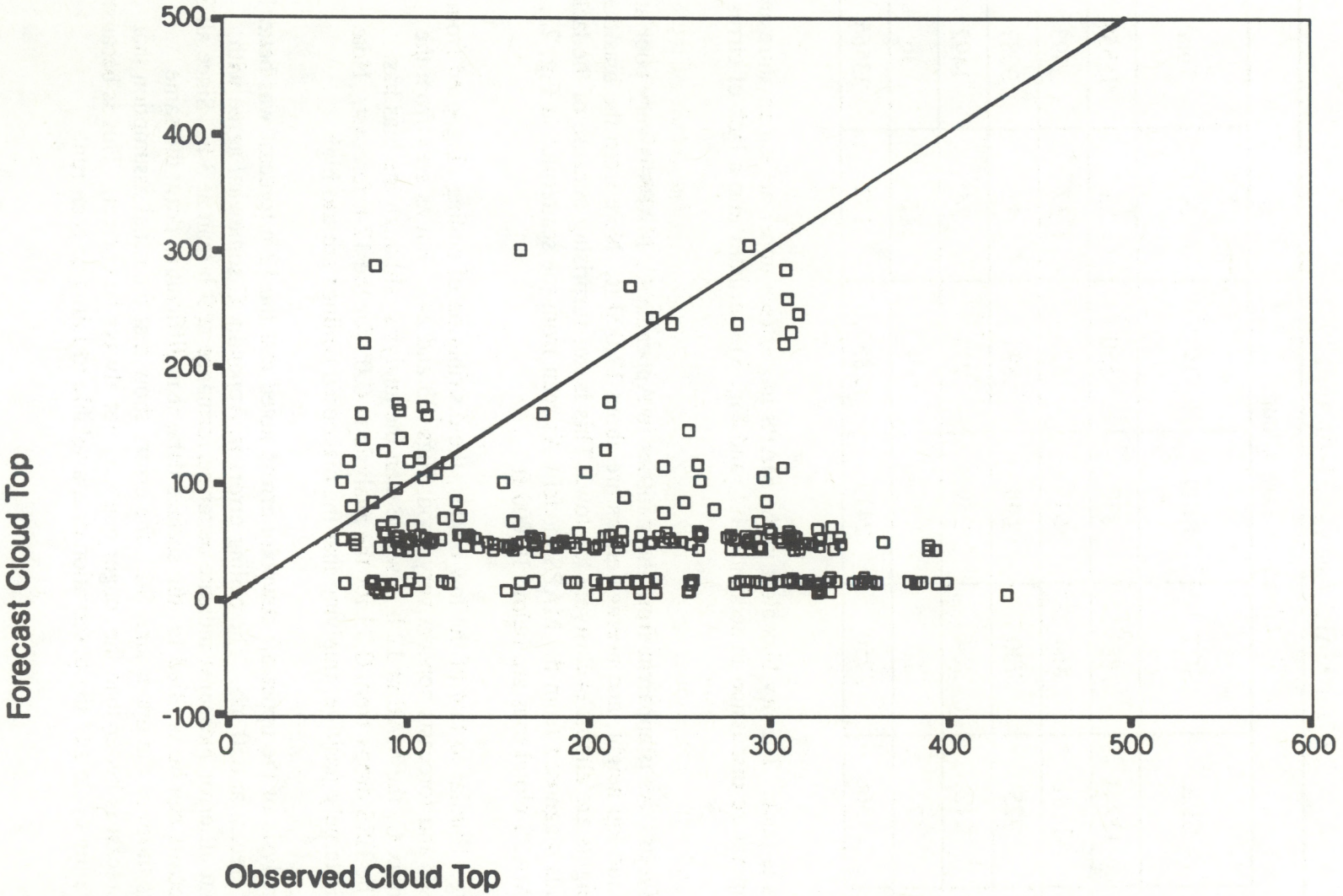

Fig. 22. Scatterplot of observed cloud-top heights versus MAPS analyses of cloud-top heights in hundreds of ft.

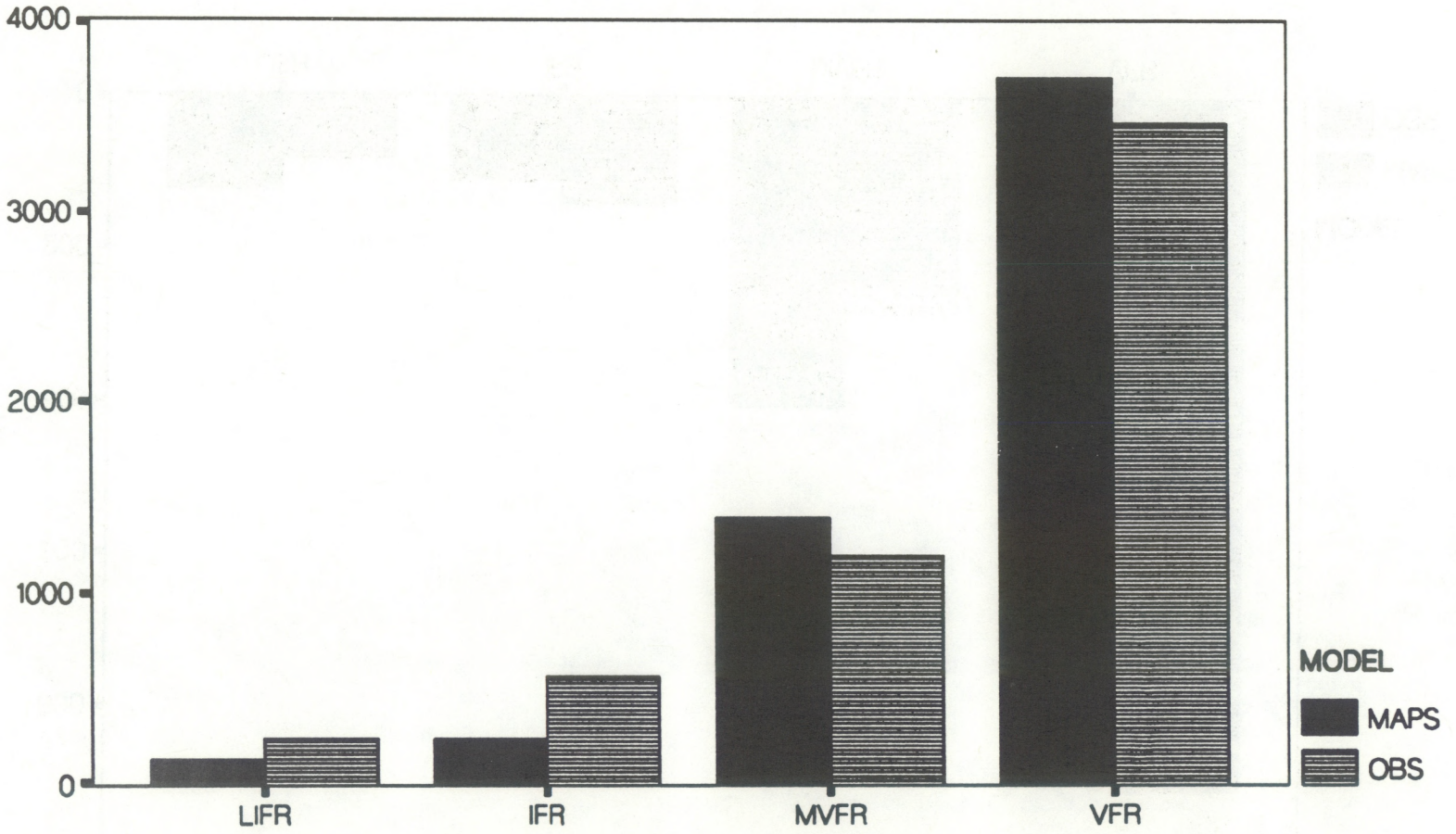

Fig. 23. Distribution of observed and MAPS analyses of ceiling and visibility by category.

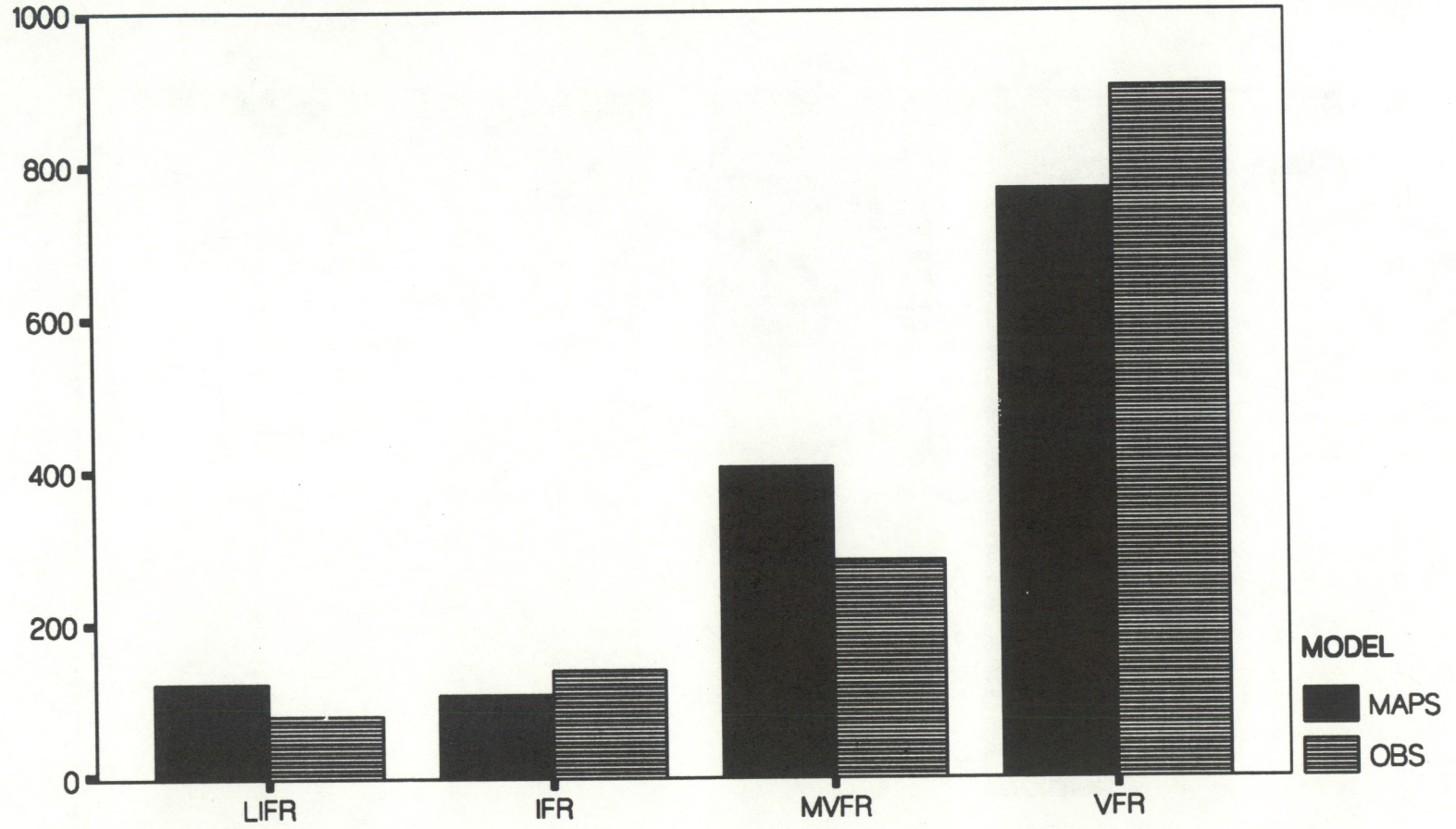

Fig. 24. Same as Fig. 23, except for MAPS 12-h forecast data.

Examination of the contingency table (Table 9) shows that, except for visual flight rules (VFR), the MAPS analysis was more often wrong than right. This same finding holds true for the forecast data (not shown). Probably most disturbing is the large number of busts in the MAPS data (i.e., four lower-left and four upper-right cells). For example, when MAPS analyzed low instrument flight rules (LIFR) or instrument flight rules (IFR) conditions, they were actually verified as marginal visual flight rules (MVFR) or VFR 63% of the time. These numbers are even worse for the 12-h forecast data (85% of LIFR/IFR forecasts verified as MVFR/VFR). This results in very low PODs and high FARs.

Table 9. Contingency Table of Diagnosed Ceilings by Observed Ceilings for MAPS Analyses, 1-10 April 1991

	Observed Ceiling				
Diagnosed Ceiling	LIFR	IFR	MVFR	VFR	Total
LIFR	**7**	17	16	27	67
IFR	13	**52**	70	27	162
MVFR	86	212	**438**	648	1384
VFR	99	194	574	**2990**	3857
Total	205	475	1098	3692	5470

Visibility. From the errors in Appendix C, it is obvious that MAPS has difficulty diagnosing visibility from analyzed and forecast SAVs. Note that the MAE and ME are nearly equal, and the ME is positive at all times. Thus it appears that MAPS over forecasts visibility in almost all cases.

However, as seen in the contingency table (Table 10), there are times when MAPS detects conditions as LIFR or IFR and they verify as VFR. If the dominant VFR category is not considered in calculations, the resulting PODs are less than 0.10, whereas the FARs are about 0.80-0.90 (0.98 at 12 h). As with ceiling, MAPS in general diagnoses conditions lower than VFR less often than they occur.

Figure 25 shows that MAPS also has difficulty correctly specifying the type of visibility obstruction. The MAPS distribution changes somewhat by 12 h (not shown), there is more blowing snow and less fog and haze. Presumably these changes are related to the increasing wind speed error in the forecasts.

The distribution of haze in Fig. 25 looks quite good, but less than 6% of the haze events in MAPS actually verified. Blowing snow was observed only 6 times, and yet it was diagnosed from MAPS analyses 109 times. For the entire 10-day data set, blowing snow

Table 10. Contingency Table of Diagnosed Visibilities by Observed Visibilities for MAPS Analyses, 1-10 April 1991

	Observed Visibility				
Diagnosed Visibility	LIFR	IFR	MVFR	VFR	Total
LIFR	**9**	12	6	51	78
IFR	14	**20**	41	116	191
MVFR	42	59	**88**	263	452
VFR	184	362	1043	**21135**	22724
Total	249	453	1178	21565	23445

was observed only 37 times. The average wind speed in these cases was about 15 kt, and the maximum wind speed was 21 kt. All the events had snow falling at the time, but it was usually S- and occasionally S, but never S+. This would imply that blowing snow is very difficult to predict.

The MAPS analyses and forecasts showed an average wind speed of about 25 kt for blowing snow events and range of 13 to 45 kt. Thus, the positive bias of blowing snow events could be related to the strong positive bias of wind speed, since frozen precipitation in MAPS had a negative bias.

Obviously, MAPS needs improvement on visibility forecasts. It is unknown whether the problems lie in the AIV algorithm or in the model itself. The AIV algorithm is a regression equation that uses precipitation rate, relative humidity, and wind speed in its computation. As shown in other sections, MAPS in general is slightly dry, too windy, and lacks precipitation. Thus, it is possible that the algorithm is good but lacks correct input data.

Precipitation. MAPS detected only 29% of the precipitation events in the 12-h forecast. The FAR was quite high at 59%. Therefore, when MAPS forecasts precipitation, more often than not it did not occur.

MAPS does a fairly good job of discerning the correct phase. The contingency table (Table 11) of the analysis data shows that when MAPS diagnosed liquid precipitation, it was rarely freezing or frozen. When MAPS said the precipitation was frozen, there was an equal chance of it being liquid. The other two categories have such a small sample size that it is difficult to draw any conclusions. Although only the analysis results are shown, the same findings hold true for all forecast times.

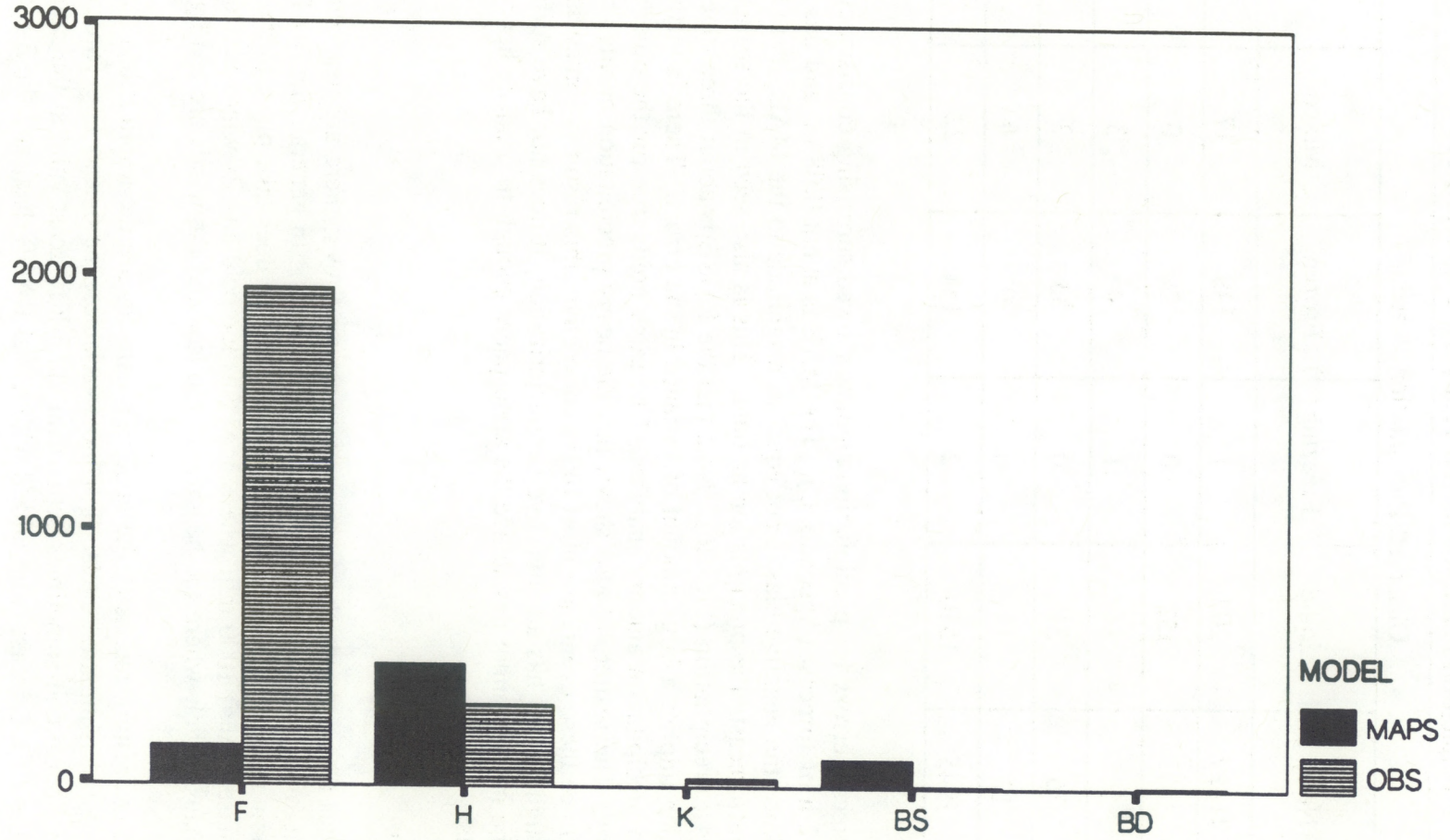

Fig. 25. Distribution of observed obstructions to visibility versus MAPS analyses of obstructions to visibility. F is defined as fog, H is for haze, K is smoke, BS is blowing snow, and BD is blowing dust.

Table 11. Contingency Table of Diagnosed Precipitation Phase by Observed Precipitation Phase for MAPS Analyses, 1-10 April 1991

	Observed Precipitation Phase					
Diagnosed Precip. Phase	None	Liquid	Freezing	Frozen	Mixed	Total
None	**21497**	1102	1	149	17	22766
Liquid	359	**352**	0	2	0	713
Freezing	1	14	**1**	4	3	23
Frozen	76	31	0	**33**	2	142
Mixed	12	11	0	0	**0**	23
Total	21945	1510	2	188	22	23667

MAPS forecast errors for precipitation amount showed surprising consistency throughout the 12 h of forecast (Appendix C). The MAE is about 0.08 in. and has a negative bias. Since the negative bias is not equal in magnitude to the MAE, there are cases where MAPS overforecasts the precipitation amount. This is also seen in the scatterplot of MAPS 3-h forecast precipitation (Fig. 26). Note that the RMSE is about three times as large as the MAE, which implies a large amount of variance in the errors. There is virtually no correlation in the precipitation amount analyses. The same holds true for the other forecast times (not shown). The scatterplot also shows that the heavy precipitation events are underforecast. Since these events are most likely convective, this result is somewhat expected. Although MAPS has a convective parameterization, it does not have surface heating. Thus the model cannot destabilize the atmosphere enough to produce heavy convective events.

6.1.3. Discussion

MAPS does a very good job of analyzing upper-air SAVs; there are near-zero biases at all levels. The only exception is a dry bias at upper levels. In general, the MAPS forecasts also show no bias at most levels. A definite drying does take place in the lower levels. Other biases are at 1000 mb, where MAPS is too cold and its winds are too strong, and at the tropopause level, where the heights are too low and the winds are too light.

The most interesting results of the upper-air statistics are the implications that were drawn about the data assimilation cycle. For most variables, the 6- and 12-h forecasts were found to be nearly equal and sometimes better than the 3-h forecast valid at the same time. This is because the 12-h forecast is initialized with actual RAOB data, and the 3-h forecast

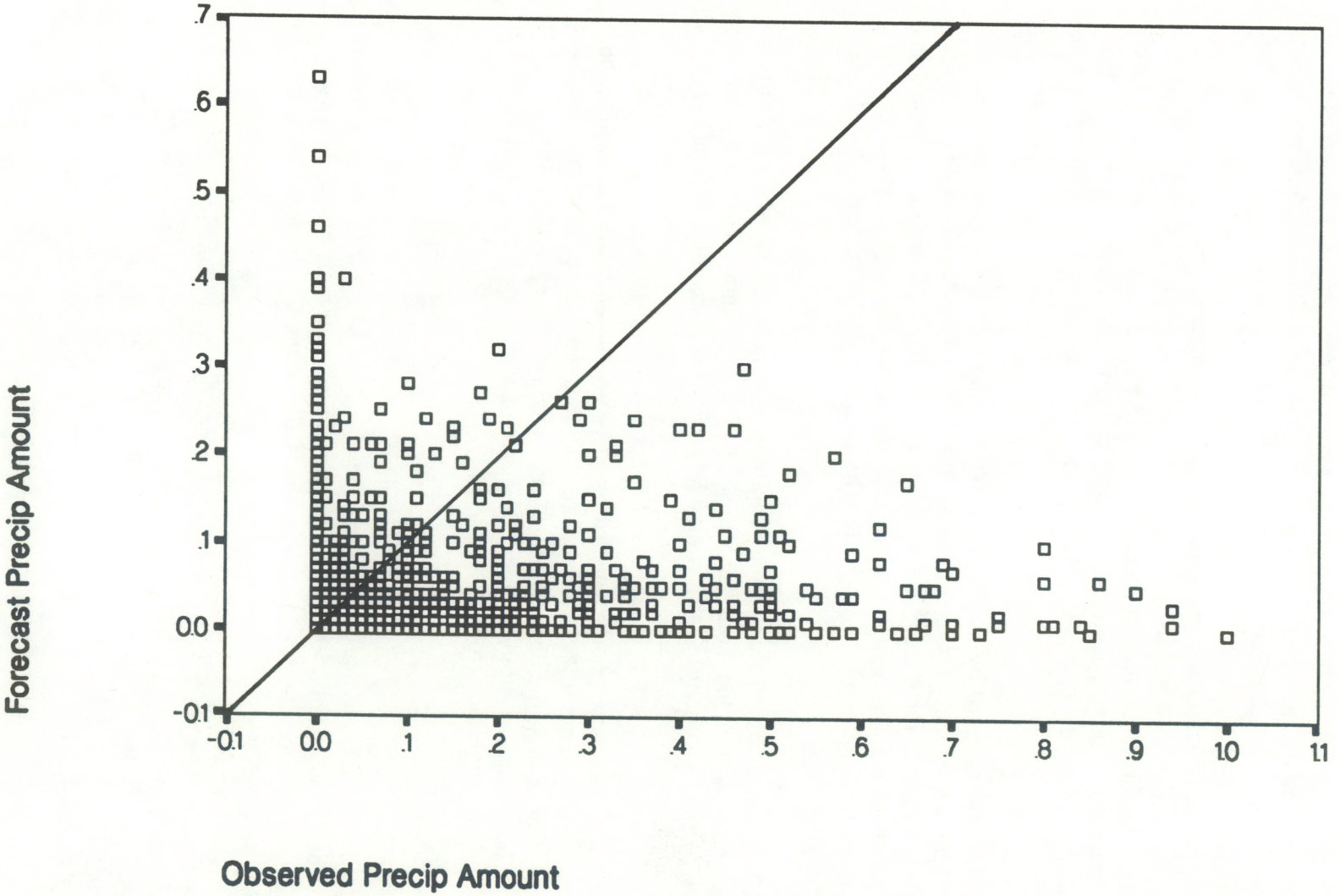

Fig. 26. Scatterplot of observed 3-h precipitation amounts versus MAPS 3-h forecast accumulations of precipitation amounts in in.

valid at the same time is initialized with three consecutive 3-h forecasts. This is most notable for moisture, since the RAOBs are the only source of moisture data, whereas winds and temperatures are available hourly from ACARS and profilers. These results contradict the usual assumption that a 3-h forecast is *always* better than a 12-h forecast.

MAPS has more difficulty analyzing and forecasting surface parameters. MAPS temperature analyses and 12-h forecasts are in error by 10°F or more 23% and 54% of the time, respectively. As shown by the upper-air results, MAPS tends to be too cold and too dry. The wind speed has a large positive bias in both the analysis and the forecast. The largest errors occur with southerly winds, which is the dominant wind direction in this data sample.

Low clouds were detected by MAPS about 65% of the time they were observed. Skill at discerning cloud amount is poor, especially for SCT and BKN categories. Results are similar, but not as good, for middle and high clouds. MAPS lack of cirrus resulted in large errors in cloud-top diagnoses from the analyses and forecasts. It is encouraging that the distribution of ceilings appeared to be good. The actual errors were near 3000 ft, and they were generally too high.

MAPS had much difficulty diagnosing visibility. It is hard to say whether that is a result of the algorithm used or the model SAVs used by the algorithm. The overforecasting of blowing snow led to the conclusion that it may be difficult to forecast based on the observations in this study.

Precipitation events were poorly detected by MAPS. The errors in forecasting the amount were consistently about 0.08 in. at all forecast times, and there was a bias toward underforecasting. Concerning the discernment of precipitation phase, if MAPS indicated the precipitation was liquid, it was rarely wrong. However, if it indicated frozen, there was an equal chance of it being liquid or frozen.

6.2. Eta

Appendix D lists the summary of Eta statistics for the SAVs and AIVs at the surface and the rawinsonde levels of 850, 500, and 250 mb at 3-h time periods out to the 24-h forecast. The following subsections discuss further details of the Eta evaluation grouped by SAVs or AIVs for the surface and upper air.

The Eta system does not have its own analysis of raw observation data. Its initial fields are provided by the NGM analysis and interpolated to the Eta higher resolution grid while taking into account the differences in the terrain of the two models. Thus, the initial hour statistics for Eta are largely a verification of the NGM analysis. The largest differences between the Eta and NGM will be near the surface, where the adjustments for differences in the model terrain were large.

6.2.1. SAVs

Upper air

Height. The initial hour height errors show only a slight negative bias in the upper troposphere (Fig. 27). At 12 h (Fig. 28), this bias is reversed above 250 mb, and there is a +25 m bias at 100 mb. Below 500 mb, there is no apparent bias. Absolute initial hour errors (Appendix D) range from 5.3 m at 850 mb to 14.3 m at 250 mb. For the 12-h forecast, the absolute errors double at 850 mb and are 50% larger at 250 mb.

Temperature and dewpoint temperature. The vertical plots of arithmetic temperature errors (Figs. 27 and 28) are impressive; there is nearly no bias at any level. The only notable exceptions are the slight positive biases at 1000 mb for the 12-h forecast (Fig. 28), and above 250 mb. Also, the upper troposphere bias changes from near zero at the initial time to positive bias at 12 h, which is in agreement with the height bias. Since all Eta runs began at 0000 UTC, the 0- to 12-h forecast would be during nighttime hours. This increasing positive bias may be a result of a lack of cooling aloft. The initial hour absolute errors (Appendix D) are all less than 1°C, and increase to only 1.4°C by 12 h.

For dewpoint temperature, the model is too moist above 600 mb in the initial hour data. Remember that Eta is initialized with the NGM, and therefore this bias can be attributed to the NGM. This bias remains at 12 h but it is larger and down to 700 mb. Below 700 mb, the forecast bias is to dry out the lower levels. Thus it appears that low-level moisture is erroneously mixed to higher levels during the model integration. Absolute errors range from 1.4°C at 850 mb at the initial time to 7.2°C for 500 mb 12-h forecast.

Winds. The u,v initial hour data look very good; there are no biases at any level. This results in good overall speed errors, with a slight negative bias at the initial hour (Fig. 29). For the 12-h forecast (Fig. 30), the only errors of note are the +2.0 m s^{-1} biases at 1000 mb and above 200 mb. The wind directions are also very good; absolute errors (Appendix D) are less than 10° at the initial time, and less than 20° at 12 h. A possible source of the RMSE could be attributed to the initialization procedure of the NGM, which uses nonlinear normal model analysis to balance the mass and wind fields. Thus, one would not expect the initial wind fields to exactly match the observational winds.

Surface

Altimeter. The MAE (Appendix D) at the initial time was only 0.03 in., and there were no significant busts (i.e., greater than 0.20 in.). A map of the errors (Fig. 31) shows that all the largest errors (greater than 0.03 in.) were located in mountainous states. In most cases, the bias in these states was strongly negative. Positive biases were found from North Dakota to Washington and California.

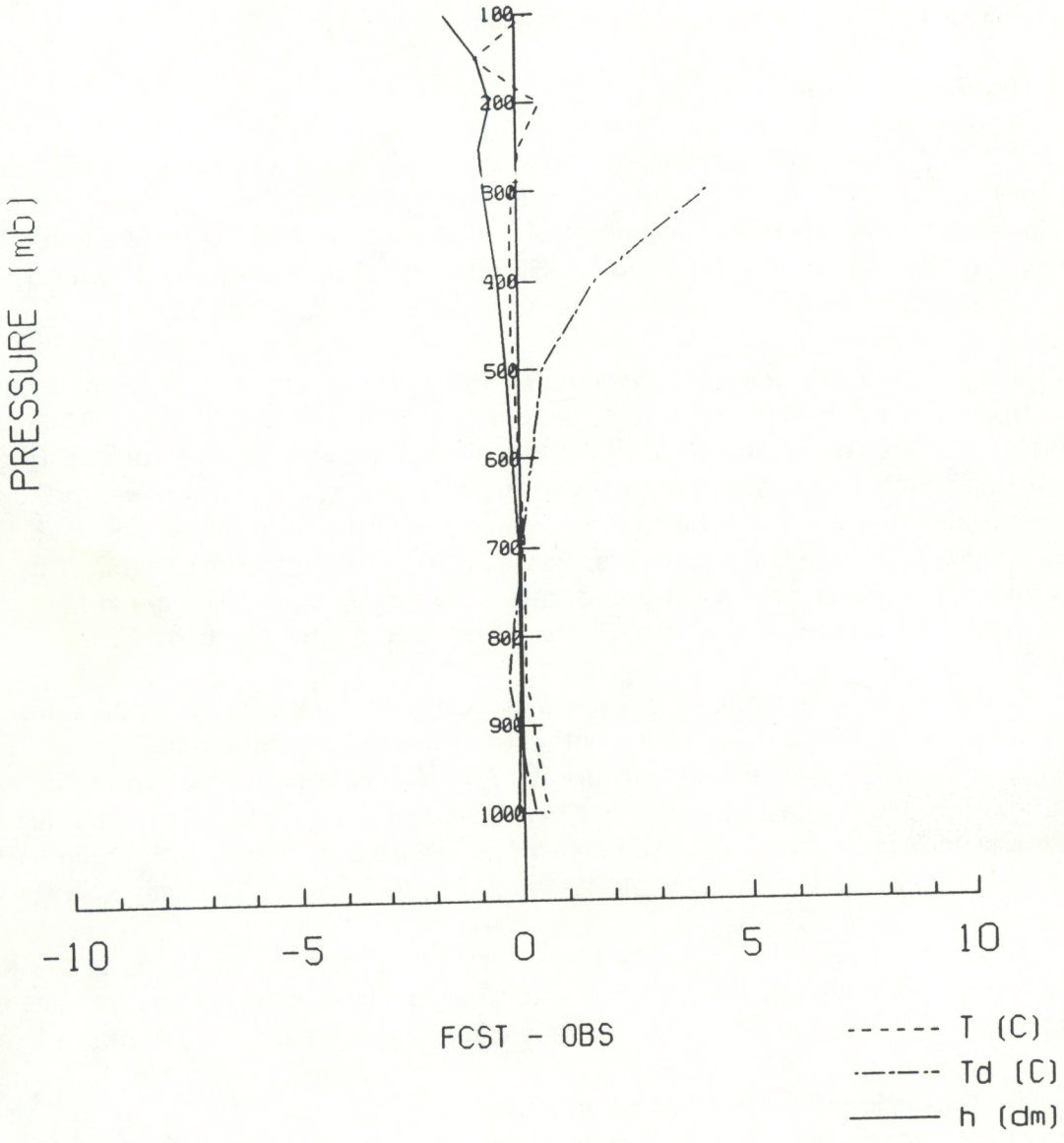

Fig. 27. Mean error (bias) of height (h), temperature (T), and dewpoint temperature (T_d) between Eta initial hour data and RAOBs.

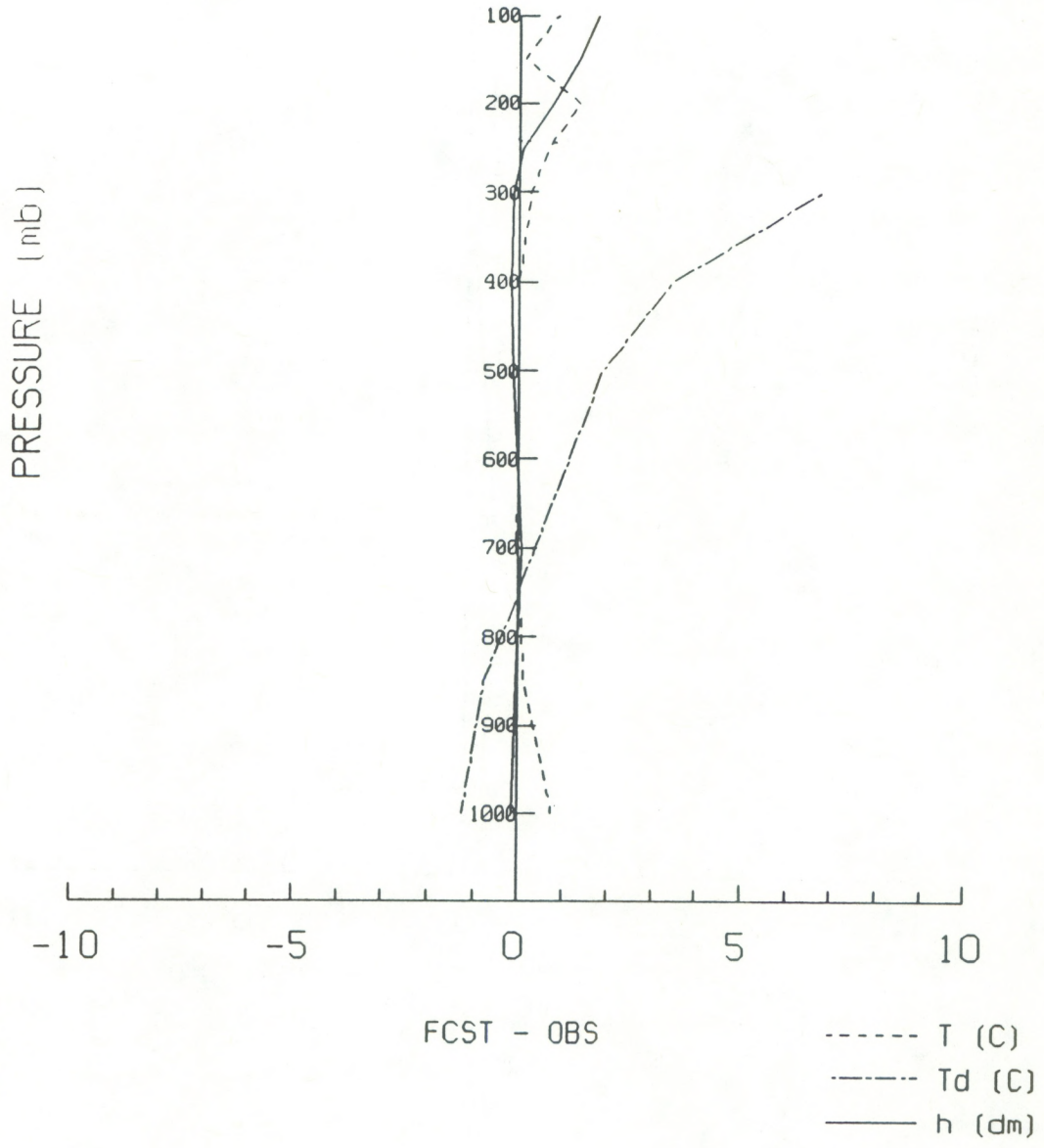

Fig. 28. Same as Fig. 27, except for Eta 12-h forecast data.

