

(NASA-TM-84995) EVALUATION OF THE SYNOPTIC AND MESOSCALE PREDICTIVE CAPABILITIES OF A MESOSCALE ATMOSPHERIC SIMULATION SYSTEM (NASA) 117 p HC A06/MF A01

CSCX 04B

N83-25317

Unclas 11565

G3/47