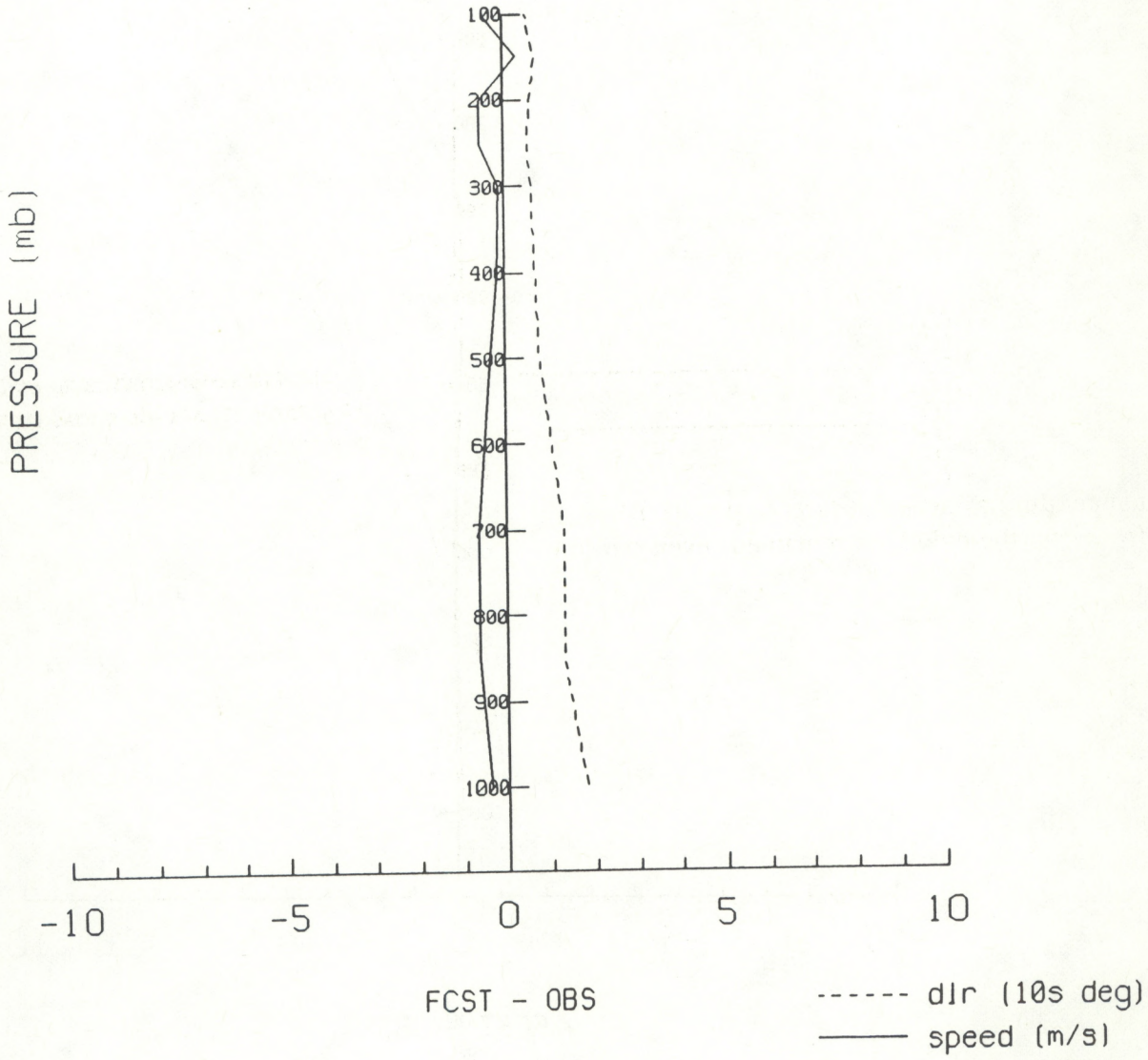

Fig. 29. Mean error (bias) of wind speed and mean absolute error (MAE) of wind direction between Eta initial hour data and RAOBs.

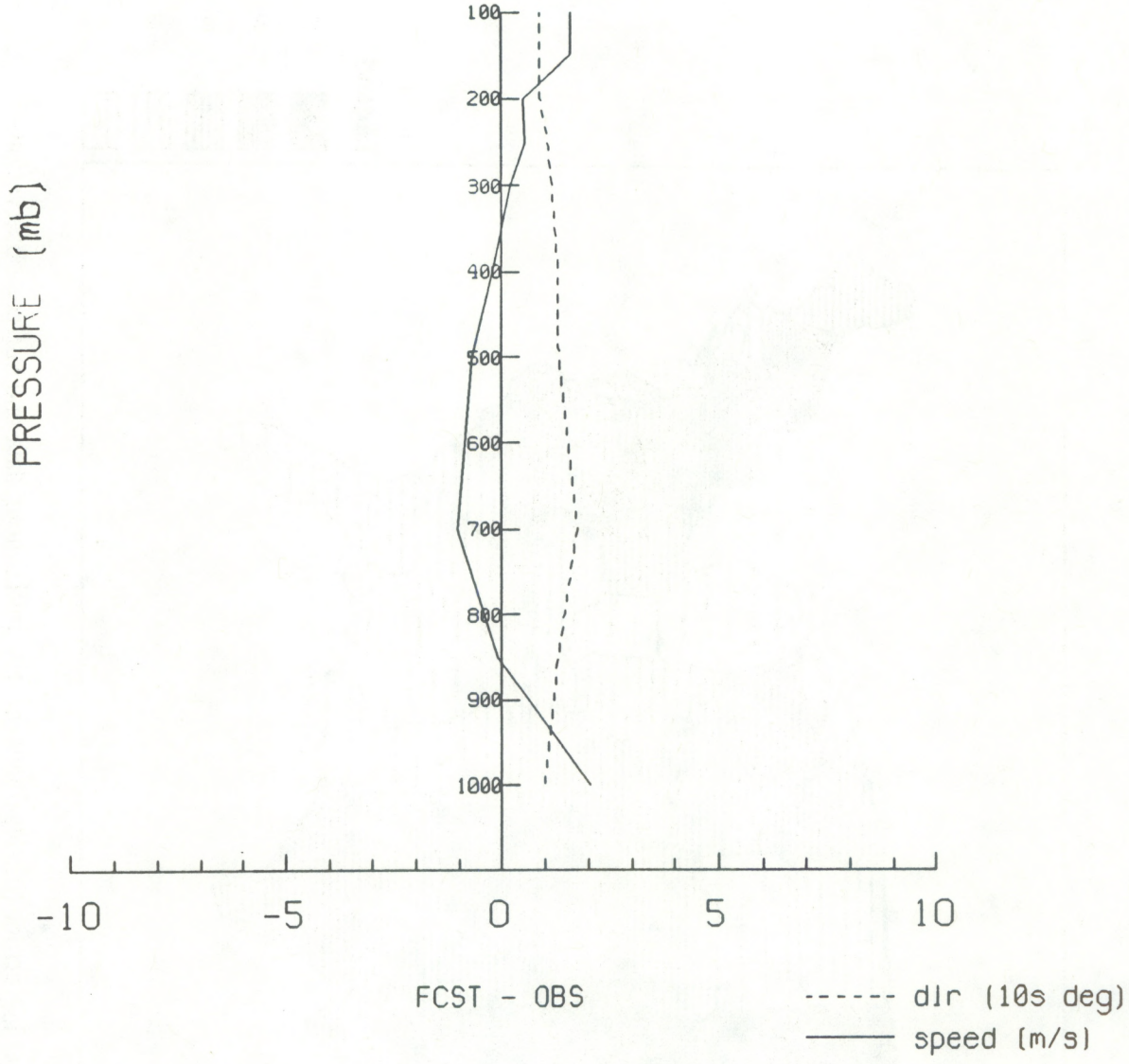

Fig. 30. Same as Fig. 29, except for Eta 12-h forecast data.

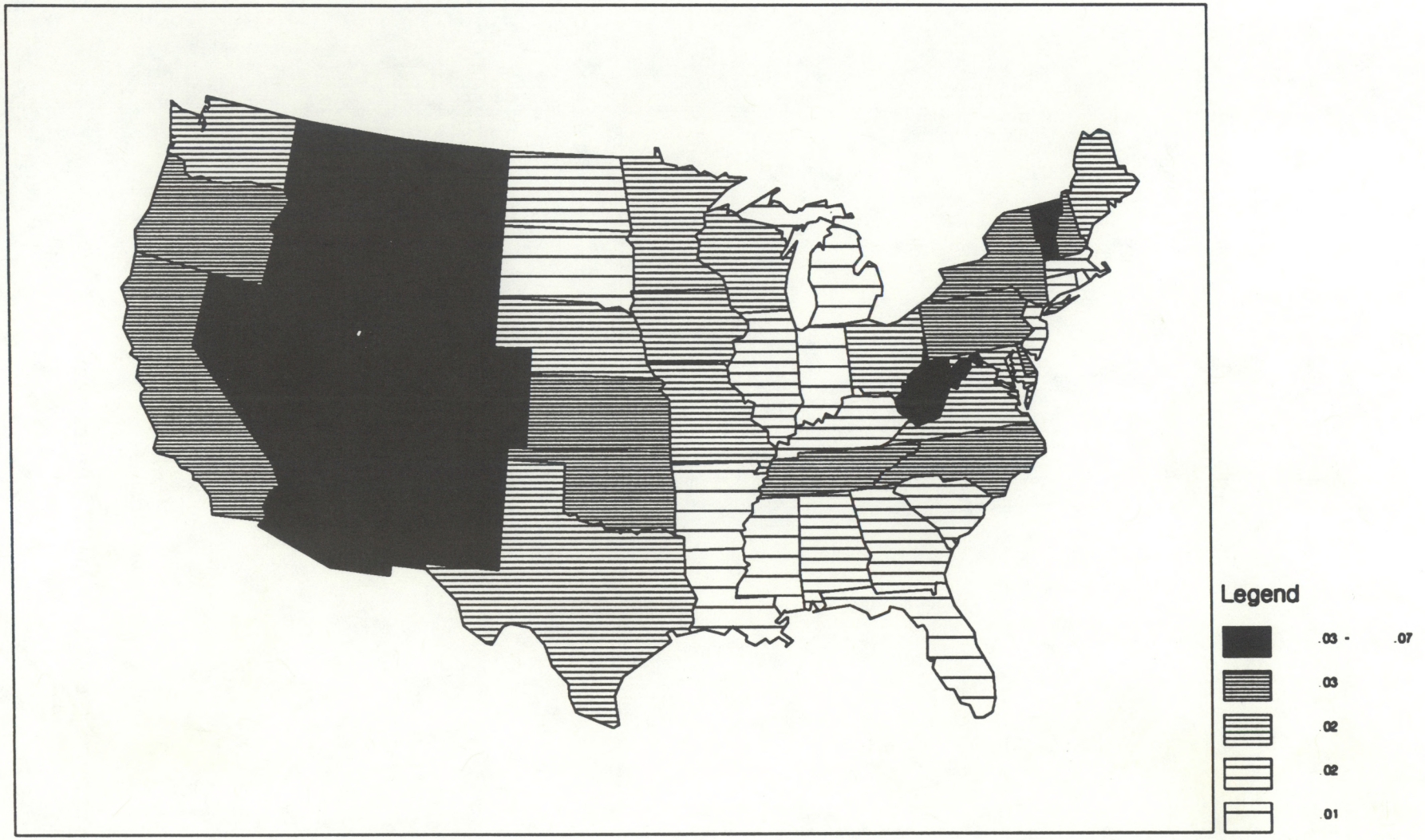

Fig. 31. Map of mean absolute altimeter setting errors (in. Hg) between Eta initial hour data and surface observations.

For altimeter forecasts, the absolute error changes were minor, rising to 0.04 in. by 12 h (Appendix D). The largest errors continue to be in the intermountain West and the Appalachian states. California now has a negative bias. This could be due to extreme local variability in altimeter setting in mountainous regions caused by cold air pooling.

Temperature and dewpoint temperature. The absolute error of 5.78°F for the initial hour temperatures is quite large, and the arithmetic error (-1.41°F) shows a cold bias. An examination of the temperature distribution (Fig. 32) shows that this cold bias results from too many temperatures in the 30 - 40°F range. From the scatterplot (Fig. 33), it is evident that Eta often analyzed a surface temperature of 32°F, even when the observed temperature was above 75°F. A closer examination of the data revealed that this occurred in the mountain stations (e.g., Grand Junction, Casper, Lander, Denver). Obviously there is a problem specifying the initial surface temperature over high terrain. These errors are the result of interpolating between the lowest model level temperature and the skin temperature to the 2-m elevation. That this results in a 32°F temperature (at 2 m) over the mountains is possibly attributed to the skin temperature also being set at 32°F because of the snow cover field in the model. This problem is well illustrated on the United States map of errors (Fig. 34), which shows an extreme cold bias over the Rocky Mountain states. Positive biases are located in the southeastern United States. As seen in the distribution chart (Fig. 32), Eta has observed twice as many 80-90°F temperatures. This makes the overall bias small.

By the 3-h forecast, the absolute error has reduced to 4.90°F and the bias is nearly zero (Appendix D). The scatterplot (Fig. 35) shows that the anomalous line of 32°F Eta temperatures is gone, but a grouping of low temperatures is still evident.

The 6-h absolute error (5.00°F) is slightly worse than at 3 h (Appendix D). By 12 h, the error has increased to 6.76°F, and there is a bias of -5.43°F. Thus, most of the error is the result of underforecasting the temperatures. The geographical distribution of errors at 12 h (not shown) is nearly identical to the initial hour error distribution.

Since all Eta runs began at 0000 UTC, the 12-h forecast is valid at about sunrise, or minimum temperature time. Therefore, it appears that Eta is cooling down too fast overnight. The temperature distribution chart (not shown) indicates 236 Eta temperatures below 20°F, where as only 12 observed temperatures were that low. Of course, many of those errors were from stations initially at 32°F.

The dewpoint temperature errors in Eta are quite large. The absolute error of 7.47°F at initial time increases to 17.86°F by 12 h. Negative arithmetic errors are nearly identical in magnitude to the absolute errors, so it is obvious that the model is affected by a dry bias. The geographical plot of errors (not shown) indicates a few states with positive biases, but by 12 h all states have a negative bias. After 12 h, the errors decrease. This is probably due to the daytime heating (note that Eta was initialized at 0000 UTC for all cases), which increases the surface fluxes and heating. A comparison of the scatterplots at 0 and 12 h (Figs. 36 and 37) is also informative. By 12 h, drying is apparent, especially for observed temperatures

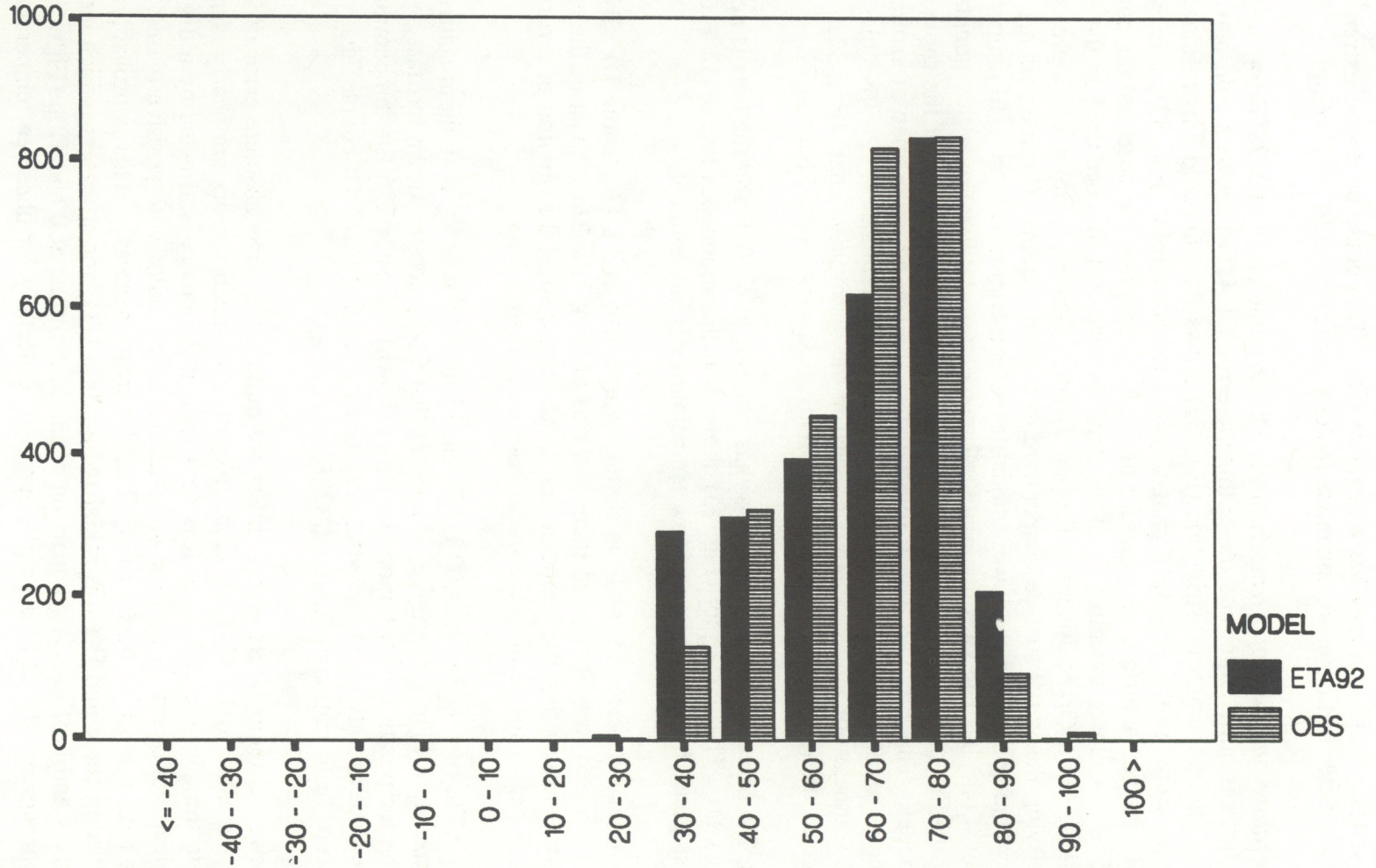

Fig. 32. Distribution of observed surface temperatures and Eta initial hour surface temperatures.

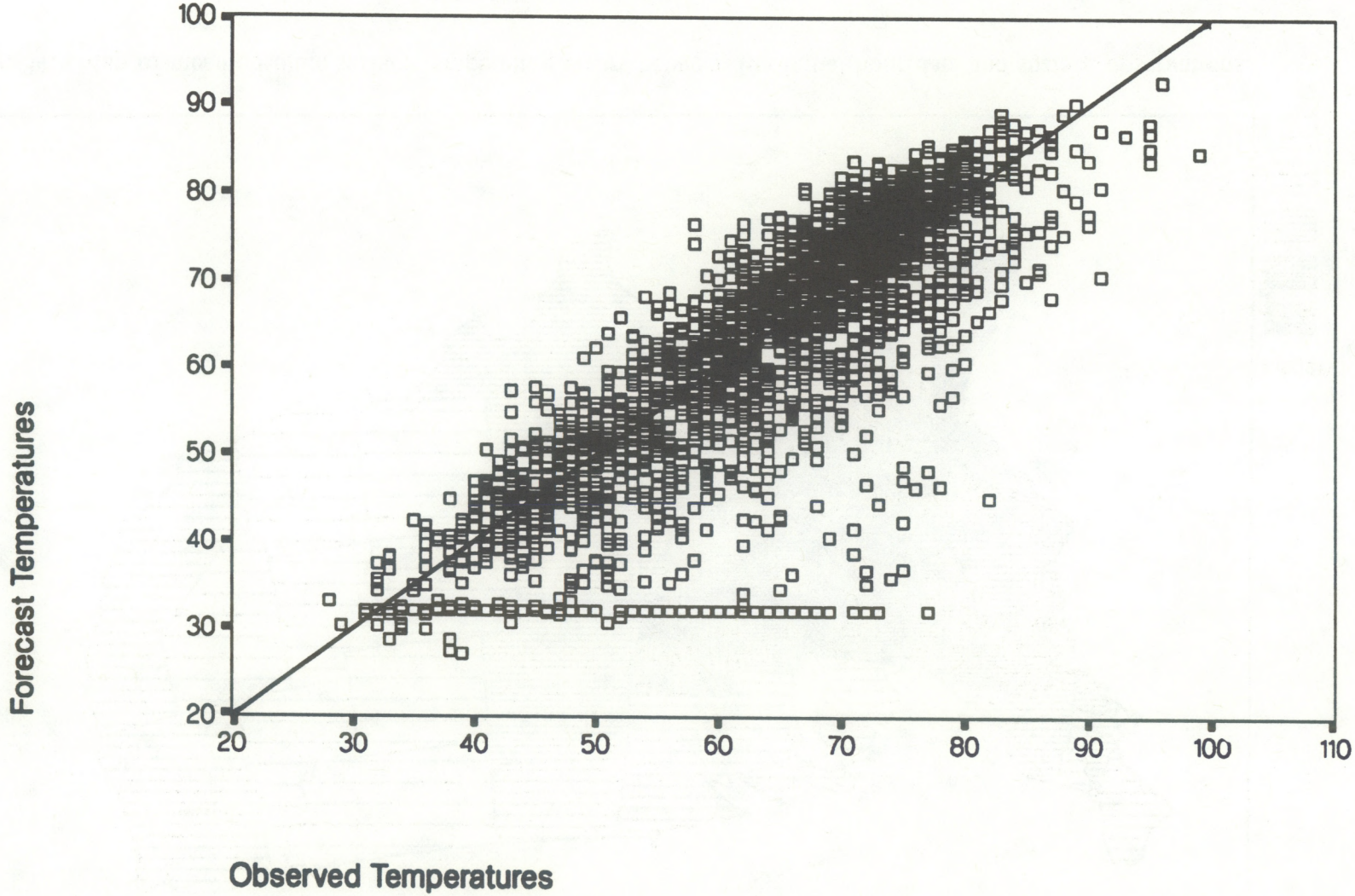

Fig. 33. Scatterplot of observed surface temperatures versus Eta initial hour surface temperatures in °F.

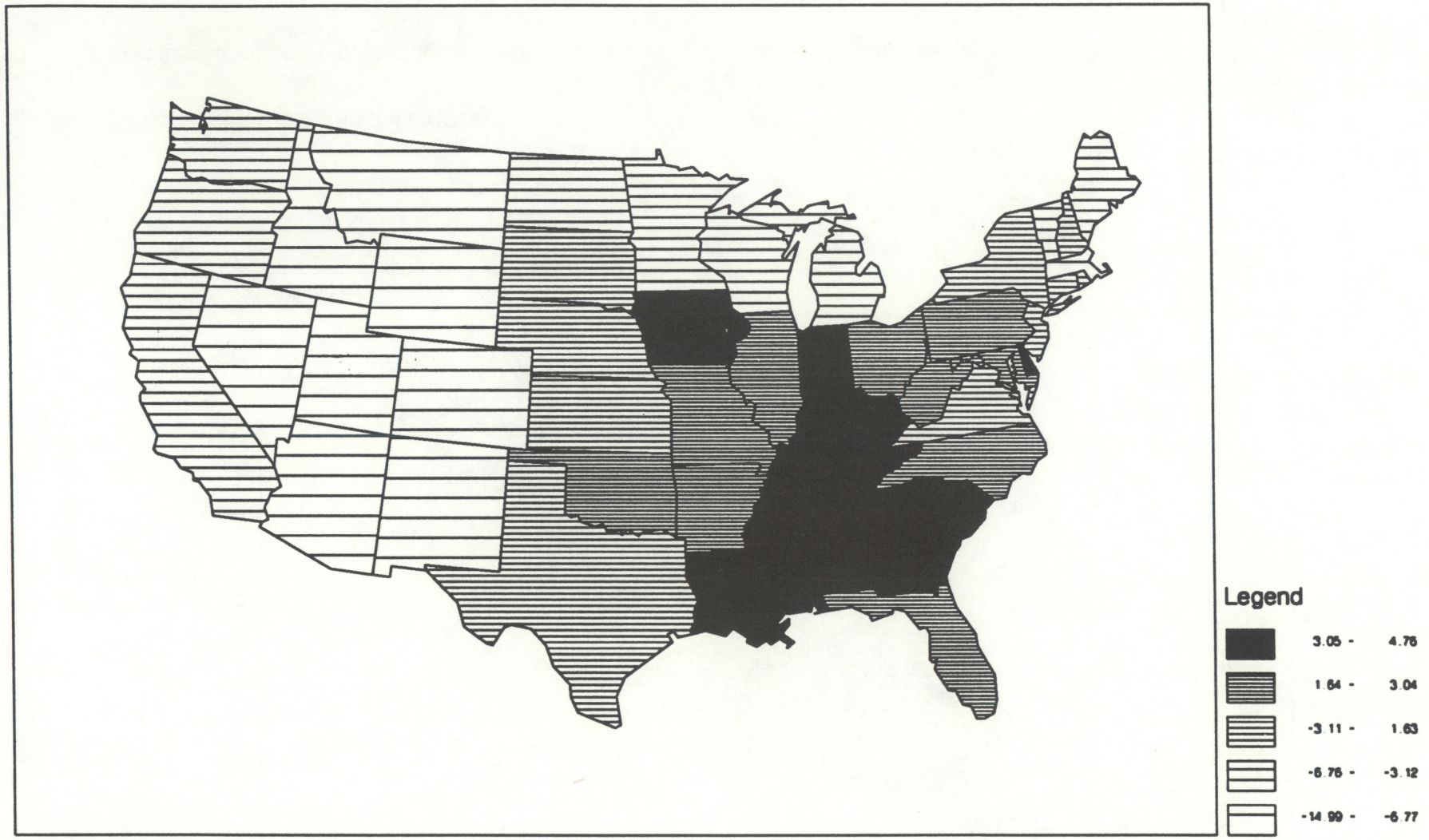

Fig. 34. Map of mean absolute surface temperature errors between Eta initial hour data and surface observations.

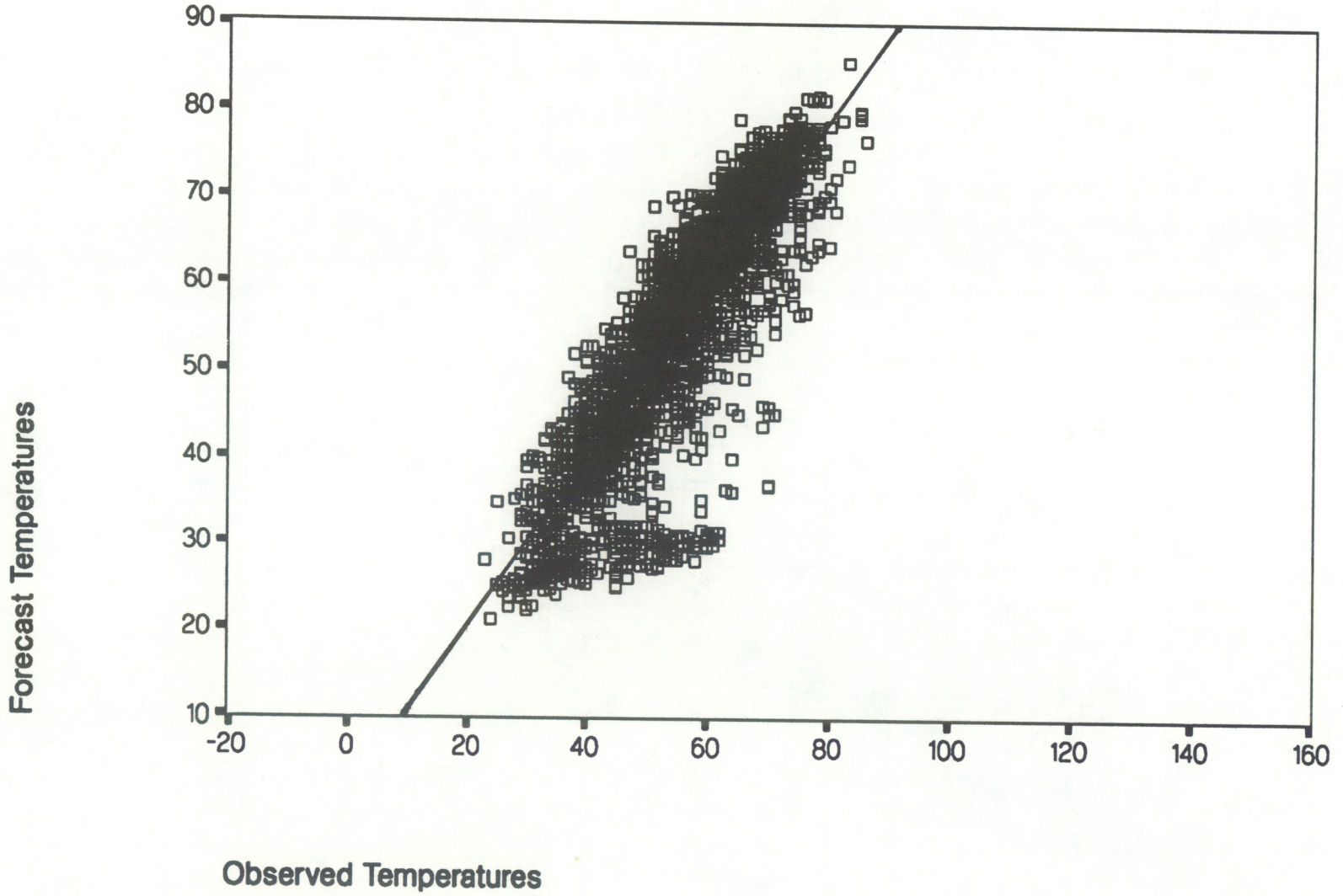

Fig. 35. Scatterplot of observed surface temperatures versus Eta 3-h forecast surface temperatures in °F.

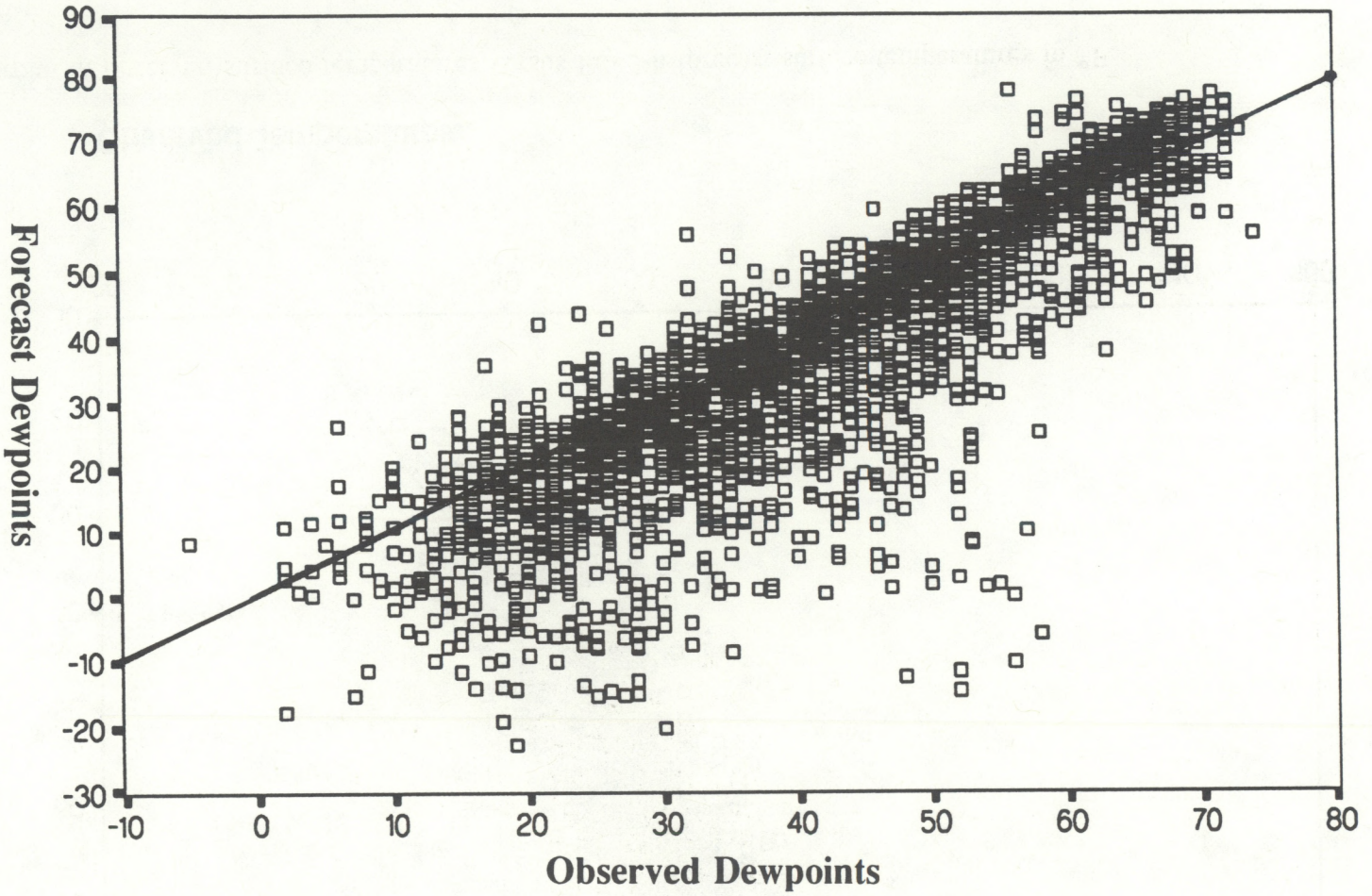

Fig. 36. Scatterplot of observed surface dewpoint temperatures versus Eta initial hour surface dewpoint temperatures in °F.

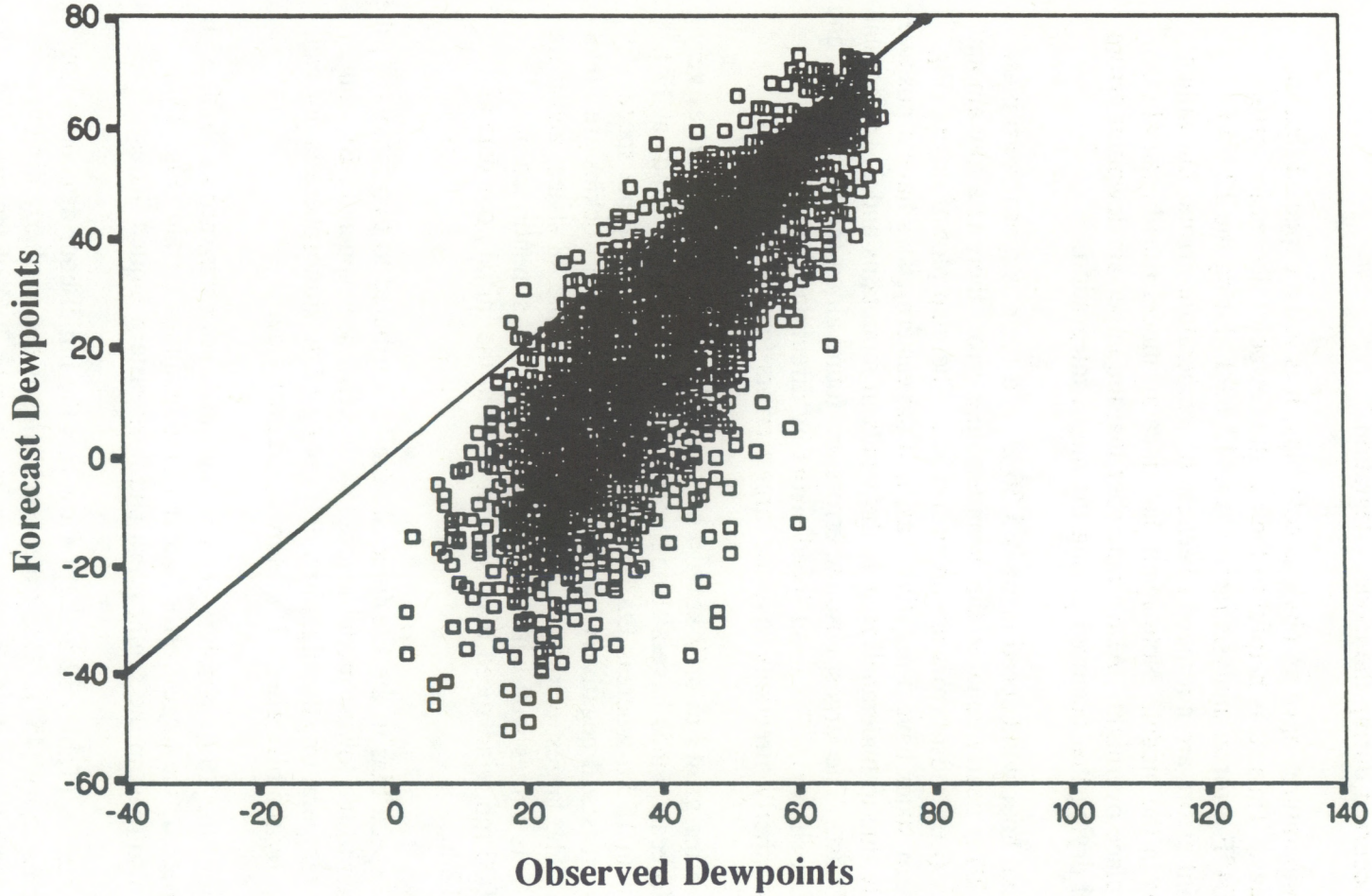

Fig. 37. Same as Fig. 36, except for Eta 12-h forecast data.

below 20°F. The distribution bar charts at 0 and 12 h (Figs. 38 and 39) also show the severe drying that has occurred during the model integration.

Since the Eta temperatures have an obvious cold bias, it is to be expected that the dewpoints have a dry bias. Thus, it is helpful to look at the dewpoint depression errors (Appendix D). The 7.04°F error at analysis increases to 12.69°F during the 12 h of forecast. Thus, dewpoint errors are not merely because of temperature errors, but rather because of a drying bias in the model. Note again that since all model runs began at 0000 UTC, this drying took place overnight. Also, remember that the upper-air dewpoint errors also showed a low-level drying that occurred during the model integration.

Winds. The initial hour wind speed error is 5.36 kt, and there is a near-zero bias. Since the RMSE (6.88 kt) is nearly equal to the mean absolute error, there is a low amount of variance in the errors (i.e., most errors are about 5-6 kt). The map plot of errors (Fig. 40) shows that the western states have a negative bias, and the eastern states have a positive bias. This is most likely due to underanalyzing 1) the winds in mountains and 2) the coastal winds along the West Coast. The largest absolute errors are found along the West Coast and in the northern states (not shown). For wind speed errors by direction (Fig. 41), the southwesterly through northerly directions have the largest errors (greater than 6 kt).

The wind speed error grows to 8.67 kt by 6 h (Appendix D), but then changes very little up to 12 h. The forecast winds have a strong positive bias. Only 116 winds at 12 h were observed greater than 15 kt, whereas Eta forecast 1220 winds to be in that speed category. The West Coast is still negatively biased, though winds in the Rockies are now too strong (not shown). The largest errors are scattered throughout the 48 states, and there is no obvious pattern. This agrees well with the upper-air errors. As mentioned in subsection 6.2.1, the 1000 mb wind speed bias is about -1 kt in the initial hour data and increases to +4 kt by 12 h.

The strong wind speed bias in the Eta forecasts can be attributed to two factors: 1) the surface winds were taken to the lowest model level (150 m AGL, see Appendix B), and 2) similar to the dewpoint temperature, the nighttime decoupling of the atmosphere is not well detected by Eta, as seen by the bias after 12 h returning to nearly zero.

For wind direction, the 26.6° absolute error in the initial hour data grows to 30.5° by 12 h. As would be expected, the largest errors occur in mountainous terrain (Fig. 42). However, large errors also exist in the southeastern and northeastern United States.

The contingency tables (Tables 12 and 13) show that the initial hour directions are usually correct +/- one category (45°). One exception to this is the northerly direction. When Eta diagnosed a north wind, it was observed to be westerly 20% of the time and southwesterly 13.5% of the time. By 3 h into the forecast (not shown), this appears to be somewhat correcting itself (only 7.5% of the northerly directions verified as southwesterly). At 12 h (Table 13), the distribution for northerly winds is about the same as at 3 h. But now

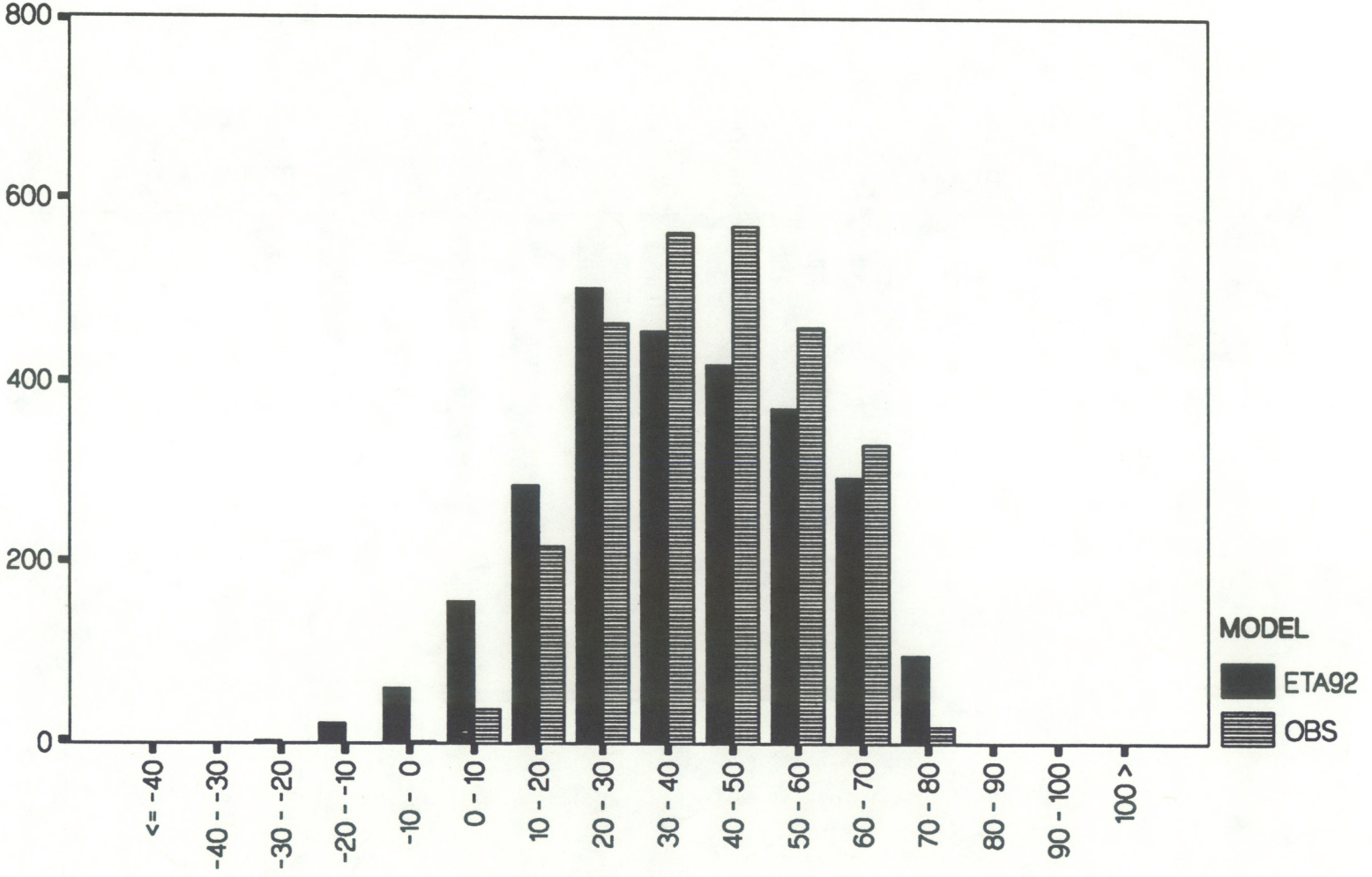

Fig. 38. Distribution of observed surface dewpoint temperatures and Eta initial hour surface dewpoint temperatures in °F.

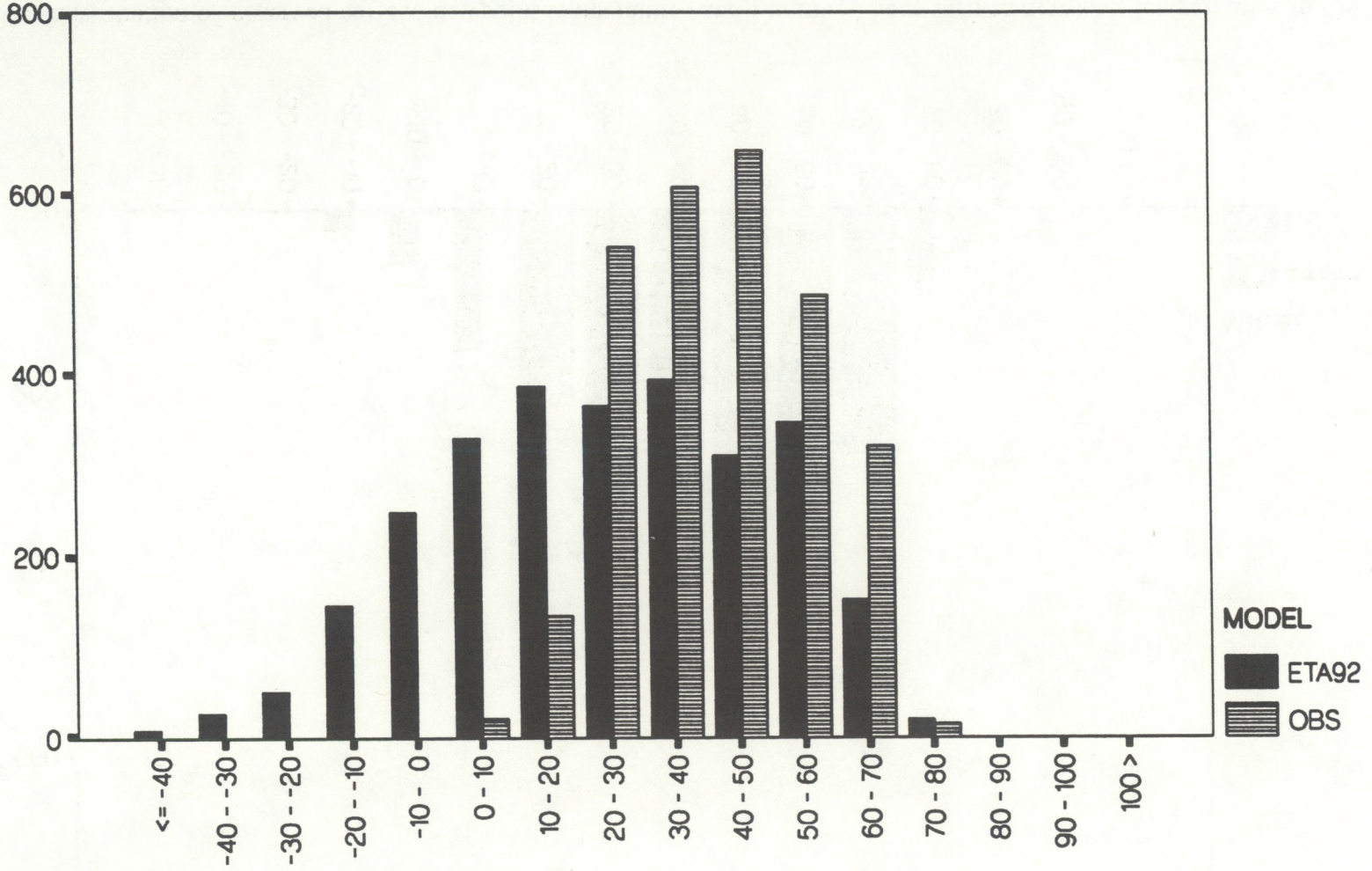

Fig. 39. Same as Fig. 38, except for Eta 12-h forecast data.

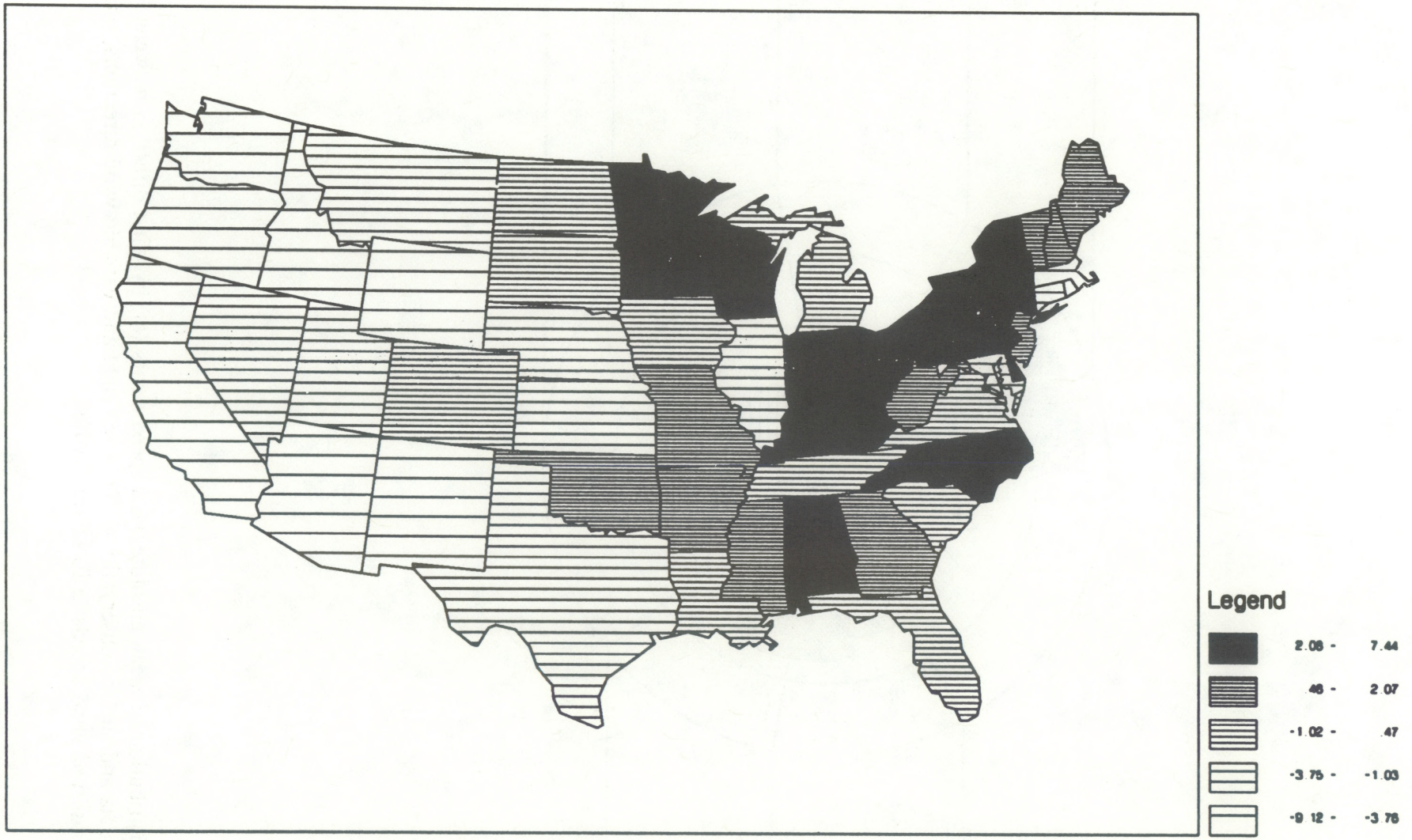

Fig. 40. Map of mean errors (bias) of surface wind speed (kt) between Eta initial hour data and surface observations.

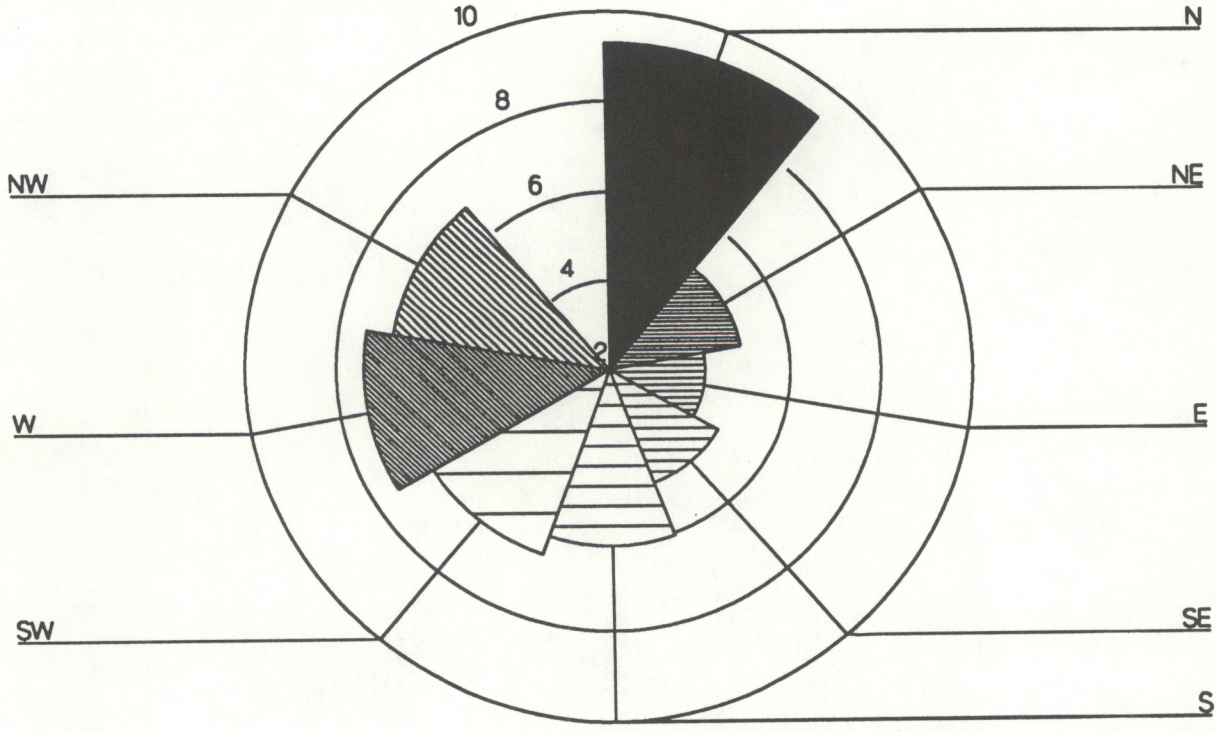

Fig. 41. Distribution of mean absolute wind speed errors (kt) by wind direction between Eta initial hour data and surface observations. To avoid light and variable wind directions, only cases with *observed* wind speeds ≥ 10 kt are verified.

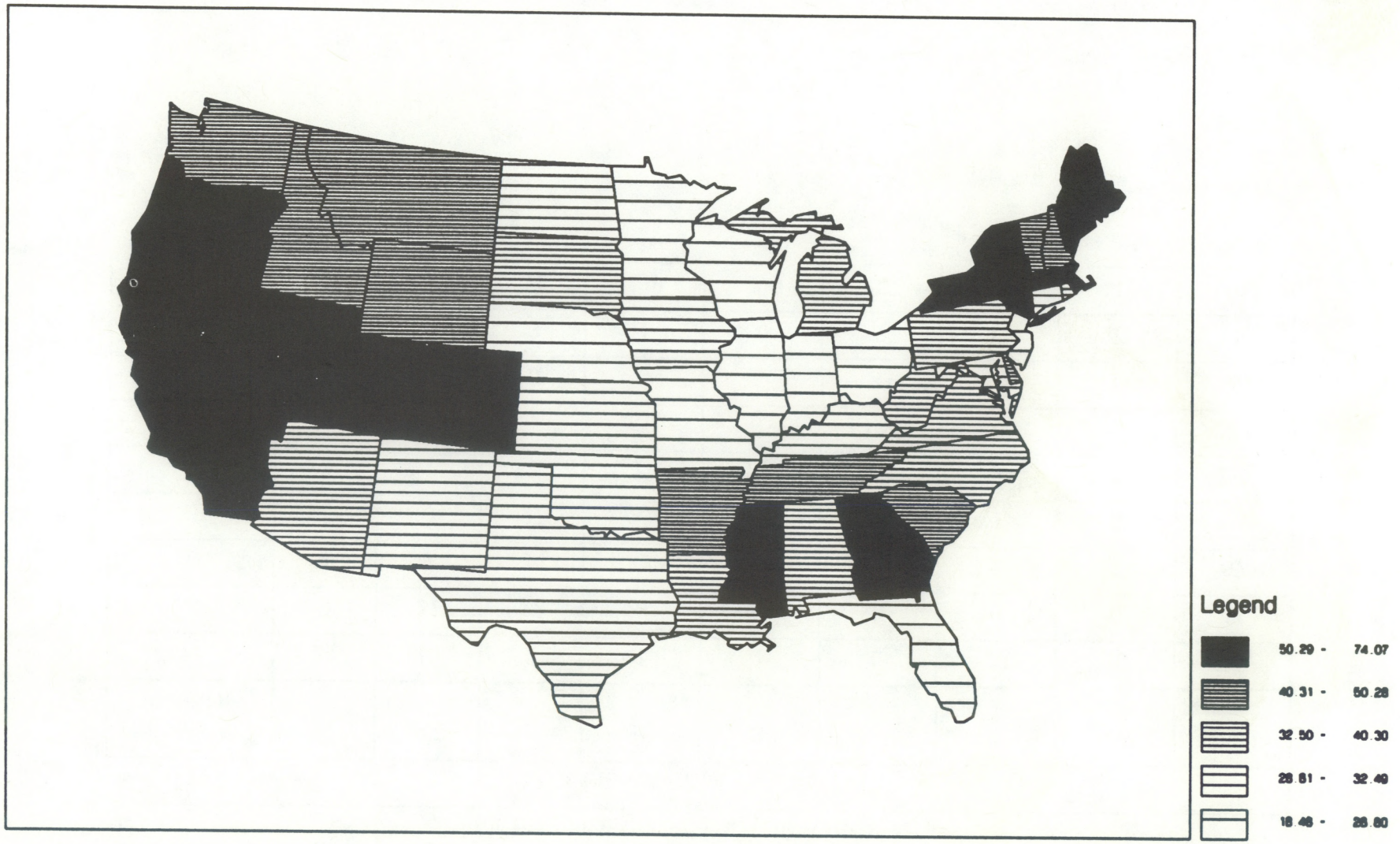

Fig. 42. Map of mean absolute wind direction error (degrees) between Eta initial hour data and surface observations. To avoid light and variable wind directions, only cases with *observed* wind speeds ≥ 10 kt are verified.

Table 12. Contingency Table of Forecast Wind Direction by Observed Wind Direction for Eta Initial Conditions, 1-10 April 1991

	Observed Wind Direction								
Forecast Wind Dir	N	NE	E	SE	S	SW	W	NW	Total
N	**39**	10	2	1	1	13	20	10	96
NE	9	**21**	12	0	20	0	0	0	44
E	4	7	**34**	18	3	0	0	1	67
SE	1	0	22	**87**	35	3	1	0	149
S	3	1	2	22	**175**	27	3	2	235
SW	1	0	1	10	65	**124**	37	6	244
W	8	0	0	0	9	33	**79**	38	167
NW	24	2	3	1	0	7	41	**62**	140
Total	89	41	76	139	290	207	181	119	1142

Table 13. Contingency Table of Forecast Wind Direction by Observed Wind Direction for Eta 12-h Forecasts, 1-10 April 1991

	Observed Wind Direction								
Forecast Wind Dir	N	NE	E	SE	S	SW	W	NW	Total
N	**20**	5	1	0	2	4	7	12	51
NE	12	**10**	0	0	1	0	1	1	25
E	2	10	**21**	6	0	0	0	1	40
SE	0	0	10	**28**	4	0	1	0	43
S	1	2	4	24	**68**	9	1	1	110
SW	0	2	1	4	64	**86**	7	3	167
W	4	1	1	0	23	26	**43**	10	108
NW	9	0	0	0	2	6	25	**26**	68
Total	48	30	38	62	164	131	85	54	612

the forecast westerly winds have verifying southerly winds 21% of the time, as compared to only 5% in the analysis data (Table 12).

6.2.2. AIVs

Surface

Clouds. In the initial hour data, Eta shows either CLR or OVC for low clouds (Fig. 43). Only 0.8% of the stations were diagnosed as SCT or BKN. When we analyze the skill scores for the yes/no analysis, Eta has a high POD (0.825) and a low FAR (0.278). Therefore, Eta does well at low-cloud detection, but has difficulty determining the amount of clouds. This holds true for the Eta forecasts, although the POD and FAR are both higher at 12 h (POD = 0.921, FAR = 0.394).

Because of the high number of cases with overcast low clouds in the initial hour and forecasts, the middle and high clouds were usually unobservable (by Eta). Thus, it is difficult to assess the accuracy of Eta for these cloud heights.

Cloud tops. Cloud-top errors are 8,800 ft at the initial hour, increasing to 12,000 ft at 12 h. The bias is strongly positive. The scatterplot (Fig. 44) shows that most Eta cloud tops are at approximately 35,000 ft. This could mean that Eta has too much high cloudiness. Another piece of data supporting this assumption is the detection of high clouds by Eta. As mentioned above, nearly one-half of the high clouds were unobservable by Eta due to the large amount of low overcast conditions. Yet, of the remaining observable cases, Eta detected more high-cloud occurrences than were observed (1050 versus 891). Also, as seen in subsection 6.2.1 regarding upper-air errors, Eta was too moist above 500 mb, especially in the 12-h forecast.

Ceiling. Since only those cases where the observation and model had a ceiling were verified, the poor distinction of low-cloud amount by Eta does not affect the ceiling verification. For the initial hour data, the ceiling error was 2,900 ft and 83% busts of over 1,000 ft. This error steadily *improves* to 2,200 ft at 12 h and a bust rate of 65%. These errors could be a result of the horizontal or vertical interpolation of the dewpoint temperature.

Precipitation. No precipitation data were available in Eta at the initial hour. The number of precipitation events is largely overforecast by Eta. At 3 h, Eta had nearly twice as many precipitation forecasts as observed (423 versus 222). By 12 h, the ratio had risen to three times as many (778 versus 237; Table 14). This results in high PODs and FARs (0.709 and 0.784, respectively) at 12 h.

Concerning the precipitation phase, Eta had only 3 of 647 liquid forecasts (12 h) verify as frozen (Table 15). On the other hand, of the 43 frozen forecasts that verified as some form of precipitation, nearly one-half verified as liquid. Since most precipitation falls

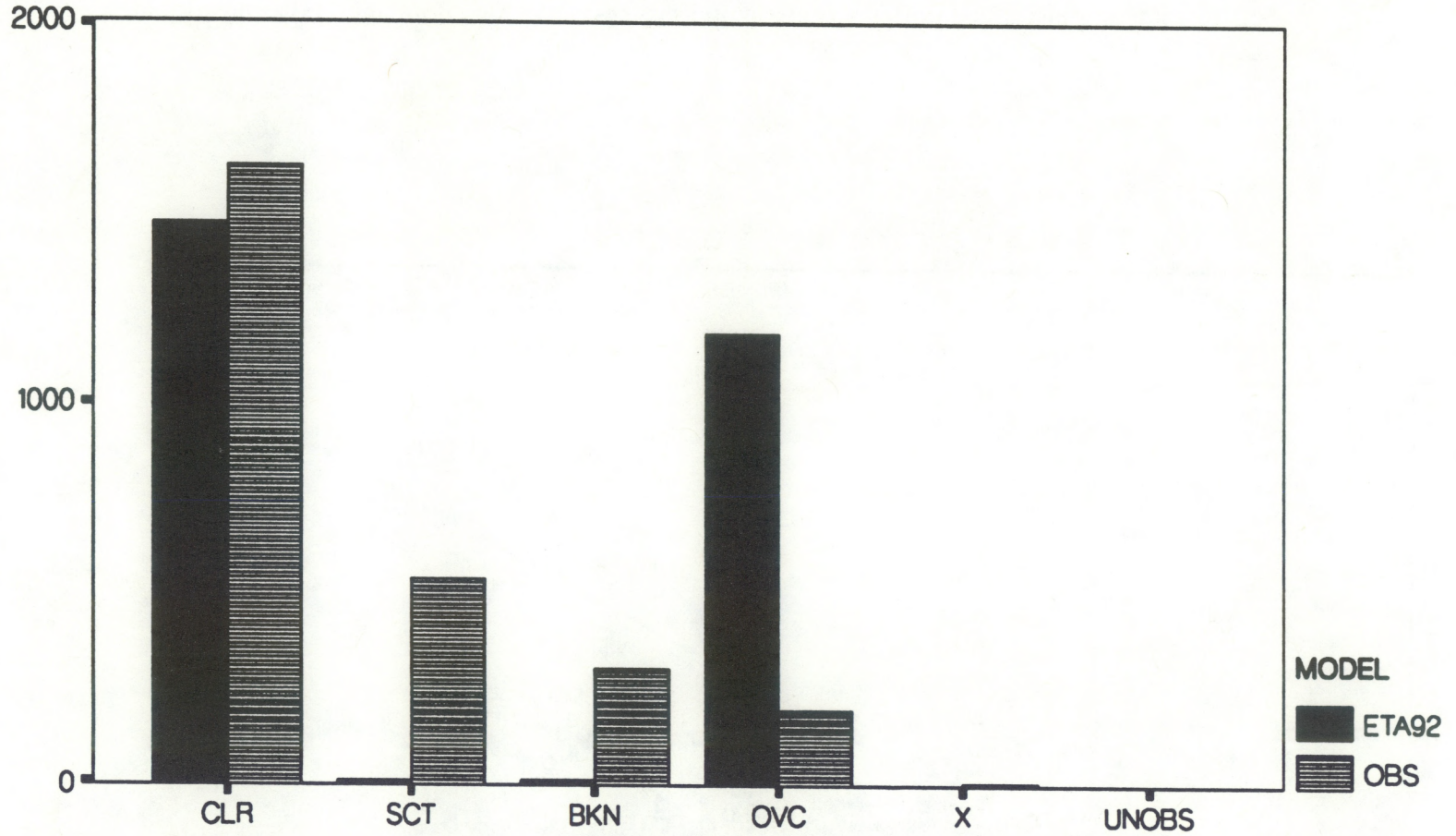

Fig. 43. Distribution of observed and Eta initial hour low cloud amounts. UNOBSD (unobservable) is defined as those cases when a lower overcast deck prevents the observation of any (higher) clouds.

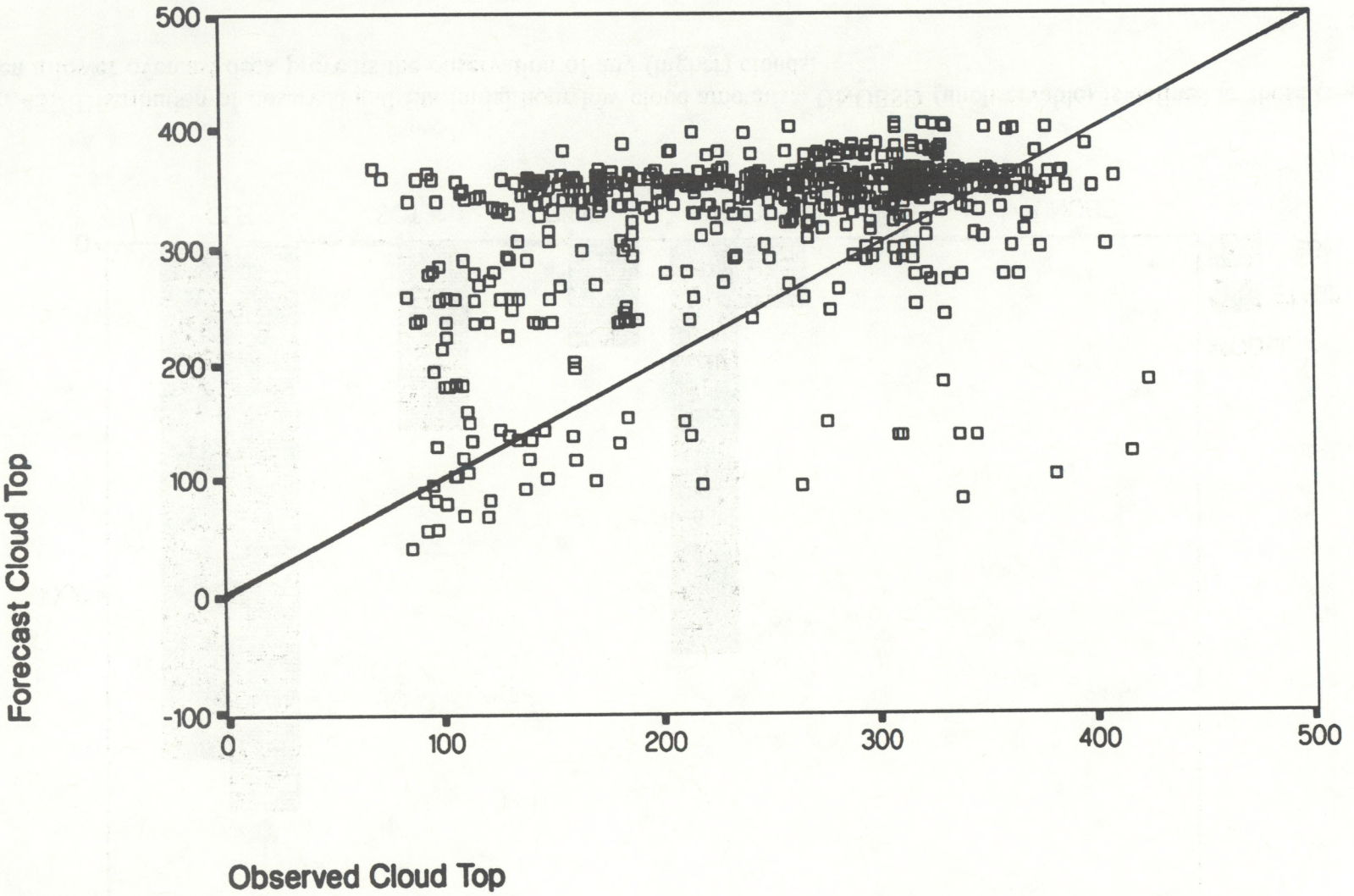

Fig. 44. Scatterplot of observed cloud-top heights versus Eta initial hour cloud-top heights in hundreds of ft.

Table 14. Contingency Table of Forecast Precipitation Occurrence by Observed Precipitation Occurrence for Eta 12-h Forecasts, 1-10 April 1991

	Observed Precipitation Occurrence		
Forecast Precipitation Occurrence	Yes	No	Total
Yes	**168**	610	778
No	69	**1979**	2048
Total	237	2589	2826

as liquid, the 50% bust rate of the frozen precipitation forecasts is probably more indicative of phase determination skill by Eta. The poorly initialized surface temperatures may have played a role in this, since the surface temperature was used to discriminate between liquid and frozen precipitation.

Table 15. Contingency Table of Diagnosed Precipitation Phase by Observed Precipitation Phase for Eta 12-h Forecasts, 1-10 April 1991

	Observed Precipitation Phase				
Diagnosed Precip. Phase	None	Liquid	Frozen	Mixed	Total
None	**1979**	62	6	1	2048
Liquid	522	**122**	3	0	647
Frozen	88	21	**21**	1	131
Mixed	0	0	0	**0**	0
Total	2589	205	30	2	2826

Eta precipitation amount errors show a surprising consistency out to 12 h and beyond (Appendix D). At 3 h, the absolute error is 0.09 in., and there is a strong tendency to underforecast. By 12 h, the error and bias are only slightly larger. Eta had only one 12-h forecast of more than 0.25 in. (in a 3-h accumulation), whereas the verifying observation was about 0.45 in. (Fig. 45). Some of the larger events (i.e., greater than 0.7 in.) were completely missed by the model. Thus, although Eta had precipitation occurring three times too often, the amounts were light and the model failed to forecast the big events. The areal

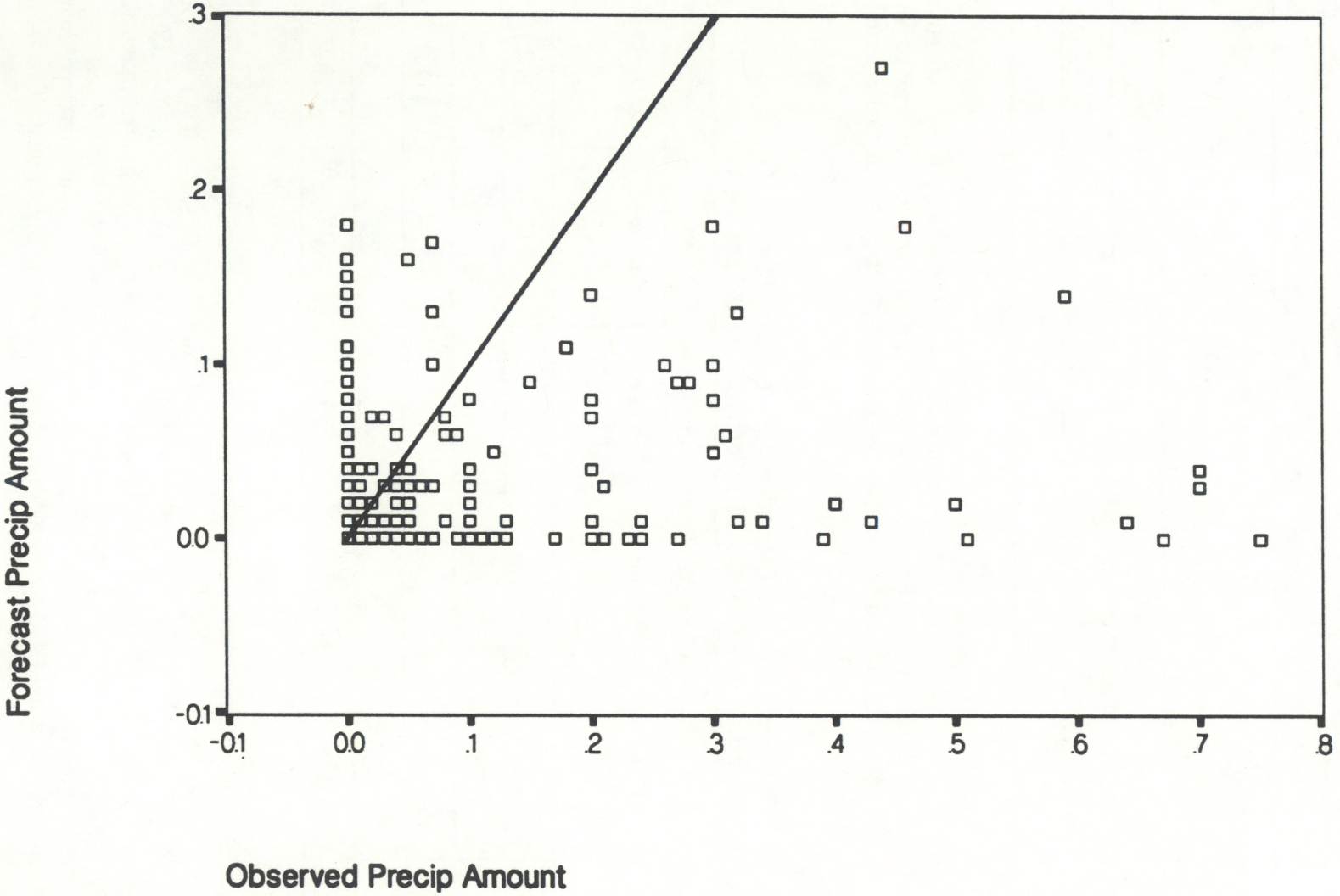

Fig. 45. Scatterplot of observed 3-hour precipitation amounts versus Eta 12-h forecast 3-h precipitation amounts in in.

distribution of errors follows that of the observed precipitation amount. The largest errors occurred in the Pacific Northwest and the Southeast.

6.2.3. Discussion

The Eta model forecasts the upper-air variables (height, temperature, moisture, winds) very well. The only analysis or forecast problem worthy of mention is moisture. The moisture bias above 600 mb at the initial time increases during the forecast, and the lower levels dry out. This has obvious effects on surface temperature (via radiation), clouds, and precipitation. Thus, although Eta appears to be performing very well above the boundary layer, the problems with moisture could have an adverse effect on many surface parameters.

Eta has a problem with surface temperatures in mountainous areas. This is largely due to the interpolation scheme used for this exercise, which included using the surface skin temperature as an input (see Appendix B). This led to cold biases at initialization. In Eta forecasts, the cold bias remained, even out to 24 h. Eta also had too many warm temperatures (above 80°F) in the analysis, which helped to offset the cold bias figures.

Eta's moisture problems aloft translated to the surface as a severe dry bias. This deficiency showed up even when dewpoint depressions were considered (to remove the temperature error). Also, Eta forecast biases improved after 12 h due to the increased mixing during the daytime.

Eta detected the presence of low clouds very well, but could not discern the amount of low clouds. It was also concluded that Eta had too much high cloudiness, which agrees with the moist bias at upper levels.

Precipitation was overforecast by Eta in terms of occurrence, but well underforecast in terms of amount. The phase forecasting also appears poor, which may be related to the surface temperature errors.

6.3. LAPS

Appendix E lists the summary of LAPS statistics for the SAVs and AIVs at the surface and at the rawinsonde levels of 500 and 250 mb. The following subsections discuss further details of the LAPS evaluation, grouped by SAVs or AIVs, for the surface and upper-air.

6.3.1. SAVs

Upper air

Since the only upper-air station within the LAPS domain is Denver, there are no verification data below 700 mb. Further, since Denver RAOB and the Platteville PRF are

the only upper-air data observation sources within the LAPS domain, generalizations about the statistics in this report should be made with caution.

A noteworthy point, though, is that the Denver RAOB is not actually analyzed by LAPS. Instead, it is entered indirectly into the model via the MAPS background field. But since the MAPS 0000 or 1200 UTC analysis is not available when the LAPS 0000 or 1200 UTC analysis is performed, the previous 3-h forecast from MAPS must be used. Because this forecast begins 9 h (or three 3-h forecasts) after the previous RAOB time, it is safe to say that the Denver RAOB is not used in LAPS. Thus, the Denver RAOB is an *independent* data source. The rest of the data used in this verification are incorporated into the LAPS analysis and thus constitute a dependent data set. For example, the Platteville wind profiler data are incorporated into the LAPS three-dimensional wind analysis (along with radar and MAPS background winds). Thus, to have the LAPS wind data interpolated back to the wind profiler site may sound redundant. However, checking the accuracy answers the question of how well the blended analyses fit the observational data. Furthermore, with MAPS acting as the background field for LAPS upper-air grid points, the accuracy of the MAPS analyses and forecasts obviously has a large effect on the LAPS analysis, and thus the statistics. In fact, above 500 mb the LAPS height and temperature field is the MAPS background field interpolated to the LAPS grids, and no additional information is added by the LAPS analyses.

Height. Height errors are essentially zero (Fig. 46) in the lower troposphere (+15 m at 700 mb). Bias dramatically increases above 200 mb to nearly +50 m at 100 mb. The absolute error at 500 and 250 mb is about 15 m, and there are small positive biases. The RMSEs at these two levels (see Appendix E) are much larger than the MAE (24 and 22 m, respectively), implying a large variance in the error.

Temperature and dewpoint temperature. The temperature bias profile (Fig. 46) is nearly identical to the height profile. The only marked biases exist above 200 mb and are strongly positive. The absolute errors at 500 and 250 mb are in the neighborhood of 1°C (Appendix E).

A coding error existed in the LAPS moisture analysis during April 1991, which prevents an analysis of the LAPS upper-air dewpoint temperature. This coding error did not affect the LAPS cloud analysis or any of the surface data.

Winds. The u,v wind component errors (not shown) do not indicate any definitive biases except for a positive u bias of 2 m s^{-1} above 250 mb. The absolute errors at 500 and 250 mb (Appendix E) are about 1.5 and 3 m s^{-1}, respectively. For the overall wind speed, the winds are too light in the lower levels (700 and 500 mb) and too strong above 200 mb. The wind speed RMSE (Fig. 47) shows little change with height.

Comparison with the profiler winds shows very good agreement with the LAPS analysis for u and v. This is somewhat expected since the profiler data are not independent

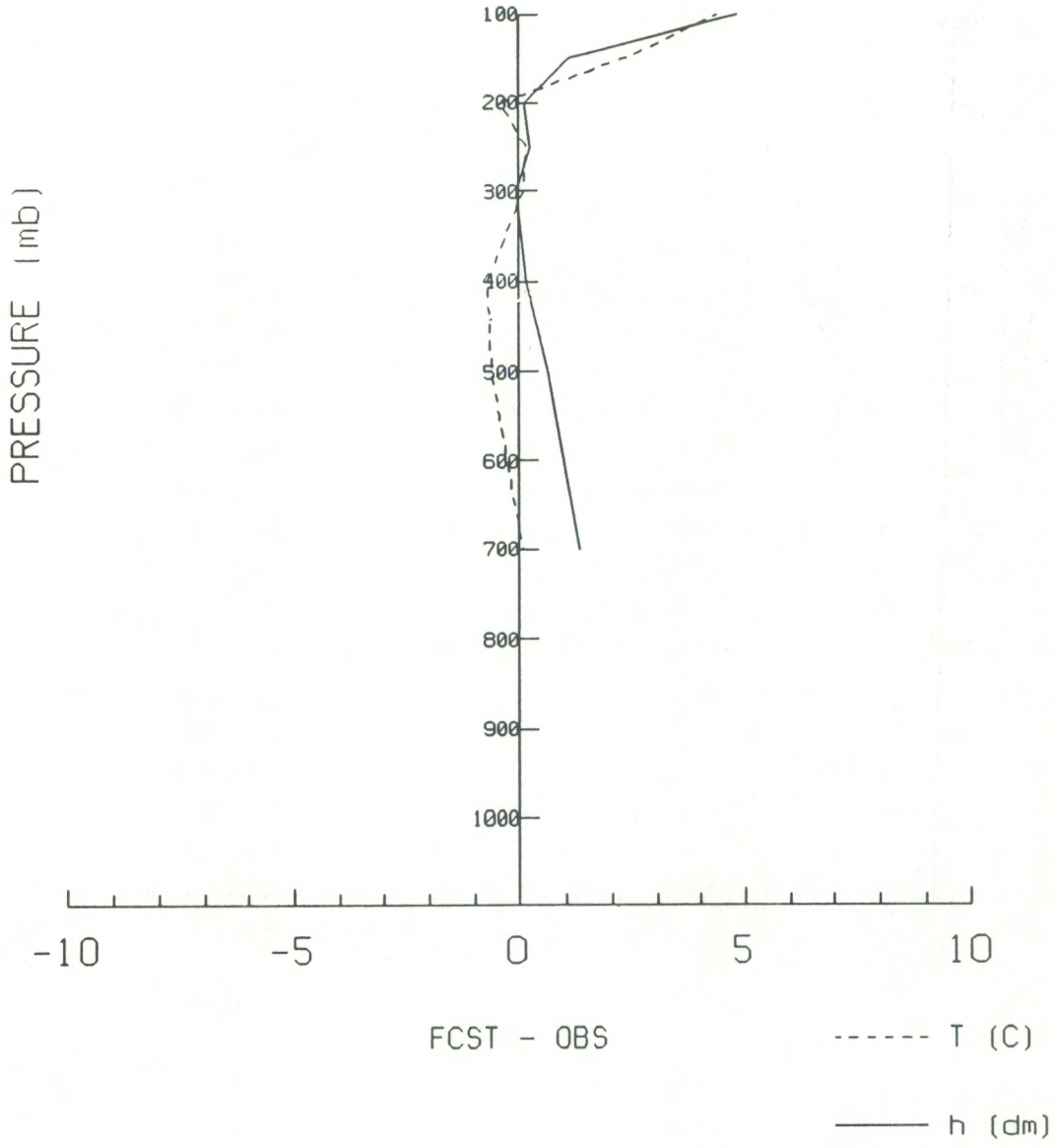

Fig. 46. Mean error (bias) of height (h) and temperature between LAPS and Denver RAOB.

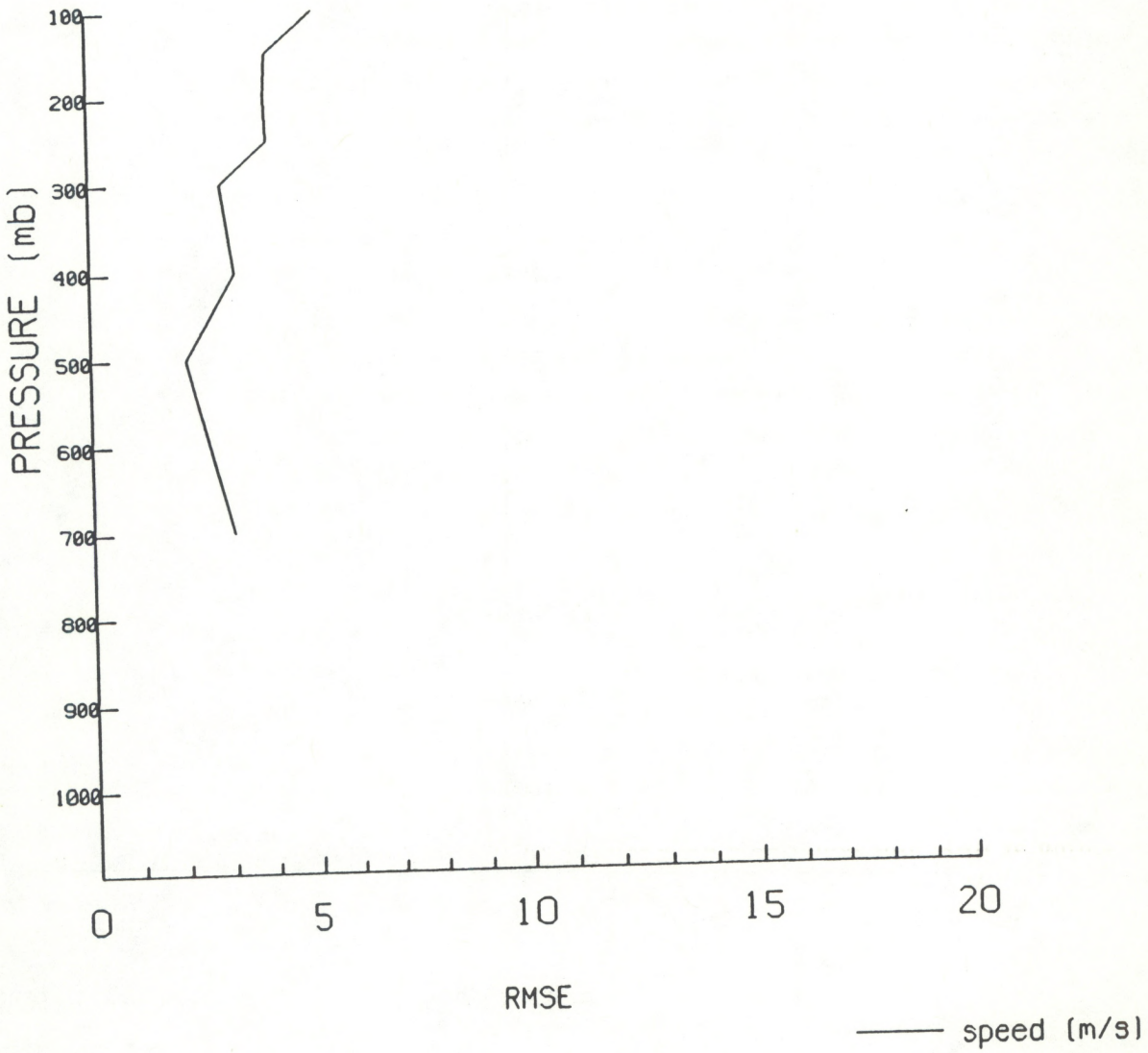

Fig. 47. Root mean square error (RMSE) of wind speed between LAPS and Denver RAOB.

observations (i.e., the profiler data were used by LAPS analysis). The vertical wind (w) has a near-zero bias in the low levels (below 3 km) but then rapidly displays a strong negative bias (Fig. 48). This is most likely due to the kinematic approach used by LAPS to compute vertical motion, which requires a vertical velocity of zero at the top of the model. For horizontal wind speed and direction, the results are nearly identical to the rawinsonde comparison.

Surface

Altimeter. In general, the LAPS altimeter analyses agree well with the observed values. The absolute error is 0.06 in. (Table 16) and there is a slight positive bias (+0.02 in.). However, there were large errors (absolute error greater than 0.20 in.) nearly 1% of the time. Most of these occurred at Scottsbluff, Nebraska (BFF), where 5.3% of the LAPS analyzed values were busts. Since this station is very close to the northern boundary of the LAPS domain, boundary problems with the analysis routine could contribute to this error.

Note also that the high elevation stations (LAR, CYS, and COS) have a negative bias, whereas the lower elevation stations have a positive bias. This could indicate that there is a problem in obtaining the altimeter setting from LAPS due to the highly variable terrain.

Temperature and dewpoint temperature. A scatterplot (not shown) indicates LAPS analyzes temperature very well. The absolute error is only 1.3°F (Table 16), which is well within the instrument precision. The bias is nearly zero. The absolute error at Trinidad, Colorado (TAD), is 2.8°F, and there is an identical positive bias. As with the altimeter setting at BFF, this could be a boundary problem associated with the analysis routine.

The absolute dewpoint error (1.2°F) is comparable to the air temperature error. But the scatterplot (not shown) shows more variance, as do the RMSE values (Appendix E). The bias is also small, but positive, which may be a result of the small positive bias of the temperature rather than a moist bias. The larger absolute dewpoint depression error of 1.8°F shows that LAPS is less accurate in the actual moisture analysis than in the temperature analysis.

As was the case with the air temperature error at TAD, the dewpoint temperature error at TAD is also large (2.2°F). Two other stations with larger errors are Akron (AKO) and Colorado Springs (COS); the LAPS dewpoint was in error at these stations by more than 10°F 4% and 5% of the time, respectively. The large error at AKO can be explained as follows. LAPS frequently analyzes a sharp moisture gradient in the vicinity of AKO. Meteorologically, it is typical in northeastern Colorado to find the eastern slopes dry and moisture near the Kansas border. This could lead to the large error if the gradient is slightly misplaced. The large errors at COS could be due to the close locations of two other stations, which may report different dewpoints. The analysis could then average these three stations, which could result in a large error at COS. This averaging affect can also be seen in the dewpoint errors comparing DEN and APA, which are about 10 miles apart. Note, however,

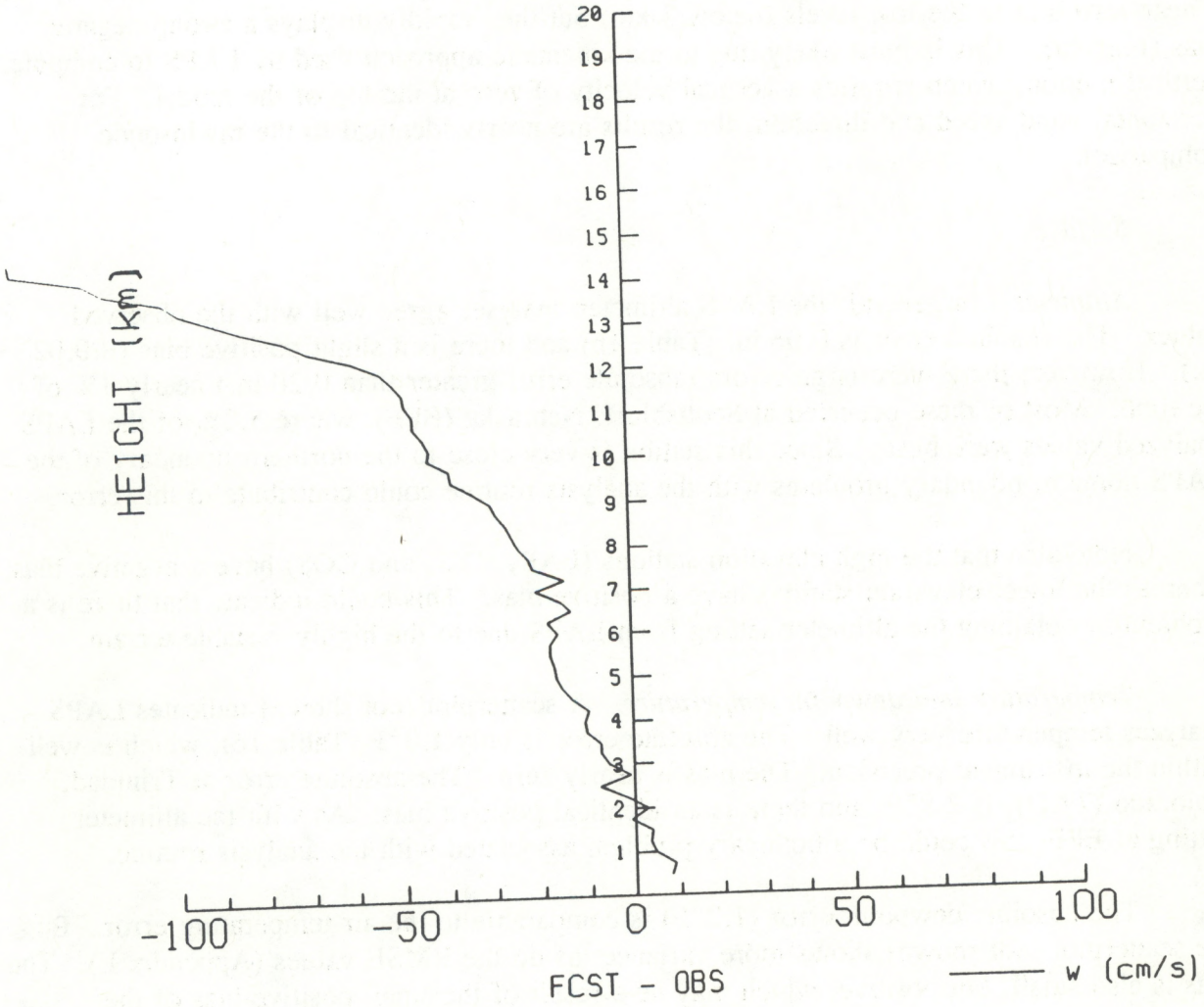

Fig. 48. Mean error (bias) of vertical wind speed (w) between LAPS and Platteville wind profiler.

Table 16. LAPS Absolute (Arithmetic) Errors by Station

	Altimeter (in. of Hg)	Temperature (°F)	Dewpoint (°F)	Wind Speed (kt)	Wind Direction (degrees)	Ceiling (100s of ft)
GLD	0.07 (+0.06)	0.89 (-0.14)	0.71 (+0.37)	1.97 (-1.68)	11.6	4.71 (-2.14)
BFF	0.10 (+0.09)	0.77 (+0.14)	0.67 (+0.06)	1.36 (-1.12)	6.3	7.43 (-1.43)
AKO	0.04 (+0.03)	1.40 (-1.29)	1.72 (+0.33)	1.30 (+0.21)	8.2	8.97 (-0.24)
LAR	0.06 (-0.05)	1.39 (+1.21)	0.77 (+0.44)	1.70 (-1.32)	8.2	17.24 (-1.94)
CYS	0.04 (-0.03)	0.88 (+0.12)	0.46 (+0.42)	1.43 (-1.04)	5.6	6.00 (-3.00)
APA	0.05 (+0.05)	1.15 (+0.20)	1.87 (-1.39)	3.01 (-2.83)	13.1	5.57 (3.22)
DEN	0.03 (+0.02)	1.04 (-0.80)	1.43 (+0.78)	2.65 (-2.58)	10.3	14.73 (-13.64)
COS	0.07 (-0.07)	1.07 (-0.21)	1.69 (+0.84)	1.50 (-1.12)	6.3	10.09 (-8.21)
LHX	0.06 (+0.05)	1.29 (-0.48)	0.93 (+0.15)	1.81 (-1.41)	9.8	10.25 (-0.08)
TAD	0.04 (+0.02)	2.80 (+2.78)	2.18 (+1.90)	2.11 (-1.60)	7.9	15.33 (-13.50)
TOTAL	0.06 (+0.02)	1.26 (+0.15)	1.24 (+0.39)	1.88 (-1.45)	8.7	9.46 (-3.26)

Note that the sample size for these first four parameters varies from 214 to 228 for each station. Since wind direction was verified only if the wind speed was $\geq$10 kt, the sample size for wind direction varies from 45 to 171. Similarly for ceiling, which was verified only if the observed ceiling was < 8,000 ft, the sample size ranges from 7 to 35.

that the biases at these two stations are of opposite signs. Examination of the data shows that APA has a higher dewpoint than DEN 60% of the time, and an average difference of 5.4°.

Winds. As with temperatures, the wind speed absolute error of 1.9 kt (Table 16) is within the instrument precision. A nearly equal negative bias tends to indicate that LAPS wind speeds are usually too weak. The distribution bar chart (Fig. 49) shows that LAPS has twice as many light winds (0 - 5 kt) as observed and only about one-half as many strong winds (greater than 20 kt).

Looking at the errors for individual stations (Table 16), we see that Denver (DEN) and Centennial (APA) have the largest errors (2.6 and 3.0 kt, respectively). These two stations are only two grid points apart in the LAPS grid. These comparatively large errors may be a result of LAPS trying to resolve differences between the two stations by essentially averaging them. Another point to consider is the collocation of the Aurora mesonet observation at DEN. This doubles the weight of the DEN observation, which places more influence on Denver's winds than on those of Centennial. Negative biases at these two stations are nearly equal in magnitude to the absolute errors, and the winds are commonly too weak.

For wind direction, LAPS is on average 8.7° in error (Table 16), which is close to the instrument precision of 5°. From the contingency table (not shown), the bias appears to be negligible, except for southeasterly and southerly winds, where LAPS verified three to four times more often in the clockwise category (south and southwest) than in the counterclockwise category (east and southeast).

As with the wind speed errors, DEN and APA have the larger direction errors (10.3° and 13.1°, respectively). However, the standard deviation at APA (18.6°) is much larger than that at DEN (9.6°), implying that the wind direction at APA is often altered in LAPS because of the wind direction at DEN. As with the wind speed, this may be due to the collocation of the mesonet station at DEN, resulting in more weight given to the DEN observation. GLD also has a large directional error (11.6°), which is most likely due to the boundary problem during analysis.

6.3.2. AIVs

Surface

Clouds. For all three cloud levels, LAPS tends to analyze either CLR or OVC (Fig. 50). For low clouds only (Table 17), LAPS has very high PODs for CLR and OVC (0.986 and 0.950, respectively), but the PODs for SCT (0.486) and BKN (0.322) are quite low. (Note that LAPS does not analyze obscurations, X, but includes this category within OVC category.) The low-cloud occurrence statistics show that LAPS does very well at detecting low clouds with a high POD (0.866) and very low FAR (0.067). From the contingency table, if LAPS says CLR or SCT, the sky cover by low clouds is rarely more than that.

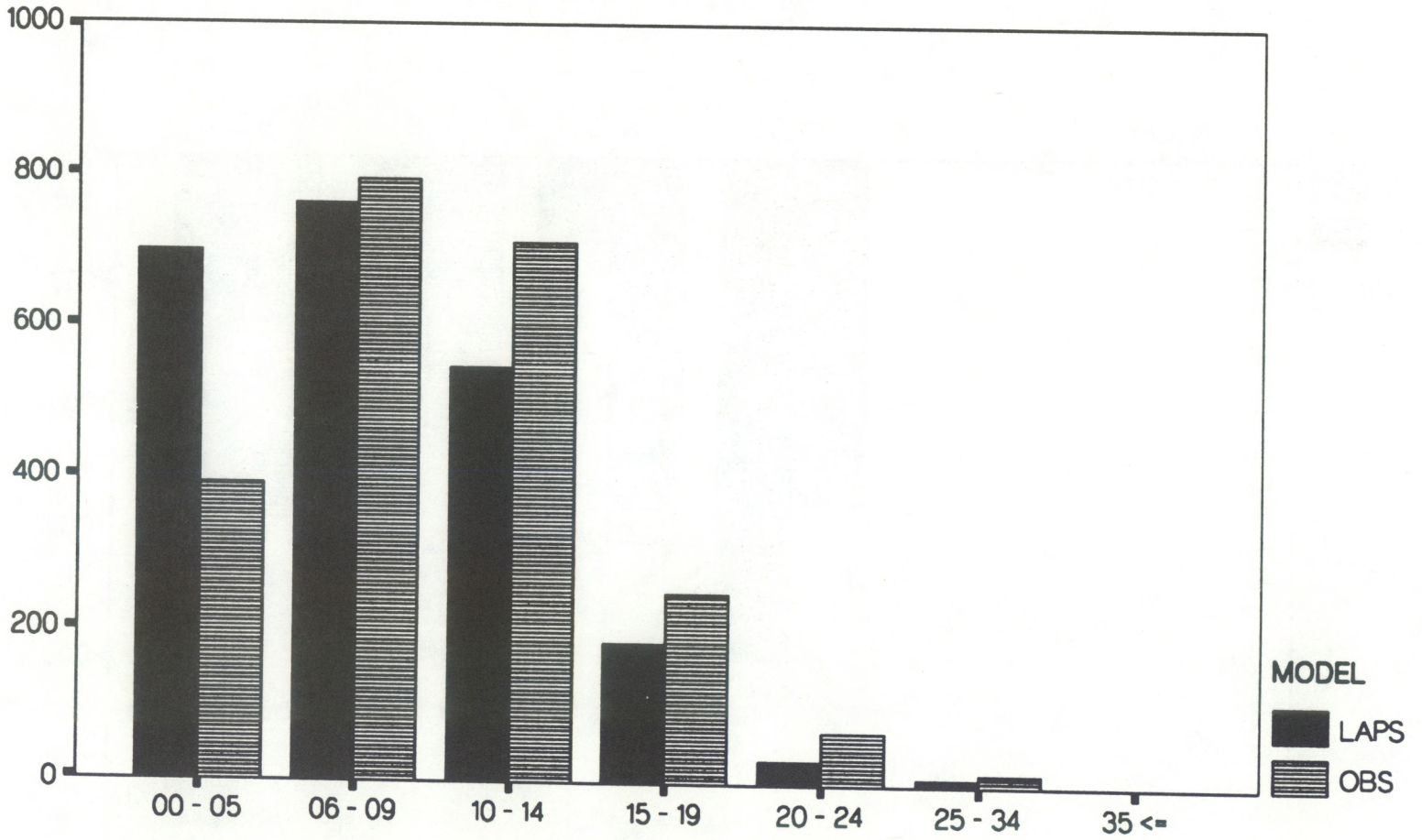

Fig. 49. Surface wind speed distribution of observed and LAPS winds in kt.

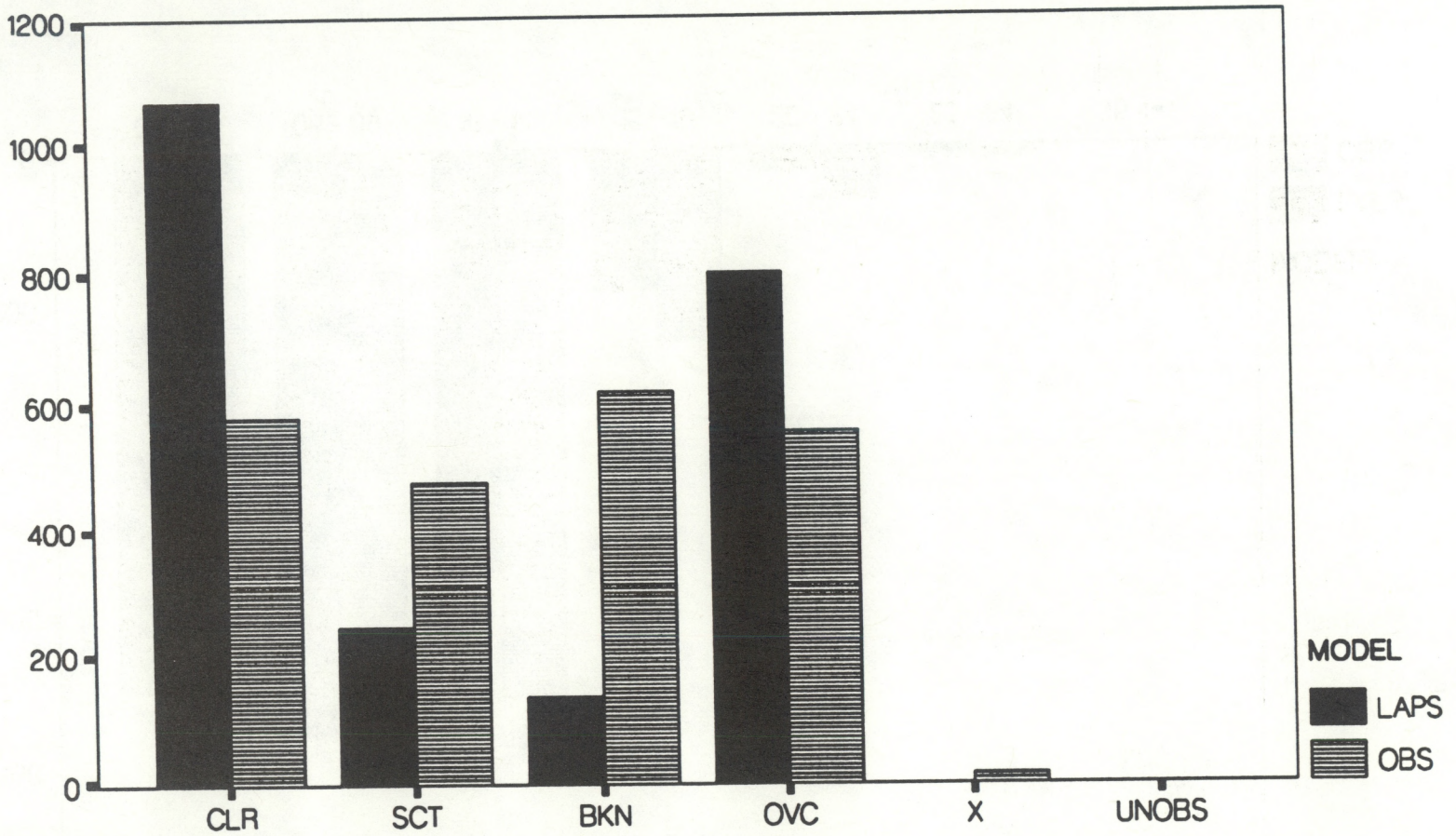

Fig. 50. Total cloud amount distribution of observed and LAPS clouds.

Table 17. Contingency Table of Analyzed Low Clouds by Observed Low Clouds for LAPS, 1-10 April 1991

	Observed Low Clouds					
Analyzed Low Clouds	CLR	SCT	BKN	OVC	X	Total
CLR	**1812**	45	7	2	0	1866
SCT	17	**107**	7	1	1	133
BKN	3	27	**29**	1	0	60
OVC	5	41	47	**76**	13	182
X	0	0	0	0	**0**	0
Total	1837	220	90	80	14	2241

However, if LAPS analyzes BKN, there is an equal chance of it being SCT. Finally, if LAPS says OVC, there is a high chance of it being BKN or even SCT. In other words, LAPS does well at detecting low clouds, but the critical discernment of cloud amount is poor.

In general, the same statistics apply to middle clouds. One notable exception is that when LAPS analyzed OVC (71 cases), there were 50% more cases that verified as BKN (109 cases) rather than OVC, and an even larger number of cases when SCT was observed (164 cases). Thus, the FAR for OVC is quite high (0.811). But overall, LAPS still performs well when detecting the presence of middle clouds (POD = 0.777, FAR = 0.130).

LAPS has difficulty detecting high clouds. The POD for high-cloud detection is only 0.360. The discrimination of cloud amount is also poor.

Cloud tops. Cloud-top errors are about 4,700 ft, and 35% of the errors are larger than 5,000 ft. From the large positive bias of +3,500 ft and the scatterplot (Fig. 51), it is evident that LAPS is often too high. The extreme cases on the scatterplot occurred at 0100 UTC 6 April, when LAPS analyzed 65,600 ft tops over DEN and APA, but only 37,000 ft tops were observed.

Ceiling. LAPS mean absolute ceiling error was 946 ft (Table 16), and there was only a small negative bias. A comparison of the errors at DEN (1,473 ft) and APA (557 ft) shows a large discrepancy for two stations in close proximity. The sample sizes are very low (11 for DEN, 23 for APA), so caution should be used when applying these statistics. In cases of low overcast, DEN tended toward higher ceilings sooner than APA. In these cases, LAPS often maintained the lower ceiling at both stations.

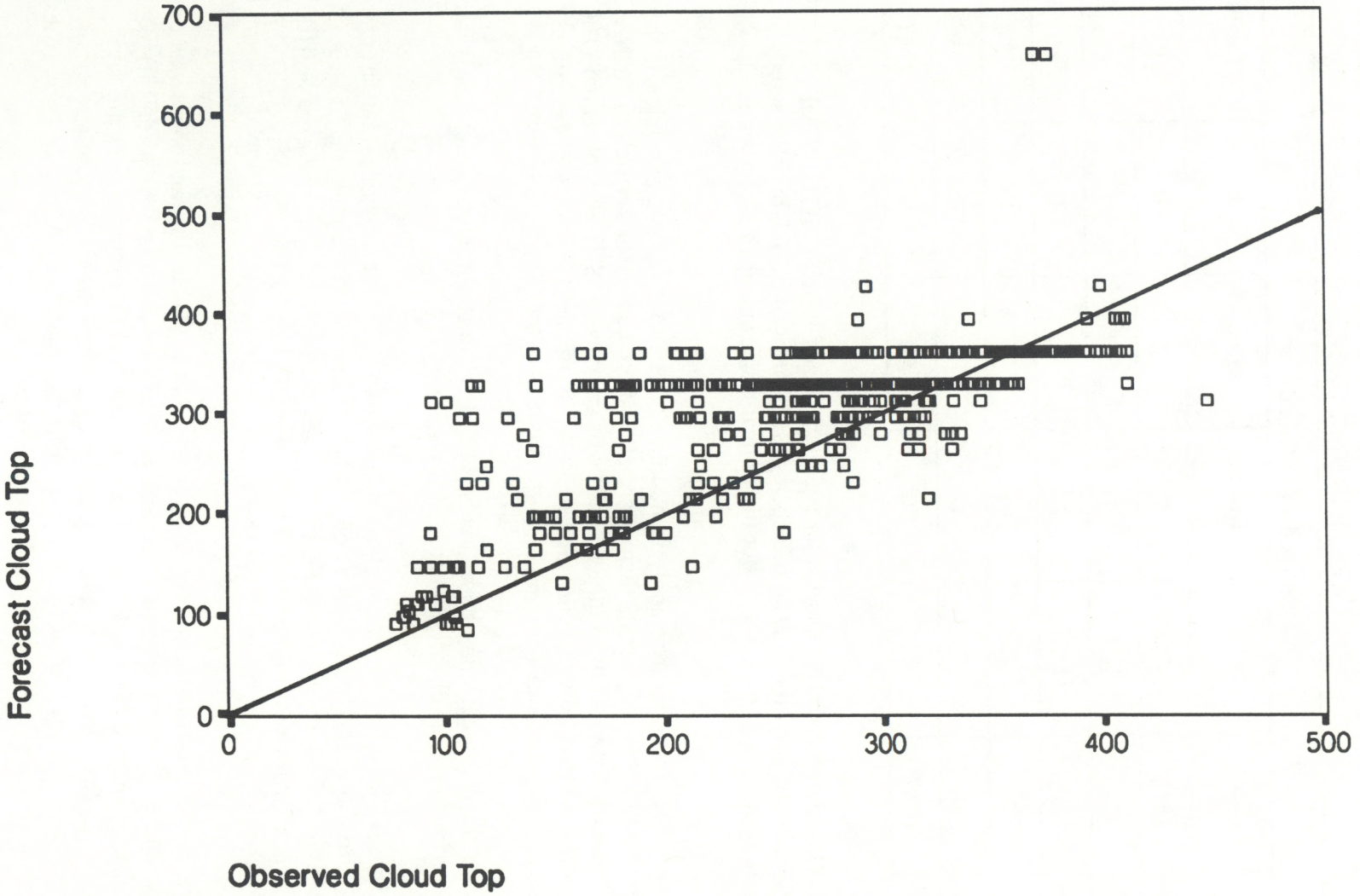

Fig. 51. Scatterplot of observed cloud-top heights versus LAPS analyzed cloud-top heights in hundreds of ft.

The distribution of ceiling categories appears to be good, because most cases fall along the diagonal in Table 18. As shown in the contingency table, all 10 times LAPS analyzed each LIFR ceiling (below 500 ft), the analysis verified. The IFR has much scatter among the four verifying categories, and about an equal chance that the actual ceiling was better than IFR, rather than IFR or worse. Only 5 of 533 occurrences that LAPS analyzed VFR conditions were they less than VFR.

Table 18. Contingency Table of Analyzed Ceilings by Observed Ceilings for LAPS, 1-10 April 1991

	Observed Ceiling				
Analyzed Ceiling	LIFR	IFR	MVFR	VFR	Total
LIFR	**10**	0	0	0	10
IFR	6	**10**	9	3	28
MVFR	0	3	**47**	30	80
VFR	1	1	3	**528**	533
Total	17	14	59	561	651

6.3.3. Discussion

The height and temperature profiles in LAPS at Denver are generally good, except for a large positive bias at 100 mb. Dewpoint temperatures were not available due to a coding error. Wind speeds are generally too weak, except above 200 mb. The wind direction is in error by 30° at 700 mb (not shown), but the error is less than 10° above that level, possibly due to terrain effects. Results of verifying against the Platteville profiler are similar.

The surface temperature, dewpoint, and wind errors are near, if not within, the instrument precision. This is to be expected because, as we pointed out earlier, the evaluation was performed using dependent data. An analysis of the errors by station reveals that LAPS seems to have more problems at some stations than others, but these problems are variable dependent. There are some interesting differences between the closely located stations of DEN and APA. For example, their closeness translates to opposite biases as LAPS tends to average these two observations.

LAPS analyzes the presence of low and middle clouds very well, but assesses the amount poorly. This is especially true when LAPS analyzes OVC. In those situations, there is often an equal or better chance that the actual sky cover will be BKN, or even SCT.

LAPS cloud tops are often too high and have a positive bias of about 3,500 ft, which may in part be due to the method of deriving the verification cloud top data. Ceiling errors are relatively small (946 ft), resulting in good discrimination of the IFR/VFR categories. As with winds, there is an interesting difference in the errors at DEN and APA, but the low sample size prevents any assumptions about the source of this difference.

6.4. RAMS

For this study, RAMS was run over four different regions (see Section 4 and Fig. 1). The Colorado and Central Plains runs differ significantly from the Florida and Northeast runs. The former were run with a coarse topography field (used by MAPS), and had a fine mesh (22 km) grid nested inside a larger coarse mesh grid. The latter runs used a high-resolution topography field and a non-nested 30 km mesh grid. For this reason, the statistics were computed separately for the two sets of runs.

The coarse (80 km) topography field used in the Colorado and Central Plains runs is highly smoothed, resulting in large differences between the model and actual station elevations, especially for the Colorado domain. This has detrimental effects on many of the surface and upper-air variables. For example, the surface temperature could contain large errors because the station altitude in the model is much higher (cold) or lower (warm) than the actual station height. Also, winds are highly influenced by terrain, and a smooth model terrain could result in large differences between the modeled and observed winds.

In addition, the number of model runs is small, resulting in a limited data sample from which to compute statistics. The stratification based on the previously mentioned differences in the model simulations further reduces the size of the data sample. Conclusions drawn from these results should be done with caution because they may not represent the true model characteristics.

Also of note is the RAMS initialization. The model was started 6 h before the initial time and nudged toward the lateral boundary conditions (MAPS analyses). Thus, the 0 h RAMS data are an initialization field, not an analysis, and thus one should not expect the field to match the observations as closely as an actual analysis. The results should be interpreted as such. All the runs were initialized at 0600 and 1800 UTC using MAPS analysis fields. Thus, any biases in the MAPS fields are probably inherited by RAMS. Also, since the MAPS fields did not contain any rawinsonde data (0600 and 1800 UTC analysis times), the analysis is largely the result of two consecutive 3-h forecasts, and some additional data from ACARS and wind profiler. Thus, even the statistical results of the MAPS upper-air analysis data (from 0000 and 1200 UTC) are likely to be better than what was used to initialize RAMS.

Appendix F lists the summary of RAMS statistics for the SAVs and AIVs at the surface and the rawinsonde levels of 850, 500, and 250 mb for the Colorado and Central Plains and the Florida and Northeast areas.

6.4.1. SAVs

Upper air

Height. RAMS appears to have a consistent positive height bias at all levels (Figs. 52 and 53). The bias increases rapidly near the top of the model. The cause of the large negative bias at 1000 mb for the Florida and Northeast domains is unknown and is probably not a function of the model, but rather of the height derivation.

The error statistics in Appendix F show that the height errors increase slightly during the 12-h forecast. The source of the anomalously large errors for the Colorado and Central Plains domains at 12 h is unknown but is probably in the derivation of the heights and possibly the coarse model terrain.

Temperature and dewpoint temperature. The data in Appendix F and Figs. 52 and 53 show little or no temperature bias, except near the surface and the top of the model. As mentioned, the topography field for the Colorado and Central Plains domains was too high in many places. This may have caused the large errors in the 850 mb temperatures for these domains. Thus, the Florida and Northeast statistics are probably more representative of the model. By 12 h, the bias profiles are largely unchanged (figures not shown), and the overall errors have actually decreased in some instances (Appendix F).

The dewpoint statistics for both regions show that the model is too dry in the lowest levels and too moist in the midtroposphere. This pattern is similar at 12 h (not shown) except in the eastern domains, where the dry bias extends up to 400 mb. This is also seen in Appendix F for the 500 mb ME. Remember that because RAMS is initialized using MAPS analyses, it inherits MAPS errors and biases.

Since the 850 mb temperature bias was too warm for the western domains, a dry bias in the dewpoint means that the model is very dry at low levels. This is well illustrated in the dewpoint depression errors. This bias is not as pronounced for the eastern domains, so again, this may be due to the coarse terrain or differences in the MAPS initial fields.

Winds. Although RAMS has a negative bias for wind speeds above 500 mb in the Colorado and Central Plains domains (Fig. 54), the bias is nearly zero throughout the entire troposphere for the Florida and Northeast domains (Fig. 55). The wind direction errors also display a similar dichotomy between the eastern and western domains.

In general, the error characteristics change very little during the 12-h forecast (Appendix F). Some levels actually show an improvement, such as the 850 mb winds for the eastern domains. The wind speed errors are all below 10 kt at 12 h. The wind direction errors increase considerably at 12 h in the western domains but are relatively unchanged in the eastern domains.

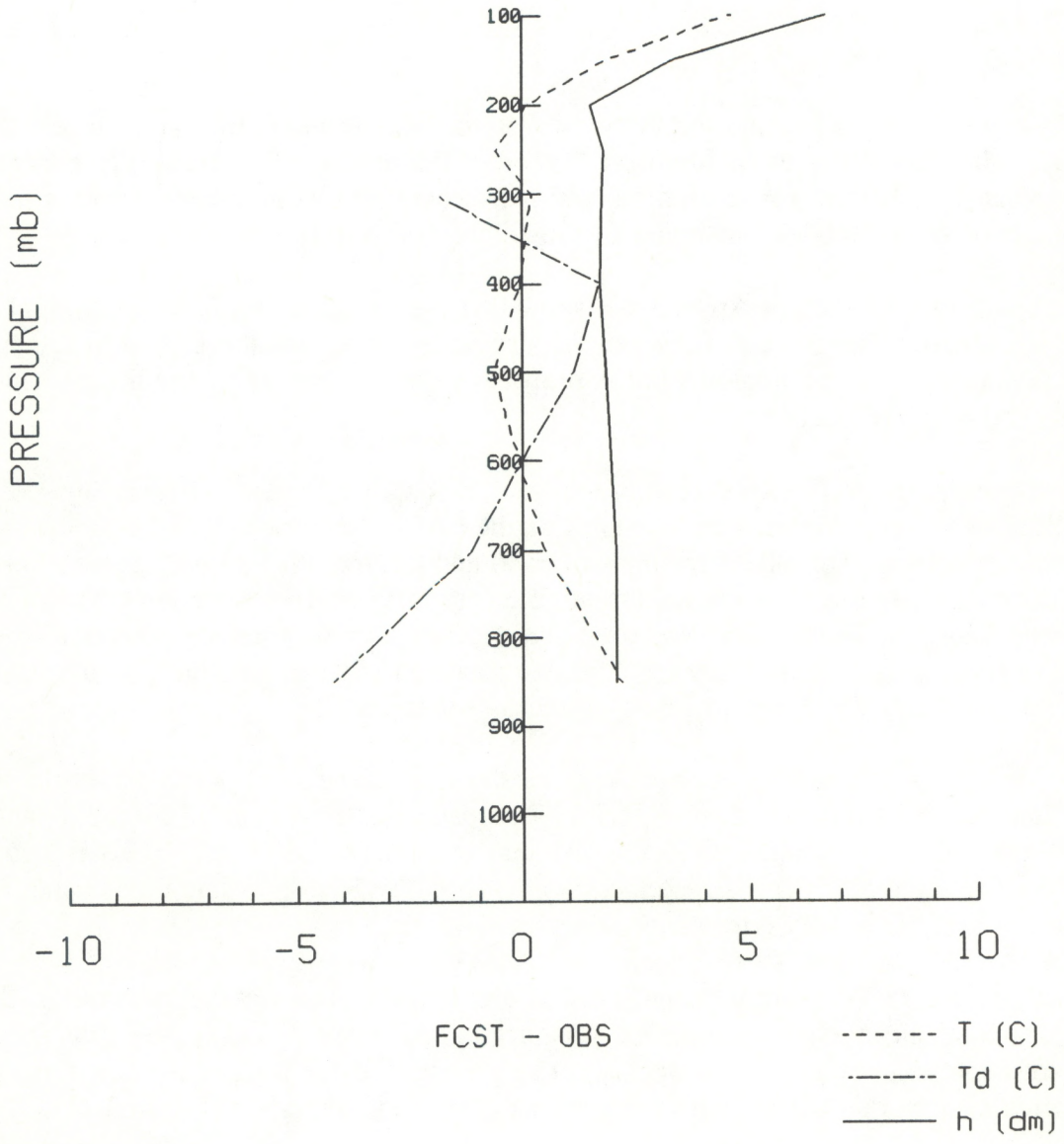

Fig. 52. Mean Error (bias) of height (h), temperature (T), and dewpoint temperature (T_d) between RAMS initial hour data and RAOBs for the Colorado and Central Plains domains.

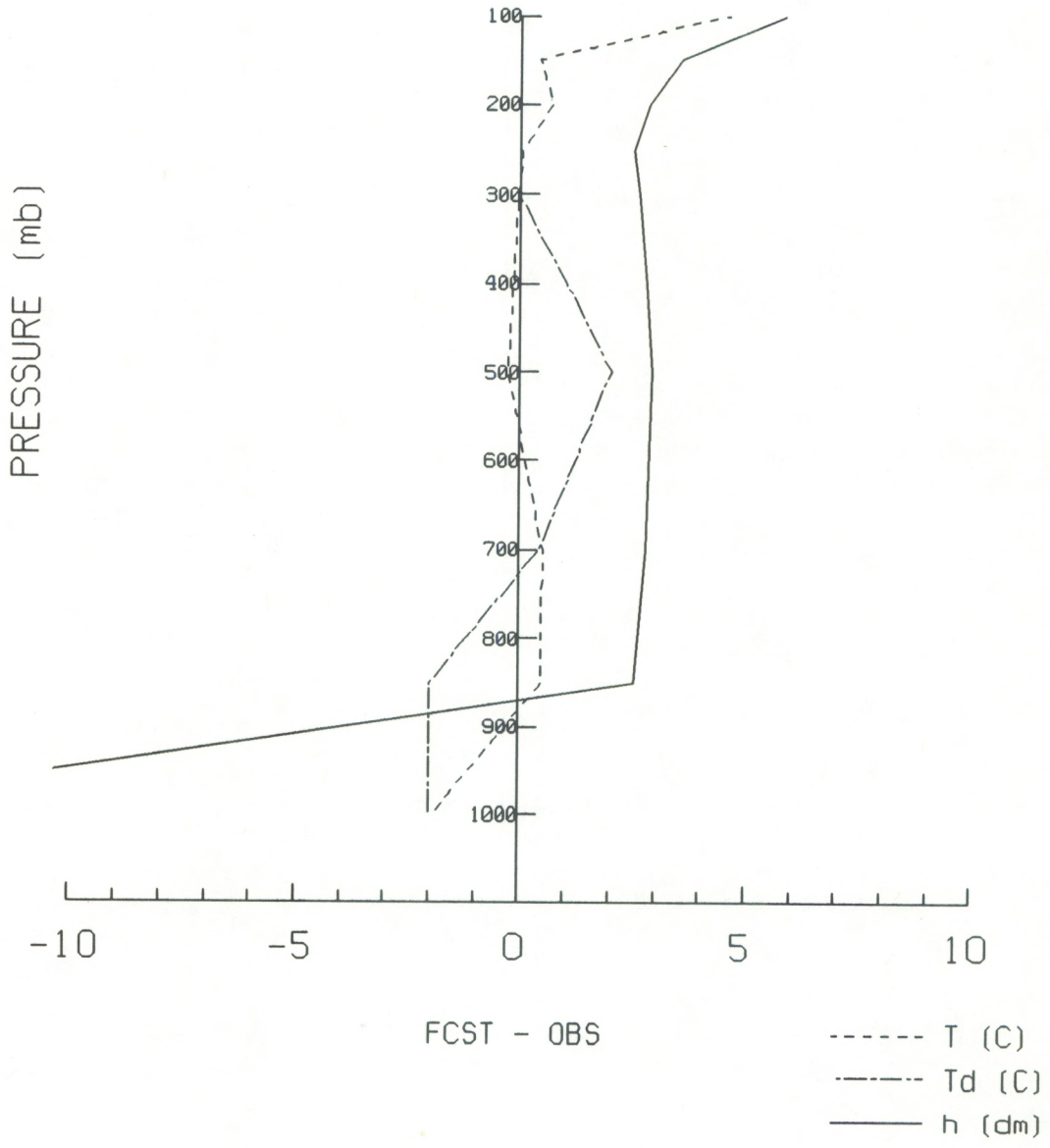

Fig. 53. Same as Fig. 52, except for the Florida and Northeast U.S. domains.

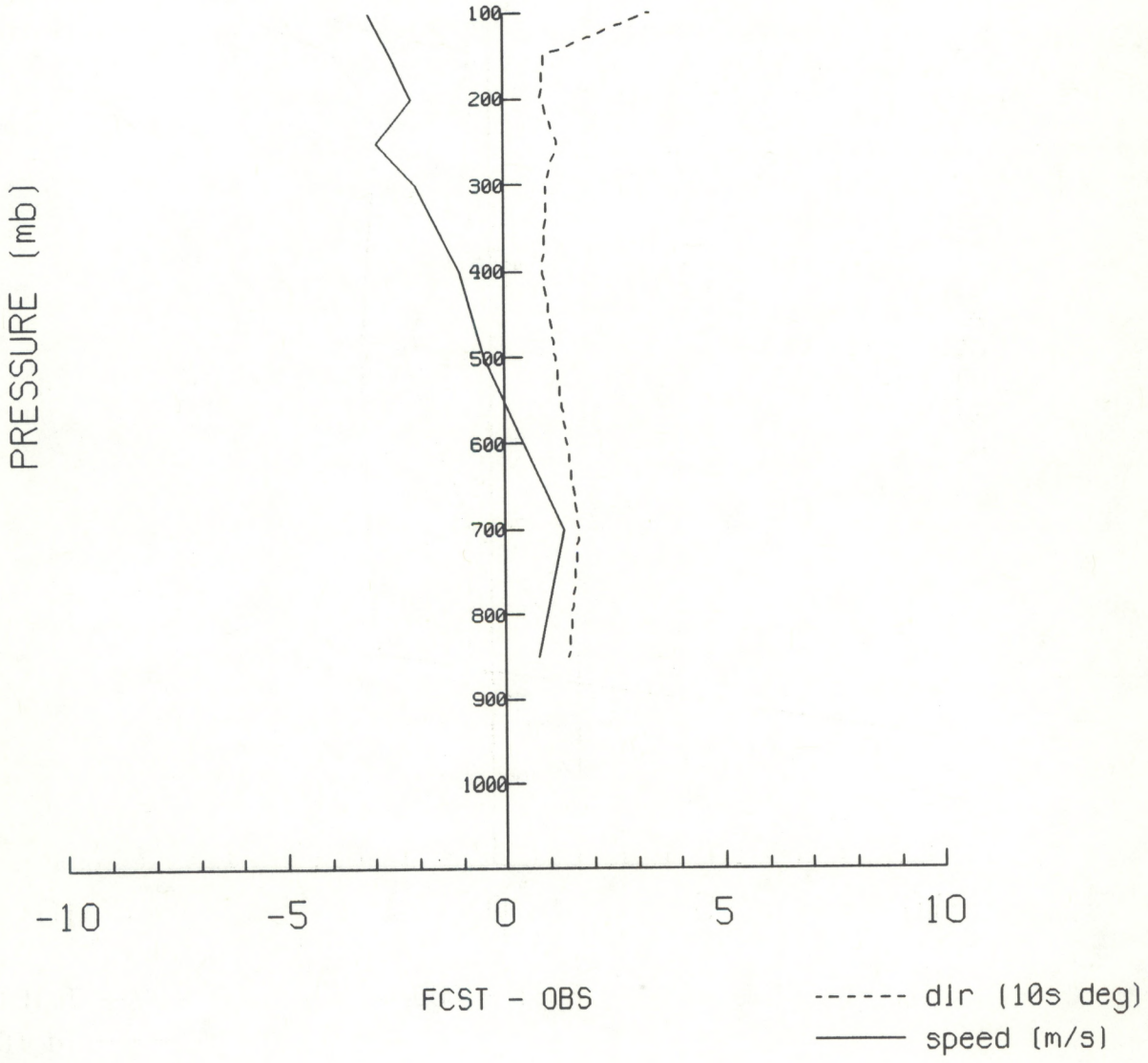

Fig. 54. Mean error (bias) of wind speed and mean absolute error of wind direction between RAMS initial hour data and RAOBs for the Colorado and Central Plains domains.

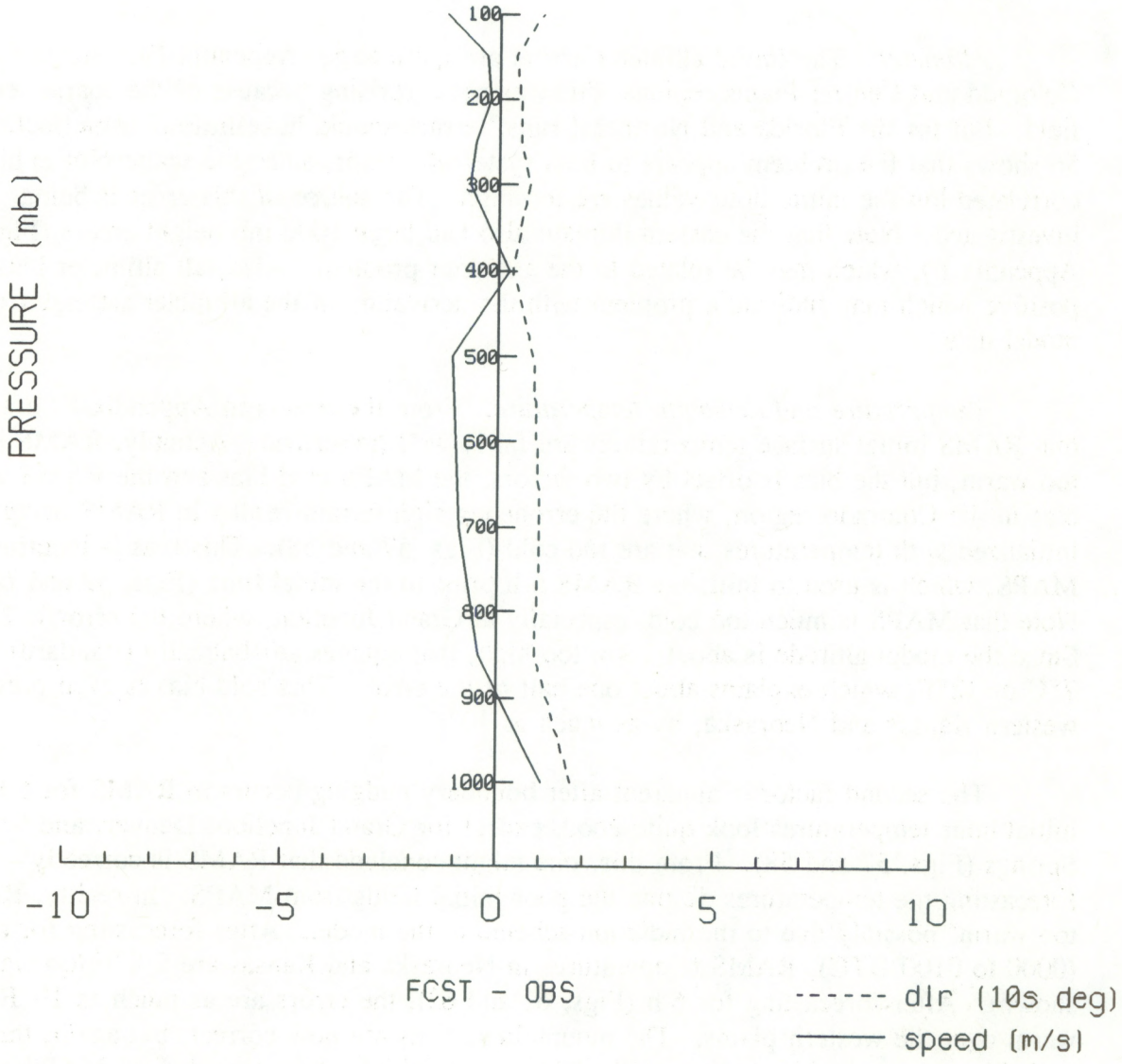

Fig. 55. Same as Fig. 54, except for the Florida and Northeast U.S. domains.

Surface

Altimeter. The initial altimeter errors are quite large (Appendix F). For the Colorado and Central Plains regions, this was not surprising because of the coarse terrain field. But for the Florida and Northeast runs, terrain should have little if any effect. Figure 56 shows that the problem appears to be a systematic error, since the scatterplot is highly correlated but the initial hour values are too high. The source of this error is being investigated. Note that the eastern domain also had large 1000 mb height errors (Fig. 53 and Appendix F), which may be related to the altimeter problem. Also, all altimeter biases are positive which may indicate a problem with the derivation of the altimeter settings from the model data.

Temperature and dewpoint temperature. From the results in Appendix F, it appears that RAMS initial surface temperatures are fairly well prescribed. Actually, RAMS is much too warm, but the bias is offset by two factors, the MAPS cold bias and the RAMS warm bias in the Colorado region, where the erroneous high terrain results in RAMS being initialized with temperatures that are too cold (Figs. 57 and 58). This bias is inherited from MAPS, which is used to initialize RAMS 6 h prior to the initial time (Figs. 59 and 60). Note that MAPS is much too cold, especially at Grand Junction, where the error is 23°F. Since the model altitude is about 1 km too high, that equates adiabatically (standard) to about 7°C or 12°F, which explains about one-half of the error. This cold bias is even present in western Kansas and Nebraska, by as much as 10°F.

The second factor is apparent after boundary nudging occurs in RAMS for 6 h. The initial hour temperatures look quite good, except for Grand Junction, Denver, and Colorado Springs (Figs. 57 and 58). From this, one might conclude that RAMS is correctly forecasting the temperatures despite the poor initial fields from MAPS. In reality, RAMS is too warm, possibly due to the radiation scheme in the model. After forecasting for only 1 h (0000 to 0100 UTC), RAMS temperatures in Nebraska and Kansas are 5-8°F too warm (not shown). After forecasting for 6 h (Figs. 61 and 62), the errors are as much as 19°F too warm over the western plains. The mountain stations are now correct, but again, this is probably more coincidence than skill. Thus, poor initialization provided by MAPS combined with the warm bias in RAMS results in a small overall bias. The problem is the same for the Central Plains domain. The data in Appendix F show that the example case presented here is typical, because the temperature bias for all the western domain runs increases from +0.9°F to +6.6°F by 12 h.

For the eastern domains, the topography was more closely described. Thus, the MAPS initial temperature field (not shown) did not suffer from the cold bias problems caused by the coarse terrain field used in the western domains. Still, RAMS was too warm by the initial time (Figs. 63 and 64). The magnitude of the errors is generally smaller than that in the western domains, but there are differences as great as 10-15°F in eastern Pennsylvania and New Jersey. By 6 h (Figs. 65 and 66), the errors are still large. Note,

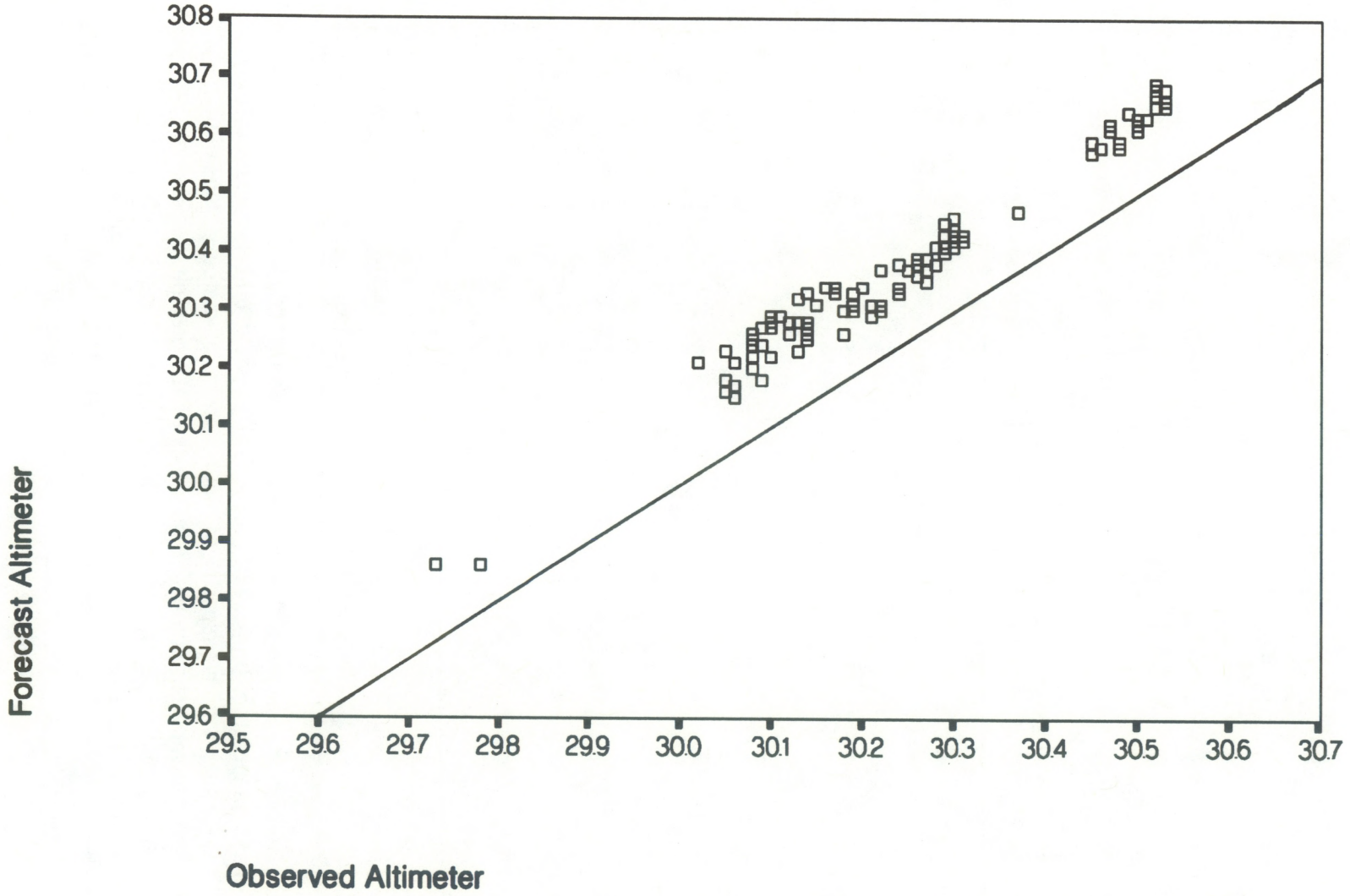

Fig. 56. Scatterplot of observed altimeter settings versus RAMS initial hour altimeter settings for the Florida and Northeast U.S. domains.

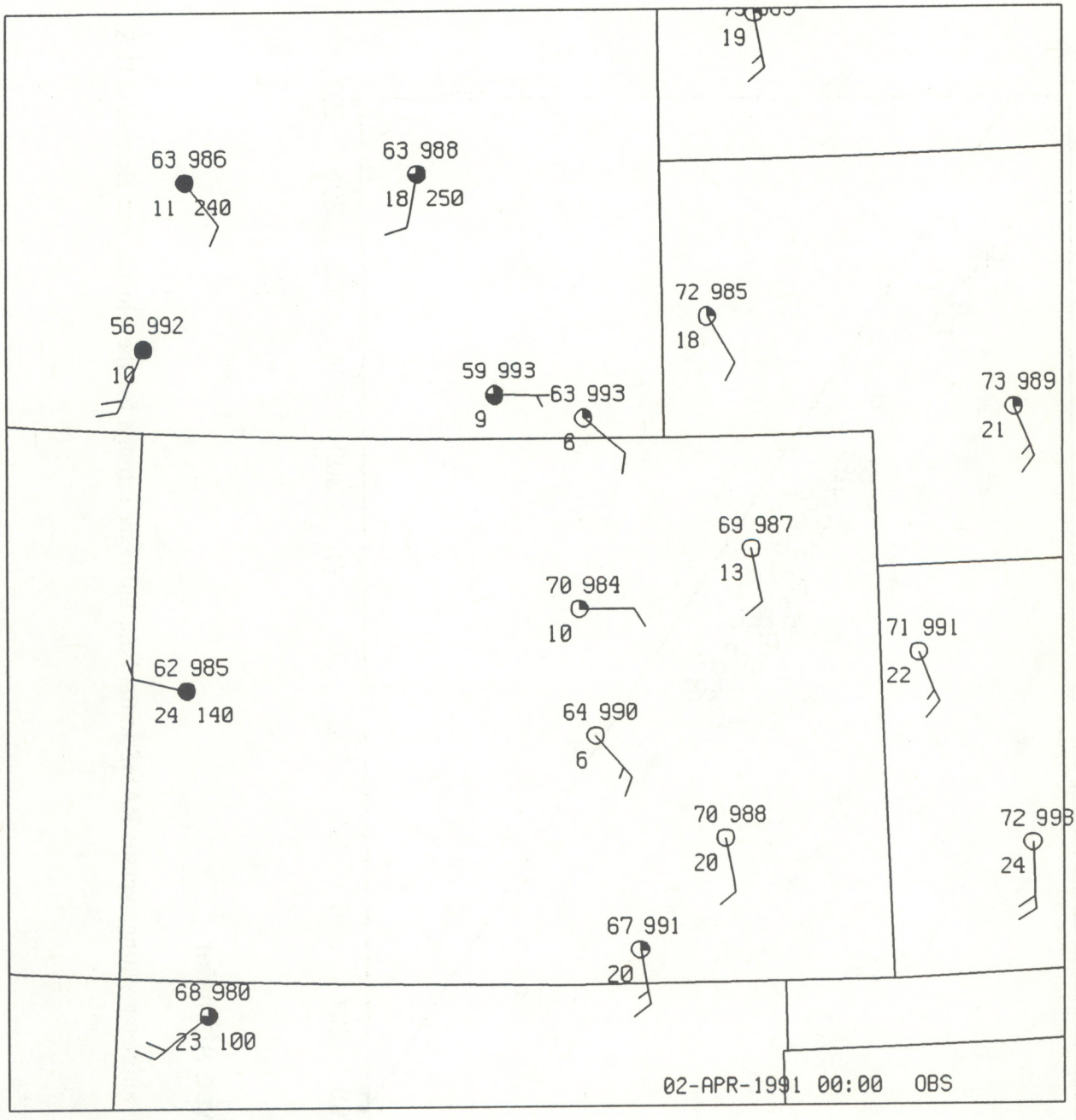

Fig. 57. Surface observations for the Colorado domain at 0000 UTC 2 April 1991. The station model displays temperature (upper left number in °F), dewpoint temperature (lower left number in °F), altimeter setting (upper right number in in. of mercury, tens digit omitted), ceiling height (lower right number in hundreds of ft), precipitation and/or obscuration type (left of cloud symbol), and hourly precipitation amount (right of cloud symbol). The wind direction and speed are shown in the traditional fashion, a full barb is 10 kt and a half barb is 5 kt. Total cloud amount is shown by CLR = open circle, SCT = one-quarter, BKN = three-quarters, OVC = black circle, X = obscuration.

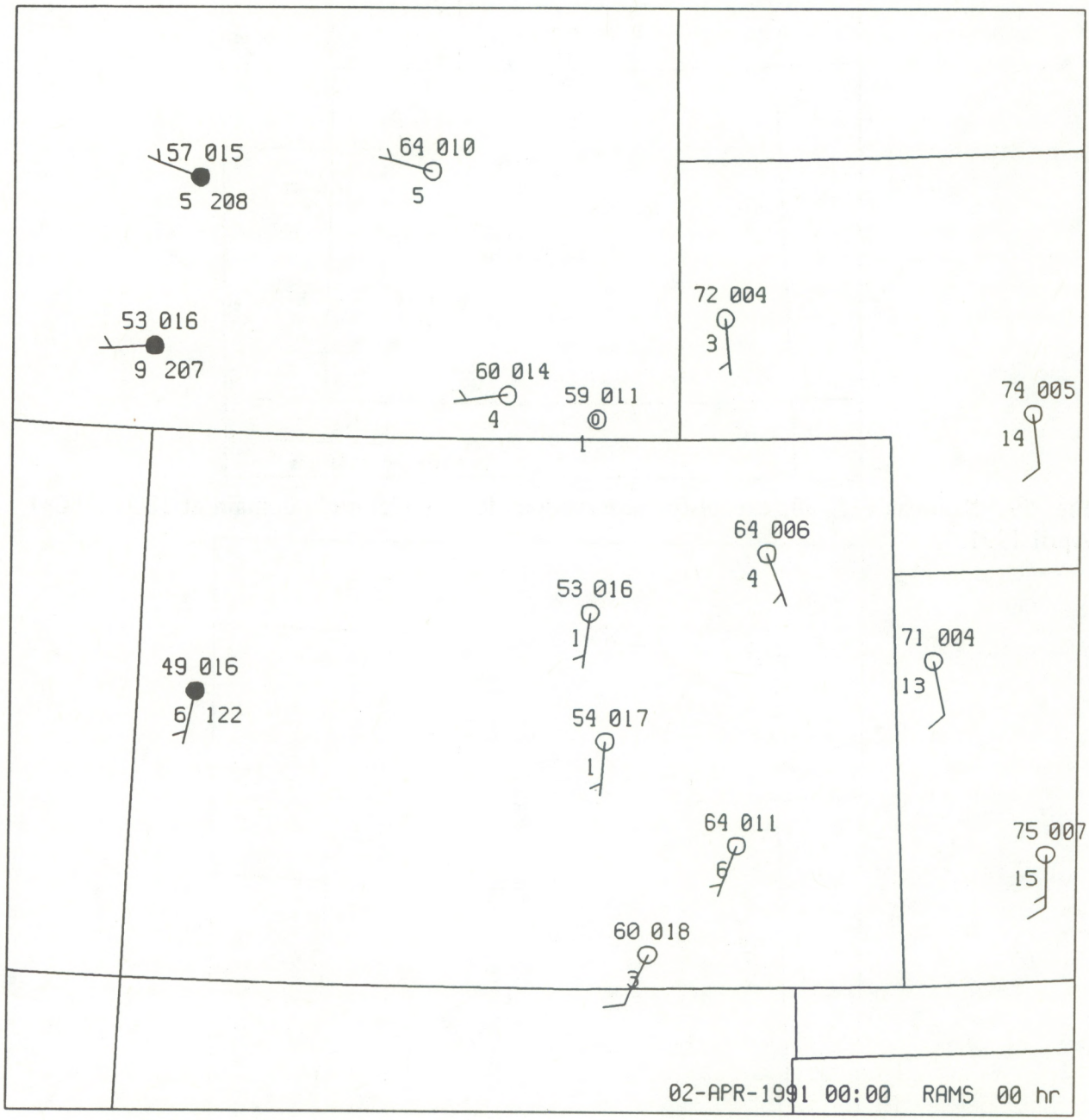

Fig. 58. Same as Fig. 57, except for RAMS initial hour. Model data are displayed in the same format as observational surface data.

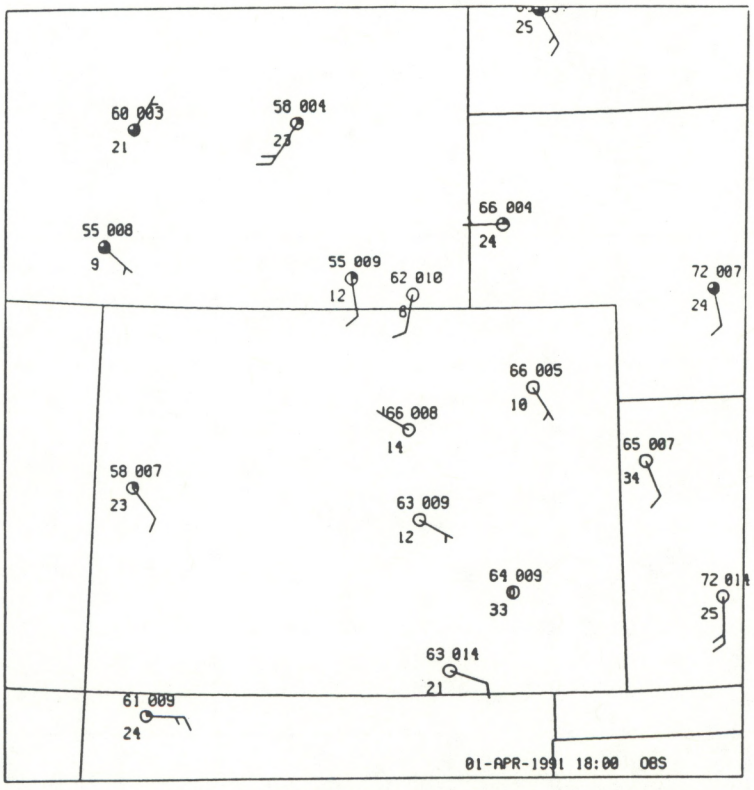

Fig. 59. Same as Fig. 58, except for observations for the Colorado domain at 1800 UTC 1 April 1991.

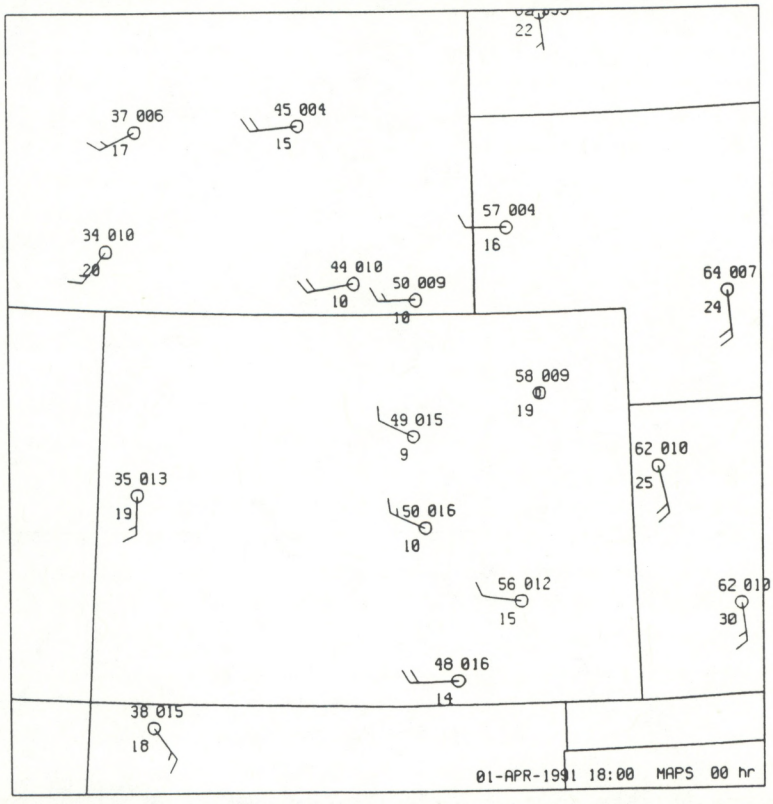

Fig. 60. Same as Fig. 59, except for MAPS analysis. Model data are displayed in the same format as observational surface data.

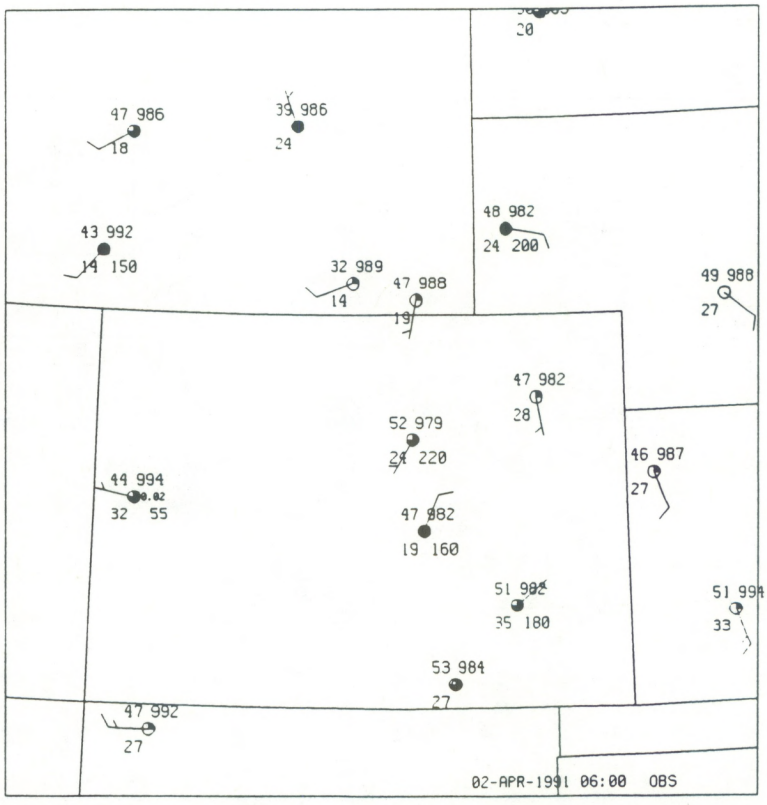

Fig. 61. Same as Fig. 58, except for observations for the Colorado domain at 0600 UTC 2 April 1991.

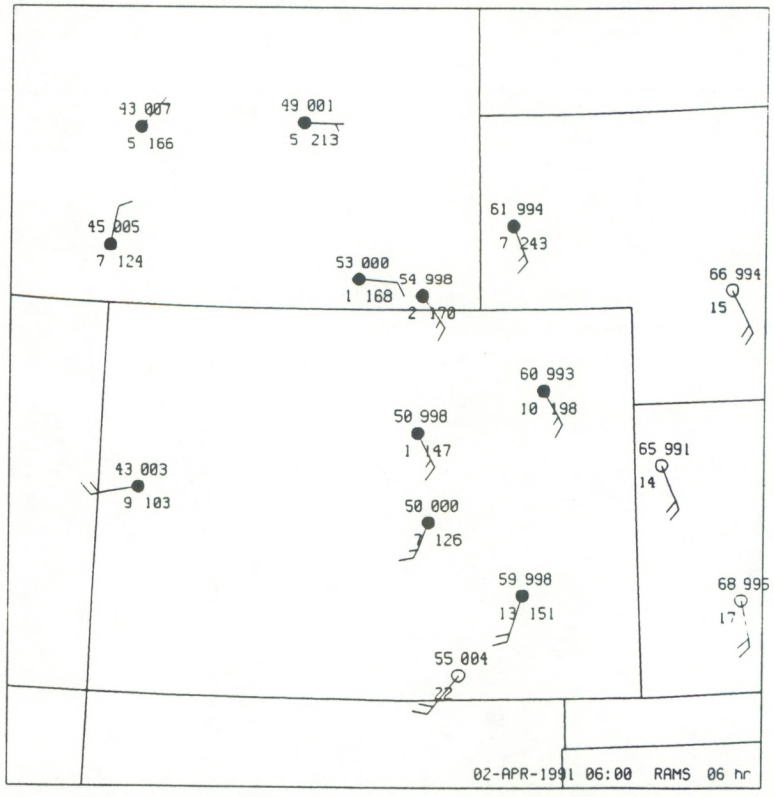

Fig. 62. Same as Fig. 61, except for RAMS 6-h forecast. Model data are displayed in the same format as observational surface data.

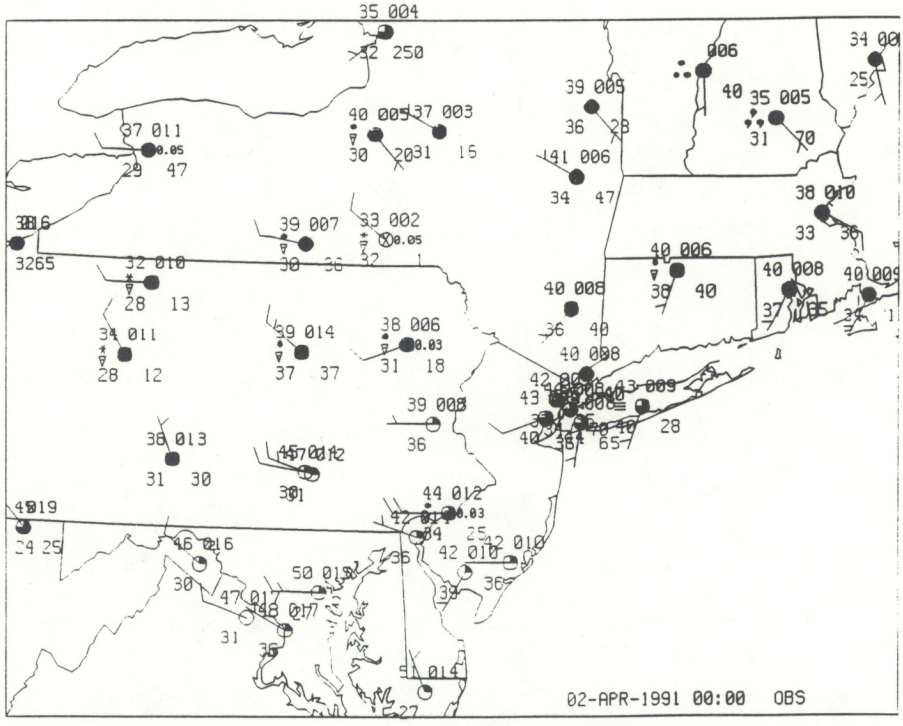

Fig. 63. Same as Fig. 58, except for observations for the Northeast U.S. domain at 0000 UTC 2 April 1991.

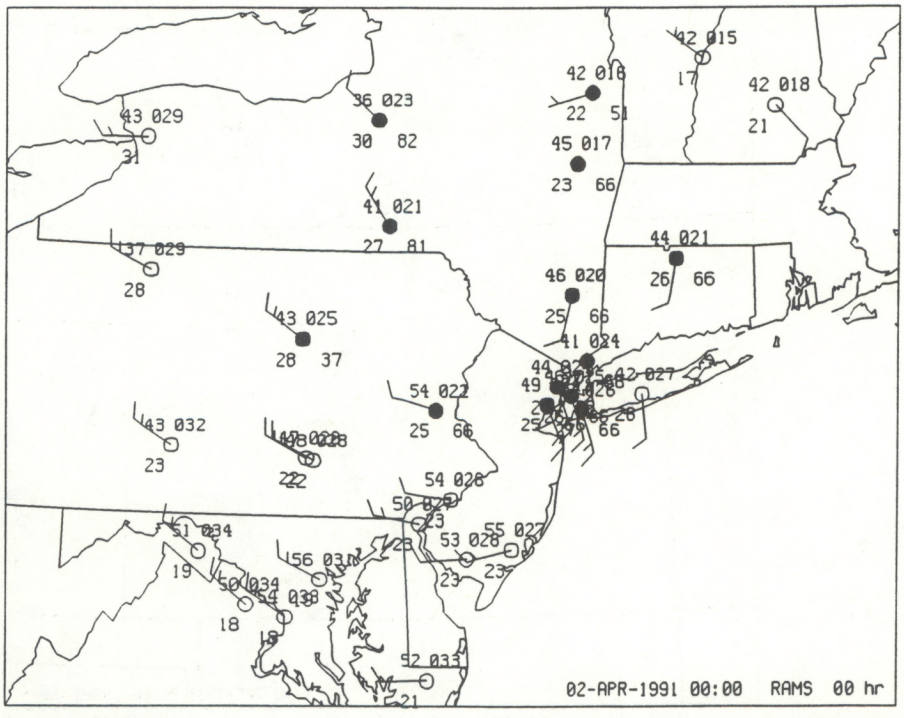

Fig. 64. Same as Fig. 63, except for RAMS initial hour. Model data are displayed in the same format as observational surface data.

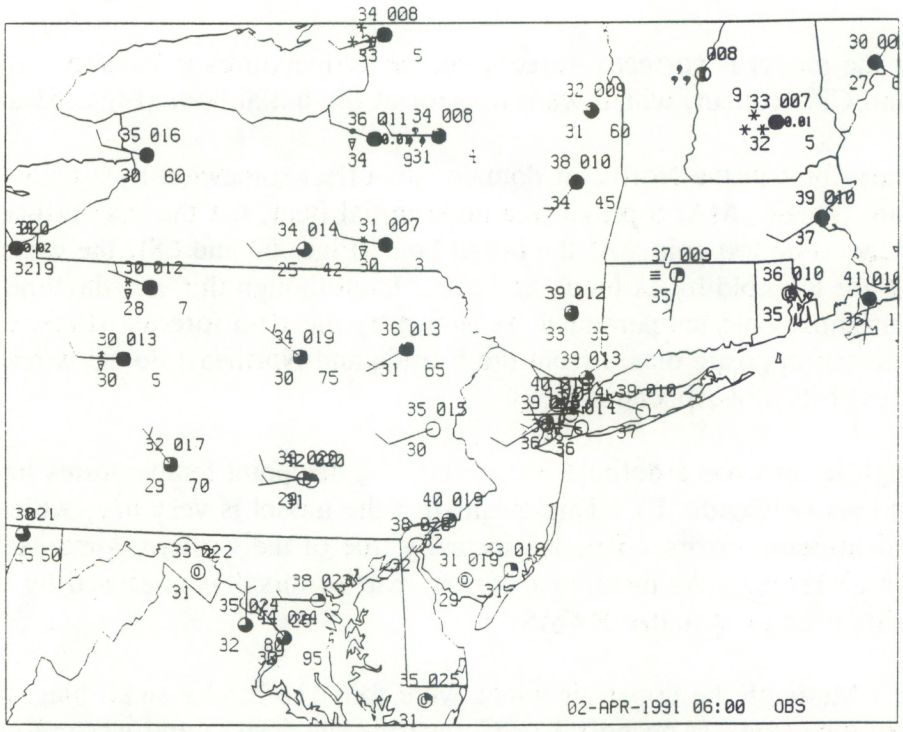

Fig. 65. Same as Fig. 58, except for observations for the Northeast U.S. domain at 0600 UTC 2 April 1991.

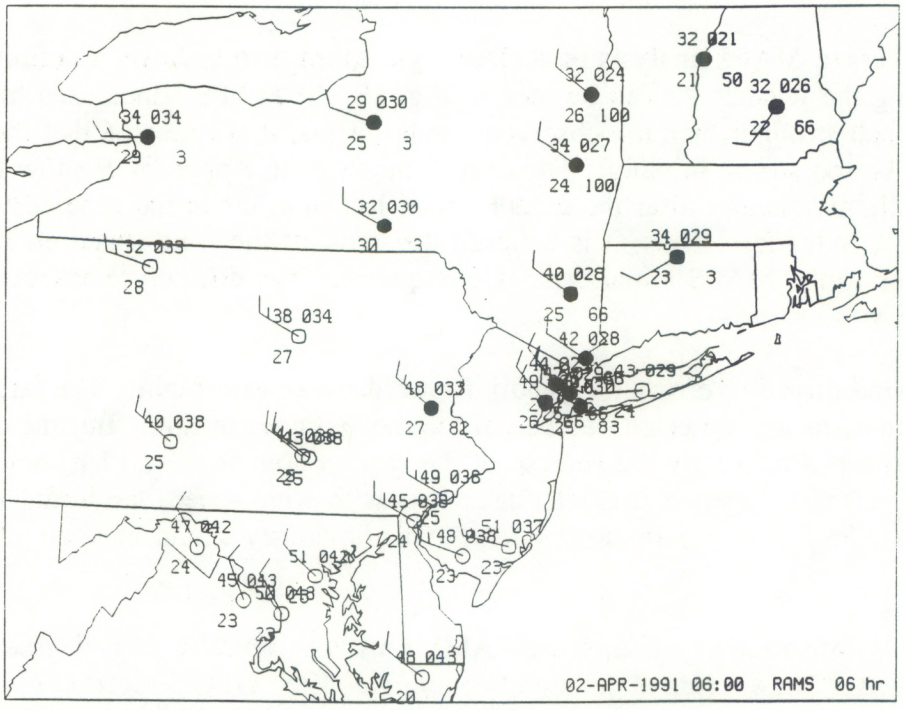

Fig. 66. Same as Fig. 65, except for RAMS 6-h forecast. Model data are displayed in the same format as observational surface data.

however, that the model is correctly forecasting the temperatures in eastern New York, New Hampshire, and Connecticut, which were in error at the initial hour (Figs. 63 and 64).

The warm bias in the Northeast domain was offset somewhat by a cold bias over the Florida domain. Again, MAPS provides a good initial field, but the sea surface temperatures in RAMS appear to be too cold. At the initial hour (Figs. 67 and 68), the eastern stations and Key West are too cold by as much as 11°F. Even though this is a daytime run, RAMS cannot bring up those cold temperatures, as shown by the 10-h forecast (Figs. 69 and 70). The combination of opposite biases from the Florida and Northeast domains results in the near-zero biases shown in Appendix F.

Although RAMS has a definite warm bias, the dewpoint temperatures have a strong negative (dry) bias (Appendix F). This means that the model is very dry, which is shown by the dewpoint depression errors. This is especially true of the western domains, which agrees with the upper-air results. As mentioned earlier, part of this bias is caused by the dry bias in the MAPS fields used to initialize RAMS.

Winds. Although the upper-air winds were found to have a small bias, the surface winds are often too strong (Appendix F). Comparing the eastern and western domains shows that the terrain differences do not appear to have much effect on the wind errors, because the numbers for the different domains are about equal. In addition, the errors are about the same at all forecast times.

The average MAPS analysis of surface wind speed errors shows a definite bias in overforecasting the winds. As can be seen in Fig. 18, the MAPS winds used to initialize RAMS are much stronger than the observed winds. Thus, it is expected that the RAMS winds would be too strong initially. The overall numbers in Appendix F show that the errors are actually slightly smaller after the model runs 6 h than at the initial time. By 12 h, the errors are slightly higher. Thus, it is believed that most of the wind speed errors in RAMS are largely due to the MAPS initial data. Consequently, it is difficult to assess the accuracy of the RAMS wind speeds.

The wind direction errors (Appendix F) are difficult to explain. The larger errors for the western domains are expected because of the mountainous terrain. But the decrease in error after a 6-h run is exactly the reverse of the eastern domain. By 12 h, both domains have again reversed their error characteristics. As with wind speed, the initial wind direction error in the MAPS fields is quite large (27°). This obviously would have an adverse effect on RAMS.

Close examination of some of the RAMS runs does reveal a high degree of accuracy in certain events. The Florida simulation beginning at 1200 UTC 5 April (Figs. 67 and 68) shows the typical easterly winds dominating the area. At 2200 UTC (Figs. 69 and 70), RAMS has correctly reversed the winds at Tampa and St. Petersburg because of a sea

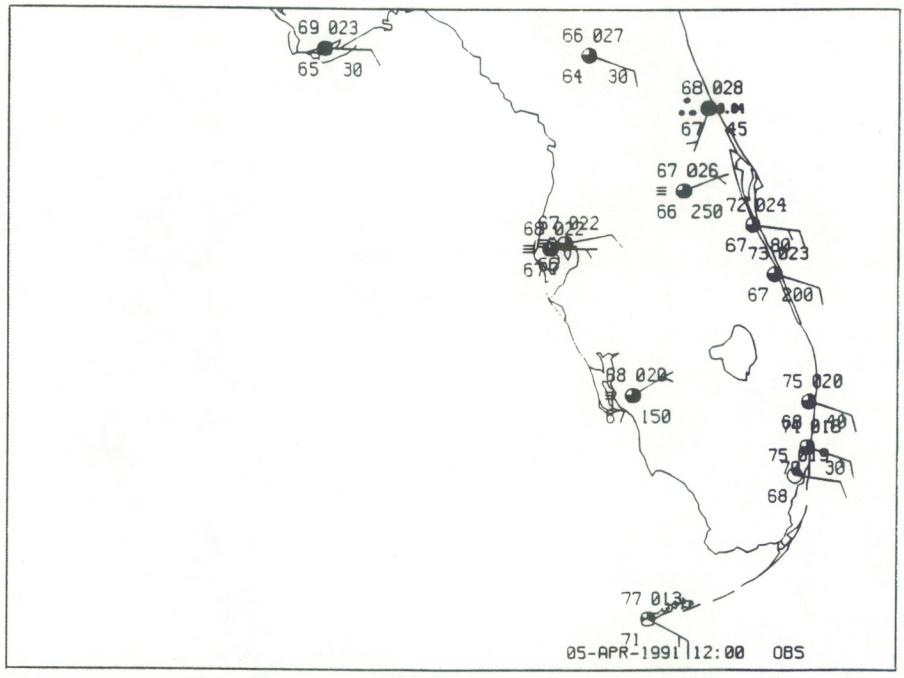

Fig. 67. Same as Fig. 58, except for observations for the Florida domain at 1200 UTC 5 April 1991.

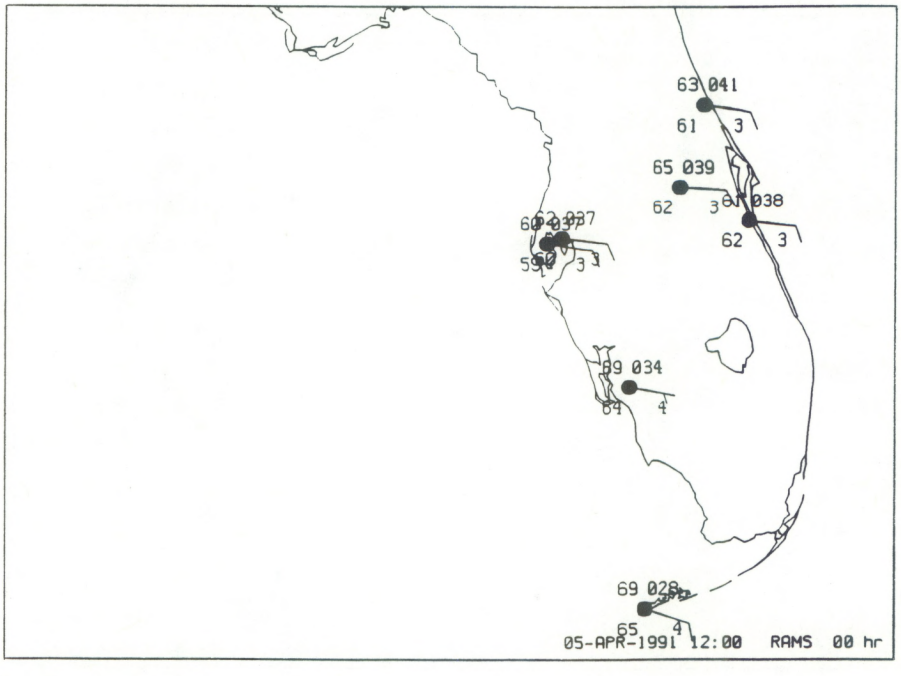

Fig. 68. Same as Fig. 67, except for RAMS initial hour. Model data are displayed in the same format as observational surface data.

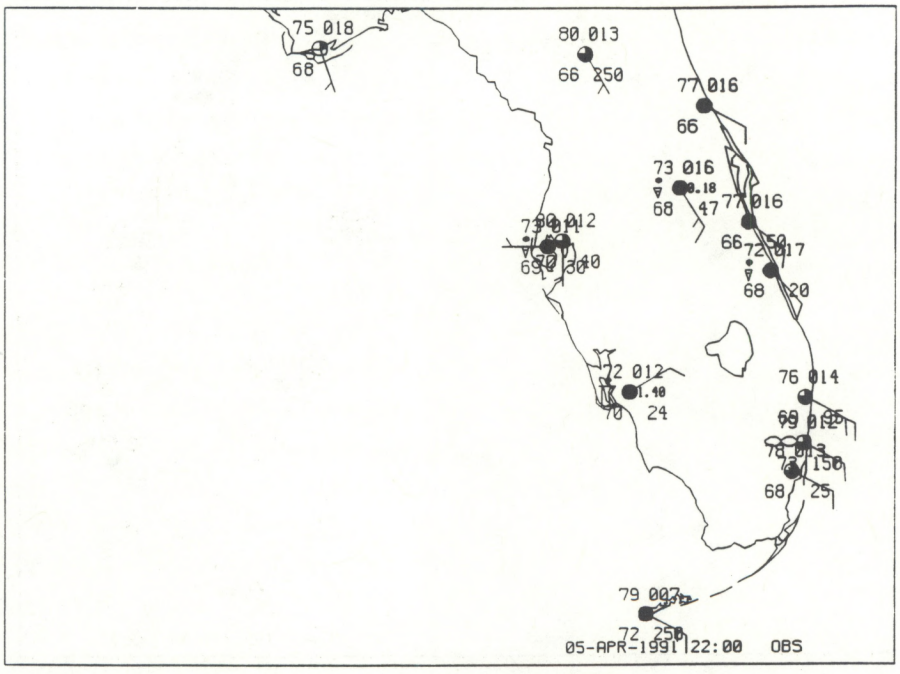

Fig. 69. Same as Fig. 58, except for observations for the Florida domain at 2200 UTC 5 April 1991.

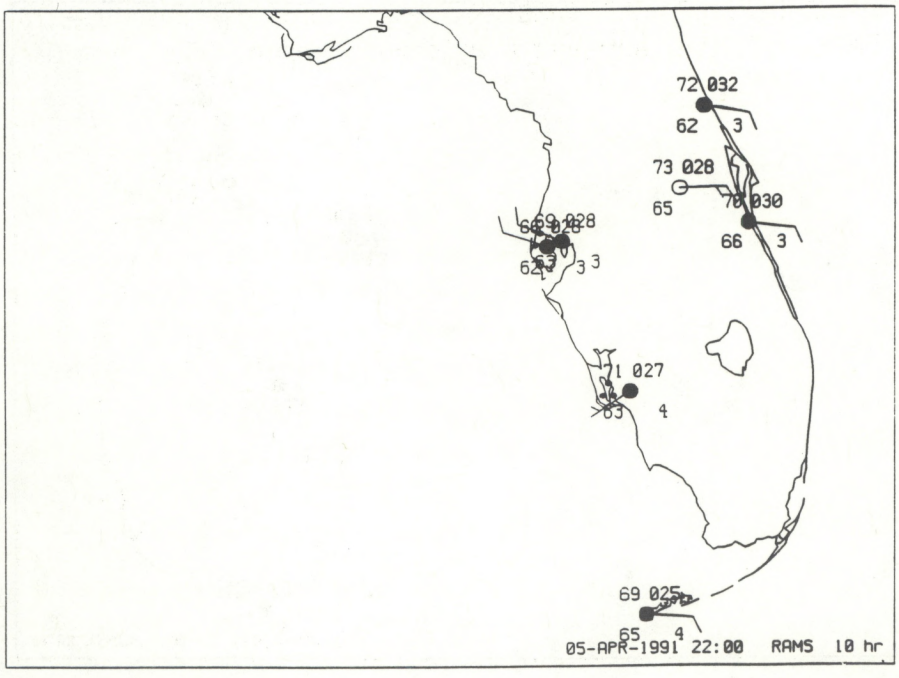

Fig. 70. Same as Fig. 69, except for RAMS 10-h forecast. Model data are displayed in the same format as observational surface data.

breeze. (The model actually generated the sea breeze at 1900 UTC, which was about 2 h too soon.) Also note that RAMS has correctly generated rain along the west coast of Florida.

In the 1200 UTC 8 April Central Plains simulation, a front slowly moves across the model domain. This front was correctly positioned in the 0600 UTC MAPS analysis (not shown) that was used to initialize RAMS. At the initial hour, RAMS has the front slightly west of its observed position (Figs. 71 and 72), as seen by the wind direction at the stations in Central Kansas and Nebraska. Note that the good temperature contrast across the front in the observations is not well defined in the RAMS data. By 2100 UTC (9-h forecast), the front has moved into Iowa but is still west of Kansas City (Fig. 73). RAMS still has the front in Nebraska (Fig. 74), yet the front is well placed in Oklahoma. The RAMS temperatures behind the front in Nebraska are much too warm. Note, however, that the dewpoint temperatures forecast by RAMS are very good, and the moist southerly flow is ahead of the front and the drier air is behind.

Problems with wind direction due to terrain are shown well in the Colorado simulation beginning at 0000 UTC 2 April. By 6 h into the forecast (Figs. 61 and 62), RAMS has a surface low in northwestern Colorado that is well defined by the wind direction. Note, however, that the Front Range and mountainous station observed winds are largely in disagreement with RAMS, presumably due to the terrain. The forecast winds over the Central Plains closely match the observations.

6.4.2. AIVs

Surface

Clouds. RAMS did not provide forecasts of cloud amount, but cloud liquid water content is explicitly forecast. If liquid water was present, clouds were assumed to exist in the model. For the western domains, the PODs at the initial time were less than 0.14 for low, middle, and high levels. The numbers were slightly better for the eastern domains but were still below 0.33. RAMS improved somewhat for the forecasts, but the PODs were never better than 0.50 for any domain or level. The FARs were all less than 0.25, which is probably due to a low bias in inferred cloudiness RAMS rather than actual skill. Based on the low POD values, it appears that the liquid water content is either poorly initialized or forecast. This would agree with the dry bias noted earlier, which is partly a result of the MAPS dry bias in the initialization fields. It may be that using just liquid water content to determine cloud presence may not be sufficient.

Some of the individual plots of observations and RAMS data provide a useful evaluation. For one of the Colorado simulations (Figs. 57-62), the model begins with cloudy skies over the mountains and clear skies to the east. By 6 h, RAMS correctly brings in clouds to the Front Range and western Plains.

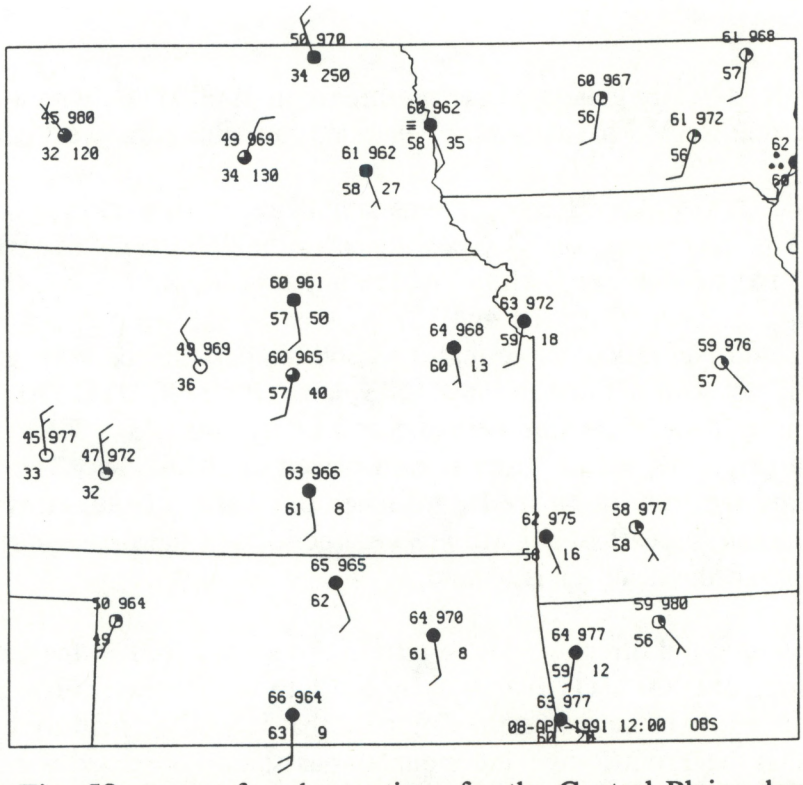

Fig. 71. Same as Fig. 58, except for observations for the Central Plains domain at 1200 UTC 8 April 1991.

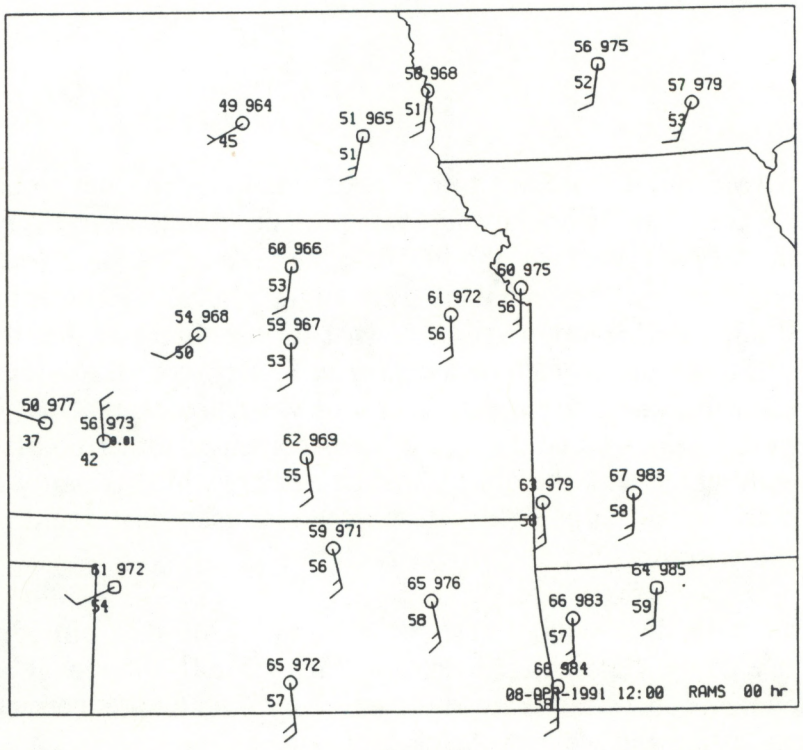

Fig. 72. Same as Fig. 71, except for RAMS initial hour. Model data are displayed in the same format as observational surface data.

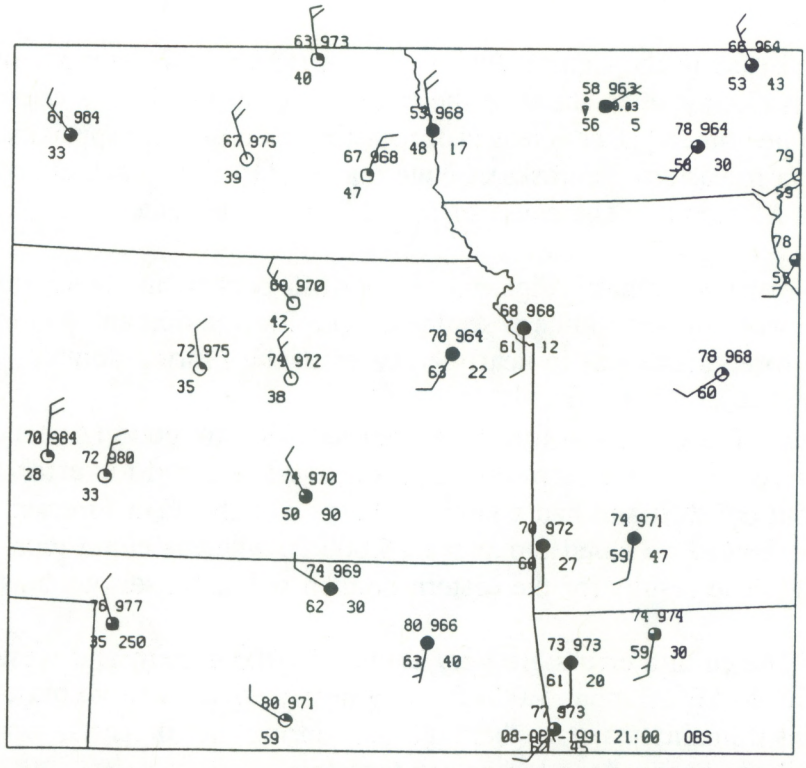

Fig. 73. Same as Fig. 58, except for observations for the Central Plains domain at 2100 UTC 8 April 1991.

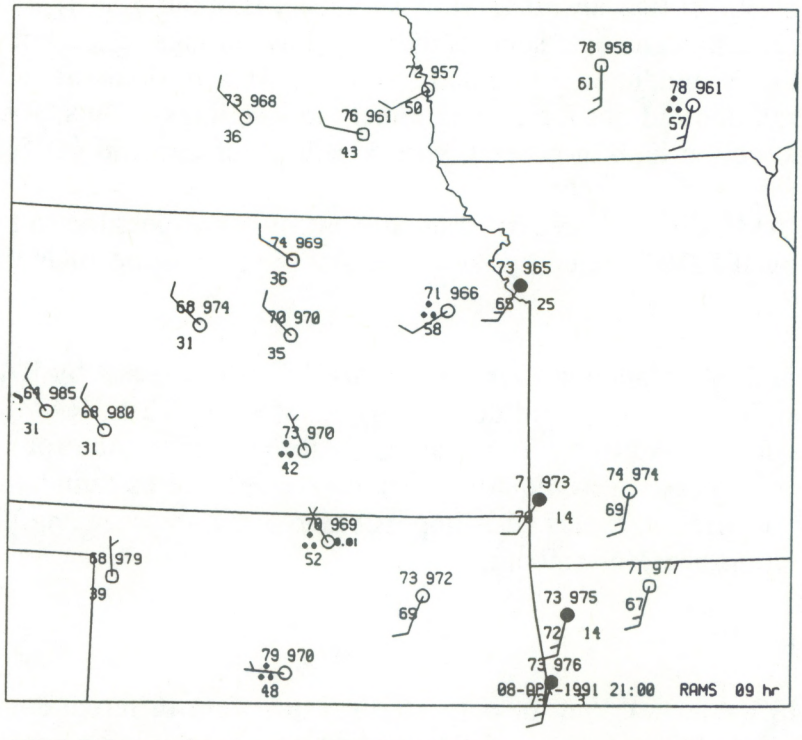

Fig. 74. Same as Fig. 73, except for RAMS 9-h forecast. Model data are displayed in the same format as observational surface data.

The frontal case in the Central Plains domain (Figs. 71-74) shows that RAMS correctly forecasts cloudy skies ahead of the front and clear conditions after the front passes. It does appear to be somewhat deficient in forecasting cloudiness, especially near the front (eastern Oklahoma to eastern Nebraska). Note that RAMS shows actual precipitation at stations that have no clouds. The cause of this is being investigated.

For the Northeast domain (Figs. 63-68), RAMS persists in forecasting too little cloudiness in the western and southern sections. The Florida domain was mostly cloudy, and RAMS correctly forecasts clouds in nearly all cases of the Florida domain.

Cloud tops. The cloud-top forecasts from RAMS look good (Appendix F). Aside from the 6-h forecasts for the eastern domains, the average cloud-top error for all four domains was about 6,500 ft and had a negative bias. For the 12-h forecast in the eastern domains, RAMS showed no cloud top above 16,000 ft, whereas cloud tops were observed as high as 43,000 ft. The results for the eastern domain suggest a serious bias problem.

Ceiling. The ceiling errors are very similar for the eastern and western domains at the initial hour: an MAE of about 5,000 ft and a nearly equal positive bias. The Colorado and Central Plains domains maintain this high bias during the forecasts, whereas the Florida and Northeast ceilings become too low in the forecasts.

As seen in the 10-h forecast in the Florida simulation (Figs. 69 and 70), RAMS predicts ceilings of 300 ft (the lowest level in the model is about 500 ft) for nearly all stations. In reality, although some stations did have low ceilings (e.g., Fort Myers), many other stations had only middle or high clouds. For the western domains, RAMS forecast fewer clouds overall and did not forecast extremely low ceilings. Thus, RAMS analyses and forecasts tended to be too high in general when a ceiling forecast did verify.

Because RAMS did not forecast cloud amount, it is not possible to evaluate the ceiling forecasts by IFR/VFR category, since the sky cover must be BKN or OVC to be considered a ceiling.

Precipitation. Precipitation events were largely underforecast by RAMS. Of the 24 events at the initial hour, only 4 were detected by RAMS. By 12 h, RAMS detected only 6 of 26 events. As seen in Appendix F, the amounts are generally underforecast. All but one of the precipitation forecasts were liquid, which always verified as liquid. Because of the very low number of forecast events, it is impossible to draw any meaningful conclusions about the precipitation phase or amount.

6.4.3. Discussion

The RAMS model was run for only 14 cases over four different domains. This small sample size makes it difficult to make detailed conclusions about the model performance.

Overall, RAMS appears to do well at upper-air temperature forecasts, but shows a dry bias at lower levels, especially over the Colorado and Central Plains domains.

At the surface, RAMS appears much too warm. This warm bias is possibly attributed to the radiation package, both in heating too quickly and cooling too slowly. Some of the bias is offset by temperatures being too cold in the high mountain stations and the coastal stations of Florida. The former was partially a result of the MAPS initialization temperature bias, and it is suspected that the latter results from sea surface temperatures that were specified too low. The wind speeds are generally too strong, which could be due in part to a bias inherited from the MAPS model. The lack of reduction of SAVs from the lowest model level to station elevation could also be a large contributing factor.

RAMS appears to do well at cloud height (tops and ceilings), but did poorly at detecting the presence of clouds. Actual cloud amount was not predicted by RAMS. In general, RAMS underforecast the areal coverage and amount of precipitation. However, the low number of precipitation events amongst the cases run precludes a detailed analysis of the actual skill of the model.

7. CONCLUDING REMARKS

An evaluation of aviation-impact variables (AIVs) derived from four numerical analysis and forecast systems was conducted over a 10-day period from 1-10 April 1991. This was the first time many of the systems had attempted to derive AIVs from state-of-the-atmosphere variables (SAVs). We found that all systems generated fairly good analyses and forecasts of SAVs. Several systems produced surface winds that were too strong and winds directions that were often 30° off. Several systems had difficulty with surface temperatures being either too high or too low. All models were too dry. More problems were found with the AIV analyses and forecasts. The best result of the evaluation was the systems' ability to produce relatively good detection of cloud presence. The difficulty arose in distinguishing between categories such as scattered or broken cloud decks. Of those systems that produced precipitation-related forecasts, precipitation events were almost always overforecast and precipitation amounts were largely under-forecast.

Many lessons were learned from this evaluation. We believe that the evaluation statistics of the SAVs were greatly affected by model topography, radiation parameterization, and moisture initialization. Improvements in these areas are expected to lead to improvements in the analyses and forecasts. Also, it is not known at this time if some of the poor results of AIV forecasts are a result of the derivation from the SAVs, or of the algorithms that produce the AIVs. In evaluating the algorithms, many errors were most likely a result of not correctly extrapolating the lowest model level data to the station elevation. Although this is a difficult task, we feel that this is an important step to take to correctly derive both SAVs and AIVs at the surface. More work is needed in these areas.

Finally, it should be noted that the MAPS and LAPS verification data used for this evaluation were from the real-time model and analysis runs conducted in April 1991. These two systems are still under development, and many of the problems associated with the results in this study have been investigated, and in many situations corrected or improved. RAMS and Eta were run specifically on the data set with the most current version of the model, but as with MAPS and LAPS, RAMS and Eta are continually under development.

8. FUTURE PLANS

Overall, we are encouraged by the results of this first baseline evaluation of model-derived AIVs. We plan to conduct additional exercises in the near future to assist the models in the development of the AGFS gridded database. The next exercise (Exercise 2) will include the unique verification data sets collected during the STorm-scale Operational and Research Meteorology--Frontal Experiment and Systems Test (STORM-FEST) (conducted from February through mid-March 1992). Also, we wish to include other meteorological groups, such as the National Center for Atmospheric Research Mesoscale and Microscale Meteorology Division and the United Kingdom Meteorological Office, to assist in improving the ability to make aviation forecasts.

9. ACKNOWLEDGMENTS

Many people have contributed to the design and implementation of the Verification Program, and in particular many scientists in the Forecast Research Division, FSL (see Appendix G for detailed information). Also, the authors wish to thank Glen Pankow, FSL Facilities Division, for his database programming support, and the FSL Facilities Division for providing the mass of verification data for the April cases. In addition, special thanks go to Robert Glahn, NWS, and John Flueck, University of Nevada-Las Vegas, for their comments on the report, and to Cecilia Girz, FSL, for her careful review of this report.

10. REFERENCES

Albers, S.C., 1992: The LAPS wind analysis. *Preprints, Fourth AES/CMOS Workshop: Forecasting in the Nineties,* September 15-18, Whistler, British Columbia, Canada, 186-195.

Benjamin, S.G., T.L. Smith, P.A. Miller, D. Kim, T.W. Schlatter, and R. Bleck, 1991a: Recent improvements in the MAPS isentropic-sigma data assimilation system. *Ninth Conf. Numerical Wea. Prediction,* October 14-18, Denver, CO, Amer. Meteor. Soc., Boston, MA, 118-121.

Benjamin, S.G., K.A. Brewster, R. Brummer, B.F. Jewett, T.W. Schlatter, T.L. Smith and P.A. Stamus, 1991b: An isentropic three-hourly data assimilation system using ACARS aircraft observations. *Mon. Wea. Rev.*, **119,** 888-906.

Betts, A. K., and M. J. Miller, 1986: A new convective scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112,** 693-709.

Brewster, K.A., 1989: Quality control of wind profiler data. Profiler Training Manual No. 2, developed for the National Weather Service Office of Meteorology, Program for Regional Observing and Forecasting Services, NOAA/ERL/FSL, Boulder, CO, 39 pp.

Brewster, K.A. and T.W. Schlatter, 1986: Automated quality control of wind profiler data. *Preprints, 11th Conf. on Weather Forecasting and Analysis,* June 17-20, Kansas City, MO, Amer. Meteor. Soc., Boston, MA, 171-176.

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*, June 22-26, Toronto, Ontario, Canada, Amer. Meteor. Soc., Boston, MA, 150-152.

Chen, C., and W.R. Cotton, 1983: Numerical experiments with a one-dimensional higher order turbulence model: Simulation of the Wangara Day 33 Case. *Bound.-Layer Meteor.*, **25,** 375-404.

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

Davies, H.C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.*, **102,** 405-418.

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.

Flatau, P.J., G.J. Tripoli, J. Verlinde, and W.R. Cotton, 1989: The CSU-RAMS cloud microphysics module: General theory and code documentation. Paper no. 451, Dept. of Atmos. Sci., Colorado State University, Fort Collins, CO 80523, 88 pp.

Flueck, J.A. and T.J. Brown, 1993: Criteria and methods for performing and evaluating solar-weather studies. *J. Climate,* **6,** 373-385.

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.

Hogg, D.C., M. T. Decker, F.O. Guiraud, K. B. Earnshaw, D.A. Merritt, K.P. Moran, W.B. Sweezy, R.G. Straugh, E.R. Westwater and C.G. Little, 1983: An automatic profiler of the temperature, wind and humidity in the troposphere. *J. Climate Appl. Meteor.*, **22,** 807-831.

Janjić, Z. I., 1990: The step-mountain coordinate: Physical package. *Mon. Wea. Rev.*, 118, 1429-1443.

McGinley, J.A., 1982: A diagnosis of alpine lee cyclogenesis. *Mon. Wea. Rev.*, **110,** 1271-1287.

McGinley, J.A., 1989: The Local Analysis and Prediction System. *Preprints, Twelfth Conf. on Weather Analysis and Forecasting*, October 2-6, Monterey, CA, Amer. Meteor. Soc., Boston, MA, 15-19.

Mesinger, F., Z.I. Janjić, S. Ničković, D. Gavrilov, and D.G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of alpine lee cyclogenesis and for a case of an Appalachian redevelopment. *Mon. Wea. Rev.*, **116,** 1493-1518.

Miller, P.A. and S.G. Benjamin, 1991: Horizontal quality control for a real-time 3-h assimilation system configured in isentropic coordinates. *Preprints, Ninth Conf. on Numerical Weather Prediction*, October 14-18, Denver, CO, Amer. Meteor. Soc., Boston, MA, 32-35.

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.

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, Sliver 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.

Schwartz, B.E., and C.A. Doswell III, 1991: North American rawinsonde observations: Problems, concerns, and a call to action. *Bull. Amer. Meteor. Soc.*, **72**, 1885-1896.

Sherretz, L., 1991: Developing the Aviation Gridded Forecast System. *Preprints, Fourth Int. Conf. on Aviation Weather Systems*, June 24-28, Paris, France, Amer. Meteor. Soc., Boston, MA, 102-105.

Smithsonian Institution, 1966: *Smithsonian Meteorological Tables*, Smithsonian, Washington, DC, 527 pp.

Stamus, P.A., 1992: A high-resolution analysis system and its use in winter weather situations. *Preprints, Fourth AES/CMOS Workshop: Forecasting in the Nineties*, September 15-18, Whistler, British Columbia, Canada, 176-185.

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.

Tremback, C.J. and R. Kessler, 1985: A surface temperature and moisture parameterization for use in mesoscale numerical models. *Preprints, 7th Conf. on Numerical Weather Prediction*, 17-20 June, Montreal, Canada, Amer. Meteor. Soc., Boston, MA, 355-358.

Tripoli, G.J., and W.R. Cotton, 1982: The Colorado State University three-dimensional cloud/mesoscale model - 1982. Part I: General theoretical framework and sensitivity experiments. *J. Rech. Atmos.*, **16**, 185-219.

Tripoli, G.J., and W.R. Cotton, 1989: A numerical study of an observed orogenic mesoscale convective system. Part 1. Simulated genesis and comparison with observations. *Mon. Wea. Rev.*, **117**, 273-304.

Weber, B.L., D.B. Wuertz, R.G. Strauch, D.A. Merritt, K.P. Moran, D.C. Law, D. Van de Kamp, R.B. Chadwick, M.H. Ackley, M.F. Barth, N.L. Abshire, P.A. Miller, T.W. Schlatter, 1990: Preliminary Evaluation of the first NOAA demonstation network wind profiler. *J. Atmos. Ocean. Tech.*, 7, 909-918.

APPENDIX A

WEATHER DISCUSSION FOR 1-10 APRIL 1991

April is commonly known for its weather extremes. Spring blizzards, tornado outbreaks, and heat waves contitute April's mixed bag of weather. For this reason, the first 10 days of April were selected as the verification period for Exercise 1. This appendix provides a brief synopsis of the weather during this period.

1 April

The month began with long wave troughs over the Great Lakes and the California coast. A weak surface low over the Great Lakes coupled with a low off the eastern seaboard produced cold temperatures, rain and snow over much of the Northeast. The tail end of a cold front from the Atlantic low produced heavy amounts of precipitation in Florida. High pressure dominated the central part of the United States as a Pacific system moved on shore.

2 April

As the complex low system exited the country to the northeast, a cold high pressure system slid southeast from Canada and dominated the East Coast weather for the next few days. The Pacific trough began to develop a surface low in the lee of the Colorado Rockies. Afternoon thunderstorms in Texas, Oklahoma, and Kansas produced up to 2.75 in. hail and spawned five tornadoes. Another system began to approach the Pacific Northwest.

3 April

The Colorado low dug southeastward to Oklahoma, producing thunderstorms over the Central Plains. Widespread hail of up to 2.75 in. and four tornadoes occurred in Texas, Kansas, and Oklahoma on the night of April 2. Rain continued in the Pacific Northwest.

4 April

Along the East Coast, a 1036 mb high was just east of Maryland. This set up a strong easterly gradient to the south, resulting in convective activity for Florida and Georgia. Miami received 2.75 in. in 24 hours. The Oklahoma low began to eject toward the northeast, bringing rain to most of the Mississippi Valley and widespread fog to the Central Plains. The Pacific Northwest remained wet, with over 3 in. in western Washington (3.32 in. at Seattle).

5 April

The low over the Central Plains continued to eject northeastward over the Great Lakes. Temperatures in the Northeast warmed 10°F in advance of the low that brought only light amounts of rain. A cutoff low over the southern Rockies produced large amounts of rain and flash floods in southern Texas. The Pacific Northwest continued to receive large amounts of rain. Some locations in Washington received 4 to 6 in. from 1200 UTC 3 April to 1200 UTC 5 April. Mud slides occurred in Oregon.

6 April

As the Great Lakes low moved out to the Atlantic, it produced rain from Miami to Maine. Amounts were generally light in the Northeast and heavy in the Southeast, with over 2 in. in central Florida. The cutoff low over western Texas brought heavy rain to the Texas and Louisiana coasts. Rain continued in the Pacific Northwest for the sixth day, but was lighter in amount and extended farther south into California and Nevada.

7 April

High pressure dominated the eastern half of the United States, and there was extensive fog in the Southeast. A Pacific front stretched from the upper peninsula of Michigan to Southern California, as a 1000 mb low formed in the lee of the Colorado and Wyoming Rockies. Rain fell from Seattle to Salt Lake City. Southerly flow ahead of the low caused fog in Kansas and Oklahoma.

8 April

The Pacific front continued to stretch from Maine to Arizona. Rain turned to snow along the Front Range of the Rockies. The high pressure system over the southeastern United States brought southerly flow to the Gulf Coast, resulting in widespread heavy precipitation from Texas to Alabama to Missouri. The Pacific Northwest remained wet. Ten in. of snow fell in Montana.

9 April

High pressure dominated the intermountain west, as the Pacific system is located in the upper Mississippi Valley. Large amounts of precipitation fell just about everywhere east of the Mississippi River. Thunderstorms with hail and tornadoes occurred from Illinois and Tennessee to western New York. Pennsylvania, Ohio, and West Virginia were especially hard hit. Yet another Pacific front brought rain to Washington and Oregon.

10 April

A strong 996 mb low was located northeast of the Great Lakes, and a cold front extended along all the Atlantic states. The next Pacific system pushed on shore bringing rain and snow to the Pacific Northwest and northern Rockies.

Summary

In general, April 1991 lived up to the aforementioned tradition of varied weather. The Pacific Northwest saw rain every day during the period, while Arizona and New Mexico were dry all 10 days (Fig. 75). Severe weather occurred over the Central Plains, while the Northern Rockies, Great Lakes, and the northeastern United States received light amounts of snow (Fig. 76). Table 19 gives some extreme values observed at the surface stations used in this exercise during the 10-day period.

Table 19. Extreme Weather Values Observed at the Surface Stations Used in Exercise 1 for the period 1-10 April 1991

Weather Values		Station	Date
Highest Temperature	100°F	Laredo, TX (LRD 72252)	8 Apr 22 UTC
Lowest Temperature	13°F	Laramie, WY (LAR 99485)	9 Apr 11 UTC
		Bradford, PA (BFD 99640)	3 Apr 11 UTC
Highest Altimeter	30.67 in.	Middletown, PA (MDT 99379)	3 Apr 13 UTC
Lowest Altimeter	29.33 in.	Akron, CO (AKO 99084)	10 Apr 23 UTC
Highest Wind Speed	50 kt	La Junta, CO (LHX 99075)	10 Apr 20 UTC
Max. 3-h Precip.	7.35 in.	Victoria, TX (VCT 72255)	5 Apr 10-13 UTC

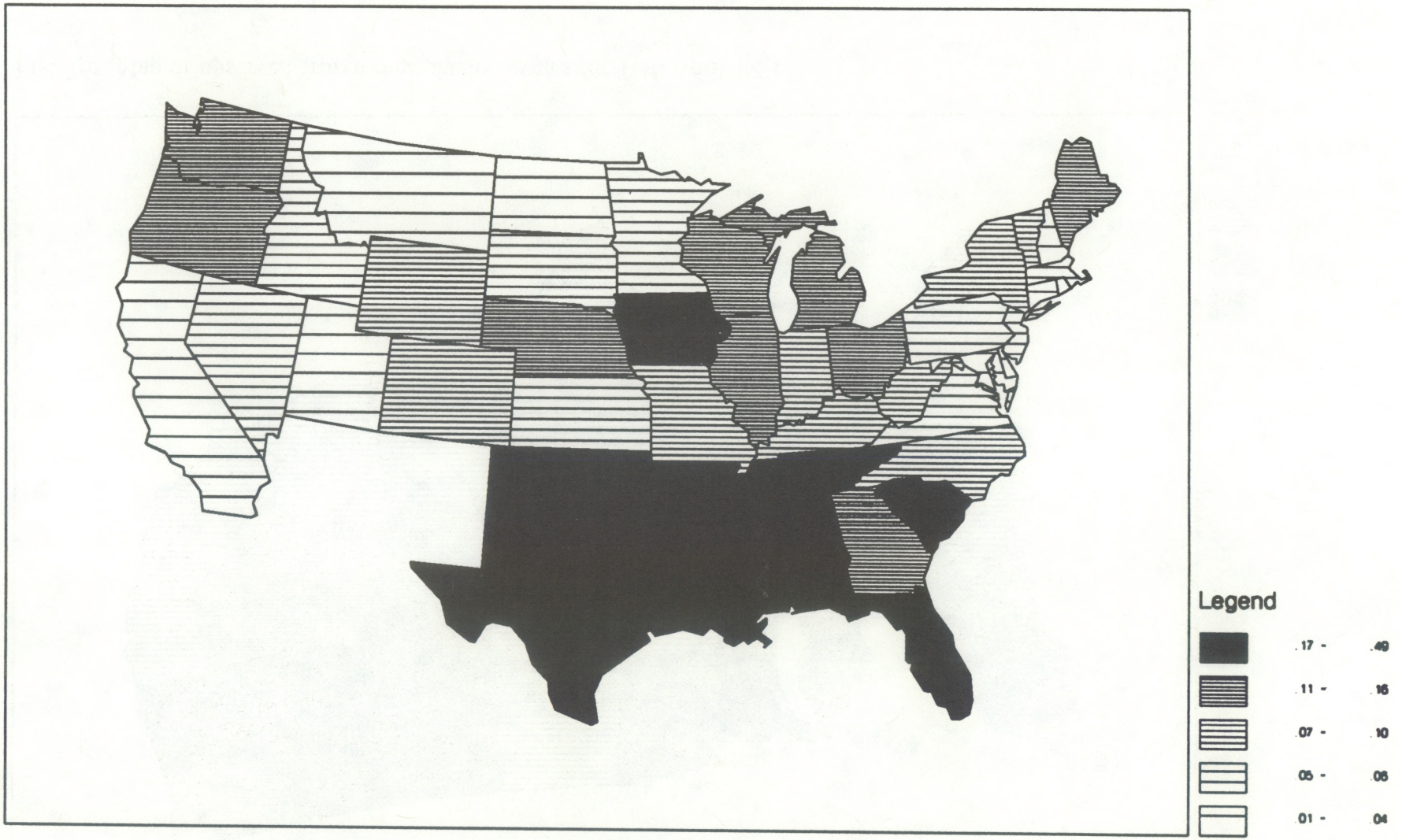

Fig. 75. Map of mean observed 3-h precipitation amounts (in.) for 1-10 April 1991.

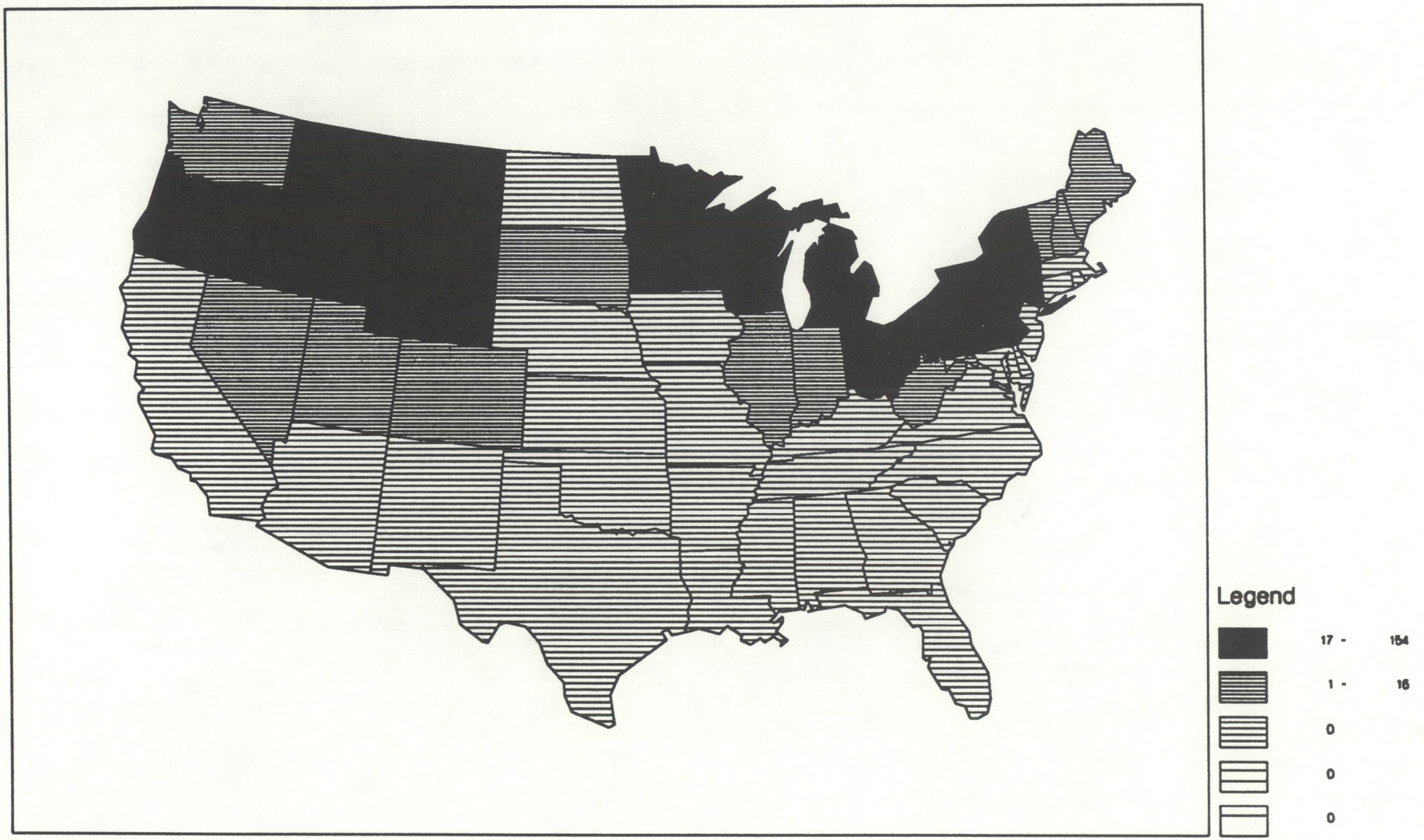

Fig. 76. Map of observed frozen precipitation events for 1-10 April 1991.

APPENDIX B

DIAGNOSING SAVS AND AIVS IN MAPS, Eta, LAPS AND RAMS

MAPS

1. Calculating state parameters

The first step in deriving all the AIVs at some horizontal location is to create a sounding of the three-dimensional SAVs. For each theta and sigma level, a value of virtual potential temperature, pressure, condensation pressure, Montgomery stream function, u component, and v component is bilinearly interpolated to the location in question. For theta levels, the virtual potential temperature value is just that which defines the level. Theta levels below the highest sigma level are thrown out. Using virtual potential temperature, pressure, condensation pressure, and Montgomery stream function, values of temperature, dewpoint, relative humidity, wet-bulb temperature, and height are calculated at each remaining sounding level. Also, u and v wind components are rotated from grid relative to true east-west and north-south components at each sounding level. This yields a sounding of all three-dimensional SAVs (plus relative humidity and wet-bulb temperature) over the location in question. Values required at height levels not in the model are linearly interpolated versus height; values for pressure levels are linearly interpolated versus the natural log of pressure.

From the sounding of three-dimensional SAVs, the values for the lowest sounding level are assumed to be values at the Earth's surface. Altimeter setting can be calculated directly from surface height and pressure. A value for accumulated precipitation can be obtained by horizontally interpolating values for convective and stable precipitation and adding the two. In order to apply some of the empirical algorithms yet to be described on initial fields, a value for the accumulated precipitation is needed. Since there is no accumulated precipitation available at analysis time, the accumulated precipitation value from the ensuing 3-h forecast is used for this purpose.

All coefficients and constant values are subjectively determined for the following AIV algorithms.

2. Diagnosing precipitation type

First an assumed precipitation rate, R, is calculated by dividing the accumulated precipitation by the 3-h output frequency. For the purpose of determining precipitation type, we define two thresholds, R' and RH', for precipitation rate and relative humidity, respectively. A precipitation type is not defined unless $R > R'$. Additionally, there must be two consecutive levels below 400 mb with a relative humidity (RH) greater than RH' for a

type to be diagnosed. The precipitation-generating level is assumed to be the highest adjacent levels where $R > R'$ and $RH > RH'$. Currently we are using $R' = 0.1$ mm hr^{-1} and $RH' = 0.95$.

The ice fraction (I) is the basic quantity being calculated for diagnosing the precipitation type, and its value is controlled by the wet-bulb temperature (T_w) sounding. The type depends on the ice fraction of precipitation arriving at the surface and the surface value of T_w. Before the ice fraction calculation is performed, two checks are made. If the surface temperature is greater than 7°C, then rain is automatically assumed. We define T_{wS} as the lowest wet-bulb temperature where super-cooled rainwater is allowed to exist. If $T_w < T_{wS}$ at the generating level and the entire sounding has $T_w < 0$°C, then we assume precipitation arrives at the surface entirely frozen. We are currently using $T_{wS} = -7$°C.

At the generating level, precipitation is assumed to be entirely liquid ($I = 0$) if $T_w > T_{wS}$; otherwise it is entirely frozen ($I = 1$). Once entirely liquid, precipitation will remain so until it falls through a level where $T_w < T_{wS}$. At that point it will become entirely solid. Once precipitation is entirely frozen, it will remain so until it reaches a level where $T_w > 0$°C. At that point it will begin to melt according to the formula

$$I = 1 - \frac{1}{E}\int_{p_0}^{p} (T_w - 0°C)\,dp \ ,$$

where p is pressure, p_0 is the pressure where T_w first becomes greater than 0°C, and E is the net area of the T_w curve where $T_w > 0$°C required for complete melting. E can be thought of as the "melting energy", though it does not have the units of energy. Currently, $E = 30$ mb °C. This formula allows partially melted precipitation to refreeze if once again $T_w < 0$°C at a lower level. This formula applies only when $0 < I < 1$ or when $I = 1$ and $T_w > 0$°C. Once entirely melted, precipitation must again fall to a level where $T_w < T_{wS}$ to refreeze. Also, once entirely refrozen, the energy calculation begins again from scratch if precipitation falls to another level where $T_w > 0$°C.

As previously mentioned, the precipitation type depends on the values of I and T_w at the surface. We define threshold ice fraction values of I_0 and I_1. If $I > I_1$, it is assumed that an observer would report the precipitation as entirely solid; if $I < I_0$, it is assumed that an observer would report liquid precipitation. An ice fraction between these values is diagnosed as mixed. Currently we use $I_0 = 0.1$ and $I_1 = 0.6$. In the case of liquid or mixed precipitation, it is diagnosed as freezing precipitation if $T_w < T_{wZ}$ at the surface. Currently $T_{wZ} = -1$°C. Solid precipitation is snow unless it was at one time completely liquid; in that case it is diagnosed as frozen precipitation.

3. Diagnosing visibility

The basic formula for diagnosing visibility is as follows:

$$\ln(V/V_0) = -\Delta_p - \Delta_f - \Delta_h - \Delta_{bd} - \Delta_{bs} + \Delta_{dry} ,$$

where V is visibility and V_0 is a value used for normal unobstructed visibility; currently 12 miles is used. The parameters Δ_p, Δ_f, Δ_h, Δ_{bd}, and Δ_{bs} represent incremental reductions in visibility due to contributions from precipitation, fog, haze, blowing dust, and blowing snow. In addition, Δ_{dry} is an increase applied for extremely dry conditions.

For calculating the contribution from precipitation, we define an effective snow rate, S, as follows

$$S = R[(1 - Q)I^{\frac{1}{3}} + Q] ,$$

where R is the precipitation rate and I is the ice fraction, as previously calculated. The value Q is how much of a reduction in visibility results from pure rain as opposed to the same rate of pure snow, and currently a value of 0.2 is used. The one-third power is an attempt to account for the fact that Stewart (1987) assumes that 90% of a snowflake must melt before it collapses. Finally, we have $\Delta_p = c_p S$, where currently $c_p = 1.2$ h mm^{-1}. In the case where no precipitation type has been identified, $\Delta_p = 0$.

The assumption is made that fog is mostly generated by high relative humidities, but is reduced by higher wind speeds. The contribution due to fog is calculated as follows:

$$\Delta_f = c_f \left[\frac{(RH - RH_f)}{(1 - RH_f)} \right]^{p_f} \cdot (U_f / U) ,$$

where RH is the surface relative humidity and U is the surface wind speed. When $RH < RH_f$, we assume no contribution due to fog. A U value less than U_f is set to U_f. Currently we have $c_f = 3.2$, $RH_f = 0.80$, $U_f = 2$ m s^{-1} and $p_f = 2$.

The assumption is made that haze is mostly generated by high dewpoints, but is reduced by higher winds speeds. The contribution due to haze is calculated as follows:

$$\Delta_h = c_h (T_d - T_{dh}) (U_h / U) ,$$

where T_d is the surface dewpoint and U is the surface wind speed. When $T_d < T_{dh}$, we assume no contribution due to haze. A U value less than U_h is set to U_h . Currently we have $c_h = 0.12°C^{-1}$, $T_{dh} = 7.2°C$, and $U_h = 0.5$ m s^{-1}.

The assumption is made that conditions must be both dry and windy to get blowing dust. The contribution due to blowing dust is calculated as follows:

$$\Delta_{bd} = c_{bd} (D - D_{bd}) (U - U_{bd}) ,$$

where D is the surface dewpoint depression. When $D < D_{bd}$ or $U < U_{bd}$, it is assumed that there is no contribution due to blowing dust. Currently, $c_{bd} = 0.0207$ s °C^{-1} m^{-1}, $D_{bd} = 16.7°C$ and $U_{bd} = 15$ m s^{-1}.

We assume that reduction in visibility by blowing snow is related to the ability of the wind to pick up already fallen snow. The ice fraction is used to determine how wet the snow is (how hard it is to pick up). The contribution due to blowing snow is calculated as follows:

$$\Delta_{bs} = c_{bs} RI^3 (U - U_{bs}) .$$

If $U < U_{bs}$ or there is no precipitation or the precipitation is all liquid, it is assumed that there is no contribution due to blowing snow. Currently, we are using $c_{bs} = 0.25$ h s m^{-1} mm^{-1} and $U_{bs} = 5$ m s^{-1}.

To allow for very high visibilities, the term Δ_{dry} is calculated.

$$\Delta_{dry} = (RH_f - RH) \ln (V_m / V_0) .$$

This term is only used when $RH < RH_f$; currently $V_m = 150$ nmi and $RH_f = 0.80$.

Before any of these phenomena other than precipitation are considered a significant obstruction to visibility, the total diagnosed visibility must be less than 7 mi. Individually,

an obstruction must reduce visibility to some fraction of its value without that obstruction before it is treated as though it would be reported in an SAO. Current values are fog, 0.5; haze, 0.9; blowing snow, 0.8; and blowing dust, 0.5.

4. Diagnosing clouds

First a sounding of cloud fraction is created. Let F_c be the cloud fraction, RH be the relative humidity, and RH_0 be the value of RH below which there are no clouds. We have

$$F_c = \left(\frac{RH - RH_0}{1 - RH_0} \right)^{p_c},$$

where $F_c < 0.1$ is defined as clear, $0.1 \leq F_c < 0.5$ is defined as scattered, $0.5 \leq F_c < 0.9$ is defined as broken, and $F_c \geq 0.9$ is defined as overcast.

The coefficients RH_0 and p_c are also assumed to have a profile versus height AGL as follows:

height agl (m)	RH_0	p_c
20000	.675	1
4500	.675	1
3000	.725	1
500	.725	1
200	.900	2
0	.975	3

A series of levels is chosen for the cloud sounding such that in the lower troposphere there are additional evenly spaced levels between levels in the model. Currently seven extra levels are inserted into the first model layer, three in the next four model layers, and one in the next three. At each cloud sounding level, values RH_0 and p_c are linearly interpolated versus height from their assumed profiles and a value of RH is likewise interpolated from the model sounding. Then a value of F_c is calculated at each cloud sounding level.

Each level in the cloud sounding is then examined, starting at the ground and working up, to identify cloud decks. A deck is recognized wherever a transition is made from one category (e.g., broken) to another and cloud fraction is increasing with height. The cloud cover of the deck is assigned the upper category. If two decks are identified within 100 m or 10% in height AGL from each other in a layer of steadily increasing RH, the lower deck is removed from consideration. Once an overcast deck is identified, no more decks are identified.

A broken or overcast cloud amount at the surface is defined to be an obscuration. If the surface cloud amount is broken, this is defined to be a partial obscuration, and additional decks may be reported. The height of obscuration is defined to be the visibility or 100 m, whichever is less. The ceiling is defined in the conventional manner, which is the height of the first broken or overcast deck (obscurations included). The cloud top is defined as the highest level at which $F_c > F_{cT}$, where currently $F_{cT} = 0.75$. While one could identify any number of decks, currently only five decks are identified. Decks of higher coverage are saved in favor of others if more than five could be identified.

Eta

The SAVs and AIVs were diagnosed from Eta model output and interpolated to the sites chosen for the verification exercise. The stair-step representation of the terrain (30 km resolution) and the interpolation procedure generated a disagreement between the elevation of the observation sites and those obtained from the model. The computation of AIVs was done assuming that each Eta station was representative of the real station; that is, no extrapolation or interpolation from model elevation to true elevation was attempted.

The following SAVs and AIVs are obtained directly from the model output. Since the Eta model does not forecast variables at the surface, it was necesary to use the following procedures:

* Temperature

Temperature was interpolated to the surface (anemometer level, 2 m AGL) using the temperatures at the lowest Eta level and the skin temperature (subsoil temperature).

* Dewpoint temperature

The dewpoint depression at the lowest Eta level was computed and the dewpoint at the surface was defined as the surface temperature minus this dewpoint depression.

* Winds

The winds at the surface were set equal to those at the lowest Eta level.

* Altimeter Setting

The the altimeter setting is computed from the sea-level pressure that is part of the output from the Eta model.

* Precipitation Phase

Precipitation with surface temperatures below 32°F was diagnosed as snow; otherwise it was diagnosed as rain.

* Precipitation Amount

The precipitation amount is directly generated by the model from the Betts and Miller (1986) convective adjustment scheme.

The remaining AIVs (ceiling, cloud heights, cloud amount, and cloud top) were diagnosed based on the relative humidity (the Eta model does not have explicit cloud water). These were estimated as follows:

* Cloud Base/Ceiling/Top

Relative humidities of 66%, 50%, and 40% were assigned to low, middle, and high clouds, respectively.

* Cloud Amount

The estimation of this parameter followed a simplified version of Krishnamurti et al. (1991) alogorithm in which the cloud amount is equal to

$$[(RH - RH_c)/(100 - RH_c)]^2 ,$$

where RH_c is 66%, 50%, and 40% for low, middle, and high clouds, respectively.

Visibility and obstruction to visibility were not estimated since they require explicit paramaterization of cloud water in the Eta model.

LAPS

1. General analysis techniques

Three files from the LAPS analyses are used in this verification exercise: the two-dimensional surface analysis (LSX), and the three-dimensional analyses of temperature (LT1) and wind (LW3). Because LAPS analyses are computed on constant pressure surfaces, heights are a derived field; they are computed hydrostatically using the LAPS 1500-m pressure (LSX) and the LAPS three-dimensional temperature field (LT1).

The LAPS surface analysis employs a two-pass Barnes analysis and a horizontal spline to compute grid-point values from surrounding data. Structures in temperature and

moisture are maintained between data stations using satellite data and "horizontal shape matching" (McGinley, 1982) to help define gradients. The LAPS three-dimensional temperature field is primarily the MAPS background temperature with LAPS surface temperatures blended up from the ground, based in part on a dry-adiabatic adjustment. LAPS winds and humidity are based on the MAPS background fields, adjusted with available data sources. More detailed descriptions of the current LAPS can be found in Stamus (1992) and Albers (1992).

In all cases, data values are horizontally interpolated from the LAPS grid to the specific locations using the four surrounding grid values, either along the surface or on constant pressure surfaces.

2. Surface comparison

Surface values are extracted from the LAPS surface analysis (LSX), with the exception of cloud parameters. These values include temperature, dewpoint, wind speed and direction, and altimeter. (Note: Precipitation and visibility values were not computed for this exercise.) Temperature and dewpoint are converted from Kelvin to degrees Celsius; wind speed and wind direction are derived from the horizontal u- and v-components of the surface wind; otherwise, no changes are made to the LAPS values of these fields.

Cloud ceiling, cloud heights, cloud amount, and cloud top values are extracted from the LAPS three-dimensional cloud analysis (LC3). First, a 42-level profile of cloud amount is interpolated from the LAPS grid to the location. This profile is then examined from the ground upward, looking for levels where the cloud amount is greater than 0.05%. When the criterion is met, a cloud is assumed; levels are then checked to see where the cloud amount drops below 0.05%. Once this criterion is reached, a cloud layer is determined.

Cloud heights are given by the lowest level of the cloud layer. Cloud amount for the individual cloud layer is assumed to be the maximum cloud amount within the layer. The *total* cloud amount is the summation of the individual cloud amount from the surface to that point. Skies are considered clear (CLR) for total cloud amounts less than 0.05%, scattered (SCT) for total cloud amounts between 0.05% and 0.55%, broken (BKN) for amounts between 0.55% and 0.95%, and overcast (OVC) for amounts over 0.95%. Obscured (OBS) cloud conditions were not separated from general overcast conditions in this verification study.

If the total cloud amount exceeds the broken (BKN) sky condition, a ceiling is reported at the lowest level of that cloud layer. Cloud tops are reported by looking from the top of the cloud amount profile down to the surface. If a cloud amount over 0.05% is found, the cloud top is reported at that level.

3. Rawinsonde comparison

Three different LAPS analyses are combined to reproduce rawinsonde data for comparison. The LAPS three-dimensional temperature file (LT1) is used to derive a horizontally interpolated profile of temperature, and that profile is integrated hydrostatically to compute heights. Wind speed and wind direction are derived from the horizontal wind components in the three-dimensional LAPS wind field (LW3).

4. Profiler comparison

The LAPS three-dimensional wind field (LW3) contains the three components of the wind for comparison with profiler reports. LAPS saves the vertical component of the wind in pressure coordinates (ω), so that a conversion is necessary to the obtain the vertical component in height coordinates, w. For this verification exercise, the standard approximate relationship was used to obtain w,

$$w = -\frac{RT}{pg}\omega \ .$$

RAMS

1. Surface

* Wind Speed

The model-predicted wind speed at the lowest level (150 m AGL) was multiplied by 0.7 to obtain the surface wind speed.

* Wind Direction

The surface wind direction was based on the reduced u- and v-components computed for the surface wind speed above.

* Temperature

The temperatures from the lowest model level and the top soil level were averaged to obtain the surface temperature.

* Dewpoint

The surface dewpoint was taken directly from the dewpoint of the lowest model level (150 m AGL).

* Altimeter

The altimeter setting was based on formula given on page 269 of the Smithsonian Meteorological Tables (1966).

* Precipitation Amount

This value was derived from the cumulus parameterization scheme.

* Ceiling Height/Cloud Top

Determination of the ceiling height and cloud top was based on presence of cloud water (> 0.0) explicitly generated by the model.

2. Upper air

All variables for RAOB and profiler comparison were directly interpolated to station location from the model grid.

APPENDIX C

LIST OF MAPS STATISTICAL RESULTS

UPPER AIR - RAWINSONDE

HEIGHT (m)

	00h	03h	06h	12h
850 mb ME	-1.0	+0.0	-0.2	-6.3
850 mb MAE	5.0	9.3	9.2	13.6
850 mb RMSE	7.2	12.0	12.1	17.1
500 mb ME	-15.8	-18.5	-18.0	-20.6
500 mb MAE	17.1	20.9	20.8	24.5
500 mb RMSE	22.1	26.6	26.8	30.9
250 mb ME	-21.4	-26.0	-23.8	-21.8
250 mb MAE	23.1	31.2	30.3	33.4
250 mb RMSE	33.2	41.0	40.1	43.3

TEMPERATURE (°C)

	00	03	06	12
850 mb ME	+0.0	-0.1	-0.2	-0.3
850 mb MAE	0.7	1.4	1.5	1.7
850 mb RMSE	1.2	1.9	2.1	2.4
500 mb ME	-0.0	-0.3	-0.2	-0.1
500 mb MAE	0.5	1.0	1.0	1.0
500 mb RMSE	0.9	1.4	1.4	1.4
250 mb ME	0.4	0.2	0.3	0.7
250 mb MAE	0.6	1.0	1.1	1.3
250 mb RMSE	1.0	1.5	1.6	1.9

UPPER AIR - RAWINSONDE (cont'd)

DEWPOINT TEMPERATURE (°C)

	00h	03h	06h	12h
850 mb ME	-0.8	-3.1	-2.6	-1.9
850 mb MAE	2.4	5.1	4.8	4.4
850 mb RMSE	3.7	7.0	6.4	6.0
500 mb ME	-0.4	+0.4	+0.3	-0.0
500 mb MAE	4.1	6.7	6.7	6.8
500 mb RMSE	5.8	8.7	8.7	9.0

DEWPOINT DEPRESSION (°C)

	00	03	06	12
850 mb ME	0.8	3.0	2.4	1.6
850 mb MAE	2.7	5.4	5.0	4.9
850 mb RMSE	4.1	7.3	6.7	6.6
500 mb ME	+0.3	-0.7	-0.5	-0.1
500 mb MAE	4.2	6.8	6.8	6.9
500 mb RMSE	5.9	8.9	8.8	9.1

U-COMPONENT ($m\ s^{-1}$)

	00	03	06	12
850 mb ME	0.1	0.2	0.3	0.6
850 mb MAE	1.5	2.7	2.7	2.9
850 mb RMSE	2.1	3.5	3.5	3.7
500 mb ME	+0.0	-0.1	-0.2	-0.4
500 mb MAE	1.6	2.6	2.8	2.9
500 mb RMSE	2.2	3.5	3.6	3.9
250 mb ME	+0.1	-0.4	-0.6	-1.0
250 mb MAE	2.1	3.4	3.7	4.2
250 mb RMSE	2.7	4.4	4.9	5.6

UPPER AIR - RAWINSONDE (cont'd)

V-COMPONENT ($m\ s^{-1}$)

	00h	03h	06h	12h
850 mb ME	-0.4	-0.7	-0.6	-0.6
850 mb MAE	1.6	2.9	2.9	3.1
850 mb RMSE	2.2	3.8	3.8	4.1
500 mb ME	0.2	0.2	0.3	0.2
500 mb MAE	1.7	2.8	2.8	3.0
500 mb RMSE	2.3	3.6	3.7	4.0
250 mb ME	+0.0	+0.0	-0.1	-0.2
250 mb MAE	2.1	3.5	3.7	4.1
250 mb RMSE	3.0	4.7	5.1	5.6

ROOT MEAN SQUARE VECTOR ERROR ($m\ s^{-1}$)

850 mb RMSVE	3.0	5.2	5.2	5.5
500 mb RMSVE	3.2	5.0	5.2	5.6
250 mb RMSVE	4.0	6.5	7.1	7.9

WIND SPEED (kt)

	00	03	06	12
850 mb ME	-0.8	+0.1	+0.5	+0.9
850 mb MAE	2.9	5.0	5.2	5.2
850 mb RMSE	3.9	6.6	6.7	6.8
500 mb ME	-0.1	-0.1	-0.2	-0.4
500 mb MAE	3.1	5.1	5.3	5.6
500 mb RMSE	4.2	6.8	7.0	7.5
250 mb ME	-0.6	-1.4	-1.8	-2.5
250 mb MAE	4.0	6.6	7.2	8.1
250 mb RMSE	5.6	8.9	9.8	11.0

UPPER AIR - RAWINSONDE (cont'd)

WIND DIRECTION (degrees)

	00h	03h	06h	12h
850 mb MAE	9.8	18.3	18.2	19.7
850 mb SD	11.8	18.1	17.3	20.0
500 mb MAE	6.8	11.2	11.5	12.0
500 mb SD	6.9	11.7	11.8	12.0
250 mb MAE	4.7	8.1	9.0	9.7
250 mb SD	5.6	9.8	12.9	13.4

SURFACE

ALTIMETER (in. Hg)

	00	03	06	09	12
ME	0.00	0.00	0.00	+0.01	-0.01
MAE	0.03	0.04	0.05	0.05	0.07
RMSE	0.04	0.06	0.06	0.07	0.08

TEMPERATURE (°F)

	00	03	06	09	12
ME	-2.8	-3.6	-4.4	-6.6	-4.3
MAE	6.5	8.1	9.3	12.8	11.3
RMSE	8.6	10.1	11.5	15.0	13.6

DEWPOINT TEMPERATURE (°F)

	00	03	06	09	12
ME	-3.9	-4.4	-4.7	-5.2	-5.4
MAE	5.1	5.7	6.2	6.9	7.1
RMSE	6.7	7.4	8.0	8.9	9.2

SURFACE (cont'd)

DEWPOINT DEPRESSION (°F)

	00h	03h	06h	09h	12h
ME	+1.1	+0.7	+0.4	-1.5	+1.1
MAE	6.4	8.1	9.6	13.1	12.3
RMSE	8.7	10.5	12.2	15.7	15.3

WIND SPEED (kt)

	00	03	06	09	12
ME	7.5	8.5	8.7	8.3	9.4
MAE	8.2	9.1	9.3	9.0	10.0
RMSE	9.9	10.7	10.9	10.6	11.5

WIND DIRECTION (degrees)

	00	03	06	09	12
MAE	27.3	27.7	28.5	29.9	32.0
SD	29.0	29.5	30.2	31.7	34.4

PRECIPITATION AMOUNT (in.)

	00	03	06	09	12
ME	n/a	-.03	-.03	-.02	-.03
MAE	n/a	.07	.08	.07	.08
RMSE	n/a	.22	.22	.20	.28

CEILING HEIGHT (hundreds ft)

	00	03	06	09	12
ME	15.0	16.0	17.1	6.1	11.8
MAE	29.0	31.6	34.9	31.5	36.6
SD	36.9	40.4	42.8	37.0	40.0

SURFACE (cont'd)

VISIBILITY (mi)

	00h	03h	06h	09h	12h
ME	4.3	4.7	5.4	6.1	7.3
MAE	4.7	5.2	5.9	6.7	7.9
SD	5.6	6.2	6.8	7.7	8.3

CLOUD-TOP HEIGHT (hundreds ft)

	00	03	06	09	12
ME	-162.9	-172.0	-173.4	-175.0	-159.6
MAE	171.7	180.1	181.3	182.8	167.8
RMSE	200.5	207.5	210.5	212.7	196.8

APPENDIX D

LIST OF ETA STATISTICAL RESULTS

UPPER AIR - RAWINSONDE

HEIGHT (m)

	00h	12h	24h
850 mb ME	-0.6	-0.5	+8.9
850 mb MAE	5.3	11.4	16.1
850 mb RMSE	6.8	19.0	21.0
500 mb ME	-3.3	-1.3	-0.6
500 mb MAE	7.5	14.0	14.6
500 mb RMSE	11.7	28.0	21.5
250 mb ME	-8.5	+0.6	-7.8
250 mb MAE	14.3	22.1	23.7
250 mb RMSE	25.3	41.5	37.5

TEMPERATURE (°C)

	00	12	24
850 mb ME	+0.1	+0.1	-0.9
850 mb MAE	0.9	1.4	1.8
850 mb RMSE	1.1	2.1	2.4
500 mb ME	-0.2	-0.0	-0.4
500 mb MAE	0.8	1.0	1.1
500 mb RMSE	1.2	1.6	1.7
250 mb ME	0.1	0.6	0.1
250 mb MAE	1.0	1.4	1.5
250 mb RMSE	1.4	2.1	2.1

UPPER AIR - RAWINSONDE (cont'd)

DEWPOINT TEMPERATURE (°C)

	00h	12h	24h
850 mb ME	-0.3	-0.8	-0.5
850 mb MAE	1.4	4.2	4.3
850 mb RMSE	4.4	6.2	6.1
500 mb ME	0.5	1.9	2.3
500 mb MAE	3.0	7.2	7.3
500 mb RMSE	4.7	9.9	9.7

DEWPOINT DEPRESSION (°C)

	00	12	24
850 mb ME	+0.4	+0.9	-0.4
850 mb MAE	1.8	4.5	5.1
850 mb RMSE	2.5	6.6	7.2
500 mb ME	-0.7	-1.9	-2.8
500 mb MAE	3.1	7.3	7.4
500 mb RMSE	4.7	10.0	9.9

U-COMPONENT ($m\ s^{-1}$)

	00	12	24
850 mb ME	-0.3	-0.5	-0.2
850 mb MAE	1.4	2.8	3.2
850 mb RMSE	2.1	3.7	4.1
500 mb ME	-0.3	-0.7	-0.2
500 mb MAE	1.7	3.2	3.3
500 mb RMSE	2.2	4.6	4.4
250 mb ME	-0.2	+1.0	+0.8
250 mb MAE	2.3	4.4	5.3
250 mb RMSE	3.0	6.2	7.0

APPENDIX D - Eta (cont'd)

UPPER AIR - RAWINSONDE (cont'd)

V-COMPONENT ($m\ s^{-1}$)

	00h	12h	24h
850 mb ME	-0.1	-0.1	+0.1
850 mb MAE	1.5	3.1	3.2
850 mb RMSE	2.0	4.2	4.2
500 mb ME	+0.1	-0.5	+0.1
500 mb MAE	1.7	3.5	3.7
500 mb RMSE	2.2	5.6	5.3
250 mb ME	+0.1	-0.8	-0.2
250 mb MAE	2.3	5.2	5.3
250 mb RMSE	3.3	8.1	7.4

ROOT MEAN SQUARE ERROR ($m\ s^{-1}$)

850 mb RMSVE	2.9	5.6	5.9
500 mb RMSVE	3.1	7.2	6.9
250 mb RMSVE	4.5	10.2	12.4

WIND SPEED (kt)

	00	12	24
850 mb ME	-1.3	-0.2	+0.4
850 mb MAE	3.0	5.5	6.3
850 mb RMSE	4.4	7.3	8.1
500 mb ME	-0.7	-1.3	-0.2
500 mb MAE	3.2	6.4	6.7
500 mb RMSE	4.2	9.5	9.1
250 mb ME	-1.1	+1.0	+0.5
250 mb MAE	4.7	8.8	10.4
250 mb RMSE	6.4	13.0	13.8

UPPER AIR - RAWINSONDE (cont'd)

WIND DIRECTION (degrees)

	00h	12h	24h
850 mb MAE	8.1	19.2	20.6
850 mb SD	11.8	25.4	22.0
500 mb MAE	6.8	14.7	15.0
500 mb SD	6.2	18.8	18.3
250 mb MAE	5.1	10.7	11.6
250 mb SD	7.0	14.1	14.3

SURFACE

ALTIMETER (in. Hg)

	00	03	06	09	12	15	18	24
ME	-.01	-.03	-.01	+.01	-.02	-.02	-.01	+.04
MAE	.03	.04	.03	.03	.04	.04	.04	.06
RMSE	.04	.05	.04	.05	.05	.05	.06	.08

TEMPERATURE (°F)

	00	03	06	09	12	15	18	24
ME	-1.4	-0.3	-1.6	-3.3	-5.4	-3.2	-2.9	-4.1
MAE	5.8	4.9	5.0	5.5	6.8	6.6	6.2	6.5
RMSE	8.5	6.9	7.1	7.9	9.3	9.4	8.9	9.5

DEWPOINT TEMPERATURE (°F)

	00	03	06	09	12	15	18	24
ME	-4.2	-9.2	-12.7	-15.2	-17.4	-8.3	-0.3	-1.3
MAE	7.5	10.0	13.3	15.8	17.9	12.0	7.9	7.1
RMSE	11.1	13.9	17.3	19.9	22.4	17.6	11.1	10.2

SURFACE (cont'd)

DEWPOINT DEPRESSION (°F)

	00h	03h	06h	09h	12h	15h	18h	24h
ME	+2.8	+9.0	+11.1	+11.9	+12.0	+5.0	-2.5	-2.9
MAE	7.0	10.1	11.9	12.6	12.7	8.7	8.2	8.6
RMSE	9.7	13.3	15.4	16.3	16.8	12.7	10.8	11.5

WIND SPEED (kt)

	00	03	06	09	12	15	18	24
ME	-0.5	+3.7	+5.6	+6.3	+6.0	+1.2	-2.5	+0.5
MAE	5.4	7.5	8.7	8.9	8.7	6.1	5.8	7.2
RMSE	6.9	9.1	10.6	11.0	10.8	7.9	7.4	8.9

WIND DIRECTION (degrees)

	00	03	06	09	12	15	18	24
MAE	26.6	28.3	27.7	27.2	30.5	25.3	27.1	32.7
SD	29.9	30.9	29.7	30.0	31.5	28.1	27.9	35.3

PRECIPITATION AMOUNT (in.)

	00	03	06	09	12	15	18	24
ME	n/a	-.06	-.05	-.06	-.07	-.05	-.06	n/a
MAE	n/a	.09	.09	.09	.10	.08	.09	n/a
RMSE	n/a	.16	.20	.21	.41	.20	.31	n/a

CEILING HEIGHT (hundreds ft)

	00	03	06	09	12	15	18	24
ME	-26.2	-23.1	-21.9	-19.9	-19.2	-20.6	-25.2	-30.1
MAE	29.1	26.2	24.7	22.7	22.2	22.7	26.5	32.1
SD	19.4	29.9	29.4	27.8	27.4	27.3	29.8	18.9

SURFACE (cont'd)

CLOUD-TOP HEIGHT (hundreds ft)

	00h	03h	06h	09h	12h	15h	18h	24h
ME	71.7	87.5	96.1	112.3	112.3	106.0	102.4	80.5
MAE	88.0	99.6	109.2	120.5	121.1	115.4	110.7	96.1
RMSE	109.3	120.6	130.3	141.9	143.4	139.7	134.3	116.2

APPENDIX E

LIST OF LAPS STATISTICAL RESULTS

RAWINSONDE

	ME	MAE	RMSE
500 Height (m)	6.1	14.8	24.0
500 Temperature (°C)	-0.6	1.0	1.2
500 Dewpoint (°C)	-2.7	5.9	7.4
500 u-component (m s^{-1})	-1.0	1.5	2.0
500 v-component (m s^{-1})	0.1	1.4	1.7
500 Wind speed (kt)	-1.9	3.3	4.0
500 Wind direction (degrees)		5.1	
250 Height (m)	2.2	15.7	22.0
250 Temperature (°C)	0.1	1.1	1.8
250 u-component (m s^{-1})	-1.7	3.1	4.5
250 v-component (m s^{-1})	-1.0	2.9	4.0
250 Wind speed (kt)	-4.0	5.6	7.6
250 Wind direction (degrees)		5.9	

SURFACE

	ME	MAE	RMSE
Altimeter (in. Hg)	0.02	0.06	0.07
Cloud top (hundreds ft)	35.5	47.3	66.1
Ceiling (hundreds ft)	-3.3	9.5	15.0
Temperature (°F)	0.2	1.3	1.8
Dewpoint (°F)	0.4	1.2	2.5
Dewpoint depression (°F)	-0.3	1.8	3.1
Wind speed (kt)	-1.5	1.9	2.6
Wind direction (degrees)		8.7	

APPENDIX F

LIST OF RAMS STATISTICAL RESULTS

UPPER AIR- RAWINSONDE

	HEIGHT (m)	
	00h	12h
Colorado/Central Plains		
850 mb ME	+20.3	-417.4
850 mb MAE	21.9	442.9
850 mb RMSE	25.0	2212.1
500 mb ME	18.4	15.6
500 mb MAE	20.7	21.0
500 mb RMSE	24.2	24.5
250 mb ME	18.2	13.5
250 mb MAE	19.4	22.9
250 mb RMSE	23.9	26.8
Florida/Northeast		
850 mb ME	25.6	18.6
850 mb MAE	25.6	33.5
850 mb RMSE	28.2	37.7
500 mb ME	29.6	28.5
500 mb MAE	29.6	33.0
500 mb RMSE	33.6	36.4
250 mb ME	25.1	42.8
250 mb MAE	30.8	42.8
250 mb RMSE	37.6	50.2

UPPER AIR - RAWINSONDE (cont'd)

TEMPERATURE (°C)

	00h	12h
Colorado/Central Plains		
850 mb ME	2.2	3.7
850 mb MAE	2.5	3.7
850 mb RMSE	2.7	4.1
500 mb ME	-0.7	-1.0
500 mb MAE	0.8	1.3
500 mb RMSE	1.0	1.5
250 mb ME	-0.6	+0.6
250 mb MAE	1.1	1.4
250 mb RMSE	2.0	1.9
Florida/Northeast		
850 mb ME	+0.5	-0.5
850 mb MAE	1.0	1.6
850 mb RMSE	1.2	2.2
500 mb ME	-0.3	-0.1
500 mb MAE	0.9	0.8
500 mb RMSE	1.2	1.1
250 mb ME	+0.0	-0.1
250 mb MAE	1.4	1.0
250 mb RMSE	2.7	1.4

UPPER AIR - RAWINSONDE (cont'd)

DEWPOINT TEMPERATURE (°C)

	00h	12h
Colorado/Central Plains		
850 mb ME	-4.2	-4.6
850 mb MAE	4.9	4.7
850 mb RMSE	6.3	5.9
500 mb ME	1.1	1.8
500 mb MAE	6.7	8.3
500 mb RMSE	8.6	10.4
Florida/Northeast		
850 mb ME	-2.0	-1.2
850 mb MAE	2.6	2.8
850 mb RMSE	3.6	3.8
500 mb ME	+2.0	-0.7
500 mb MAE	5.8	6.6
500 mb RMSE	7.1	8.3

DEWPOINT DEPRESSION (°C)

	00	12
Colorado/Central Plains		
850 mb ME	6.4	8.3
850 mb MAE	7.2	8.3
850 mb RMSE	8.5	9.4
500 mb ME	-1.8	-2.8
500 mb MAE	6.8	8.8
500 mb RMSE	8.8	10.7

UPPER AIR - RAWINSONDE (cont'd)

	00h	12h
Florida/Northeast		
850 mb ME	2.5	0.7
850 mb MAE	3.3	3.2
850 mb RMSE	4.5	4.3
500 mb ME	-2.3	+0.6
500 mb MAE	5.9	7.0
500 mb RMSE	7.4	8.8

U-COMPONENT ($m\ s^{-1}$)

	00	12
Colorado/Central Plains		
850 mb ME	1.5	2.0
850 mb MAE	2.8	2.9
850 mb RMSE	3.5	3.6
500 mb ME	+0.4	-0.7
500 mb MAE	2.2	3.1
500 mb RMSE	2.7	4.0
250 mb ME	-3.0	-2.2
250 mb MAE	4.6	3.9
250 mb RMSE	5.7	5.2
Florida/Northeast		
850 mb ME	-0.3	+0.6
850 mb MAE	2.3	1.6
850 mb RMSE	2.8	2.0
500 mb ME	-1.0	-1.1
500 mb MAE	2.5	3.0
500 mb RMSE	4.4	3.3

UPPER AIR - RAWINSONDE (cont'd)

	00h	12h
250 mb ME	-0.9	-1.6
250 mb MAE	3.2	4.5
250 mb RMSE	4.8	5.3

V-COMPONENT ($m\ s^{-1}$)

	00	12
Colorado/Central Plains		
850 mb ME	1.2	1.1
850 mb MAE	2.3	3.6
850 mb RMSE	2.8	4.9
500 mb ME	-0.5	-0.6
500 mb MAE	3.2	3.7
500 mb RMSE	4.2	4.7
250 mb ME	-0.3	-1.0
250 mb MAE	3.8	5.8
250 mb RMSE	5.9	7.9
Florida/Northeast		
850 mb ME	+0.0	-1.4
850 mb MAE	2.0	2.1
850 mb RMSE	2.5	2.9
500 mb ME	-0.1	+0.1
500 mb MAE	2.4	2.3
500 mb RMSE	3.3	3.2
250 mb ME	-0.9	+0.6
250 mb MAE	3.0	3.4
250 mb RMSE	3.8	4.2

UPPER AIR - RAWINSONDE (cont'd)

ROOT MEAN SQUARE VECTOR ERROR ($m\ s^{-1}$)

	00h	12h
Colorado/Central Plains		
850 mb RMSVE	4.5	6.1
500 mb RMSVE	5.0	6.1
250 mb RMSVE	8.2	9.5
Florida/Northeast		
850 mb RMSVE	3.8	3.5
500 mb RMSVE	5.5	4.6
250 mb RMSVE	6.2	6.7

WIND SPEED (kt)

	00	12
Colorado/Central Plains		
850 mb ME	1.4	1.6
850 mb MAE	4.7	5.6
850 mb RMSE	5.6	7.3
500 mb ME	-0.9	+0.1
500 mb MAE	5.0	5.0
500 mb RMSE	6.3	6.5
250 mb ME	-5.6	-4.8
250 mb MAE	8.6	8.2
250 mb RMSE	11.0	10.7
Florida/Northeast		
850 mb ME	-0.9	-0.4
850 mb MAE	4.1	3.1
850 mb RMSE	5.1	4.4

UPPER AIR - RAWINSONDE (cont'd)

	00h	12h
500 mb ME	-2.1	-1.8
500 mb MAE	5.2	6.2
500 mb RMSE	9.0	6.9
250 mb ME	-1.1	-5.3
250 mb MAE	6.8	9.3
250 mb RMSE	10.0	10.9

WIND DIRECTION (degrees)

	00	12
Colorado/Central Plains		
850 mb MAE	13.2	19.6
850 mb SD	9.4	17.1
500 mb MAE	14.9	21.3
500 mb SD	15.9	31.6
250 mb MAE	12.0	21.0
250 mb SD	12.9	29.4
Florida/Northeast		
850 mb MAE	10.8	12.4
850 mb SD	10.0	9.1
500 mb MAE	7.7	7.0
500 mb SD	6.7	7.7
250 mb MAE	4.8	6.7
250 mb SD	4.0	6.8

APPENDIX F - RAMS (cont'd)

SURFACE

ALTIMETER (in. Hg)

	00h	06h	12h
Colorado/Central Plains			
ME	0.11	0.04	0.07
MAE	0.11	0.06	0.07
RMSE	0.13	0.08	0.09
No. of Cases	112	139	121
Florida/Northeast			
ME	0.13	0.14	0.14
MAE	0.13	0.14	0.14
RMSE	0.13	0.15	0.15
No. of Cases	100	132	127

TEMPERATURE (°F)

	00	06	12
Colorado/Central Plains			
ME	0.9	4.4	6.6
MAE	5.9	8.2	8.4
RMSE	7.5	9.6	10.1
No. of Cases	111	139	121
Florida/Northeast			
ME	0.5	0.3	0.4
MAE	4.6	6.3	5.0
RMSE	5.7	7.5	6.0
No. of Cases	96	128	123

SURFACE (cont'd)

DEWPOINT TEMPERATURE (°F)

	00h	06h	12h
Colorado/Central Plains			
ME	-8.0	-7.5	-6.0
MAE	9.8	10.2	9.3
RMSE	11.7	12.5	11.4
No. of Cases	112	139	121
Florida/Northeast			
ME	-6.8	-3.3	-3.9
MAE	7.3	8.5	7.6
RMSE	8.7	10.5	9.5
No. of Cases	96	128	123

DEWPOINT DEPRESSION (°F)

	00	06	12
Colorado/Central Plains			
ME	8.9	11.8	12.5
MAE	11.1	16.6	16.3
RMSE	13.7	19.2	18.6
No. of Cases	111	139	121
Florida/Northeast			
ME	7.2	3.6	4.3
MAE	9.0	12.8	9.7
RMSE	11.1	15.2	11.8
No. of Cases	96	128	123

SURFACE (cont'd)

WIND SPEED (kt)

	00h	06h	12h
Colorado/Central Plains			
ME	2.3	3.1	4.8
MAE	5.8	5.5	6.3
RMSE	9.0	6.7	7.3
No. of Cases	112	139	121
Florida/Northeast			
ME	4.3	3.5	5.7
MAE	5.7	5.1	6.3
RMSE	6.9	6.2	7.9
No. of Cases	100	132	127

WIND DIRECTION (degrees)

	00	06	12
Colorado/Central Plains			
MAE	33.5	23.6	33.5
SD	38.5	32.0	34.9
No. of Cases	57	84	71
Florida/Northeast			
MAE	12.8	27.3	17.2
SD	12.1	28.2	11.4
No. of Cases	22	54	42

SURFACE (cont'd)

PRECIPITATION AMOUNT (in.)

	00h	06h	12h
Colorado/Central Plains			
ME	-.02		-.05
MAE	.04		.06
RMSE	.04		.10
No. of Cases	4	0	14
Florida/Northeast			
ME		-.01	-.70
MAE		.01	.70
RMSE		.02	.70
No. of Cases	0	8	1

CEILING HEIGHT (hundreds ft)

	00	06	12
Colorado/Central Plains			
ME	39.7	18.0	55.2
MAE	48.1	42.1	71.9
SD	46.2	35.6	60.6
No. of Cases	9	20	22
Florida/Northeast			
ME	+43.5	-6.4	-12.7
MAE	56.3	35.4	39.3
SD	59.3	23.1	27.8
No. of Cases	20	26	38

SURFACE (cont'd)

CLOUD-TOP HEIGHT (hundreds ft)

	00h	06h	12h
Colorado/Central Plains			
ME	-38.5	-24.0	+9.8
MAE	57.2	55.9	67.9
RMSE	67.1	70.8	83.0
No. of Cases	12	34	29
Florida/Northeast			
ME	-59.7	-139.7	-68.6
MAE	63.5	141.4	76.7
RMSE	101.1	170.4	106.7
No. of Cases	20	27	27

APPENDIX G

LIST OF SYSTEM DEVELOPERS

MAPS

Stan Benjamin, Project Leader
(303) 497-6387

Patty Miller, Surface Objective Analysis and Quality Control
(303) 497-6365

Zaitao Pan, Physical Parameterizations
(303) 497-6062

James Ramer, AIVs
(303) 938-2072

Tracy Smith, Observational Ingest and Evaluation
(303) 497-6727

Eta

Adrian Marroquin, Project Leader
(303) 497-6202

Brian Jamison, AIVs and Quality Control
(303) 497-6079

Jennifer Mahoney, AIVs and Quality Control
(303) 497-6514

LAPS

John McGinley, Project Leader
(303) 497-6161

Steve Albers, Cloud Analyses
(303) 497-6057

Dan Birkenheuer, Upper-Air Moisture Analyses
(303) 497-5584

Craig Hartsough, Quality Control
(303) 497-6882

Pete Stamus, Surface Analyses
(303) 497-6100

RAMS

John McGinley, Project Leader
(303) 497-6161

Paula McCaslin, Programmer, 3-D Visualization
(303) 938-2077

Jerry Schmidt, Model Evaluation and Diagnostics
(303) 497-6098

GLOSSARY

ARINC	Aeronautical Radio INCorporated
ACARS	ARINC Communications Addressing and Reporting System
AGFS	Aviation Gridded Forecast System
AGL	Above ground level
AIV	Aviation-impact variable
ASOS	Automated Surface Observing System
AWOS	Automated Weather Observing System
BKN	Broken (cloud amount)
BT	Brightness temperature
CLR	Clear (cloud amount)
CONUS	Continental United States
CSI	Critical success index
CSU	Colorado State University
FAA	Federal Aviation Administration
FAR	False alarm rate
FSL	Forecast Systems Laboratory
FT	Terminal Forecasts
GLA	Goddard Laboratory for Atmospheres
HSS	Heidke Skill Score
IFR	Instrument Flight Rules
LIFR	Low Instrument Flight Rules
LAPS	Local Analysis and Prediction System
MAE	Mean absolute error
MAPS	Mesoscale Analysis and Prediction System
ME	Mean (algebraic) Error
MOS	Model output statistics
MVFR	Marginal Visual Flight Rules
NCDC	National Climatic Data Center
NGM	Nested Grid Model
NMC	National Meteorological Center
NWS	National Weather Service
OVC	Overcast (cloud amount)
OBSCD	Obscured (cloud amount)
POD	Probability of Detection
PRF	Vertical wind profiler
RAMS	Regional Atmospheric Modeling System
RAOB	Rawinsonde observation
RMS	Root mean square
RMSE	Root mean square error
RMSVE	Root mean square vector error
SAV	State-of-the-atmosphere variable

SAO	Surface aviation observation
SCT	Scattered (cloud amount)
TSS	True Skill Statistic
UPA	Upper-air rawinsonde
VFR	Visual Flight Rules
WFO	Weather Forecast Office
WSFO	Weather Service Forecast Office
WSO	Weather Service Office

*U.S. GOVERNMENT PRINTING OFFICE: 1993-774-025/89012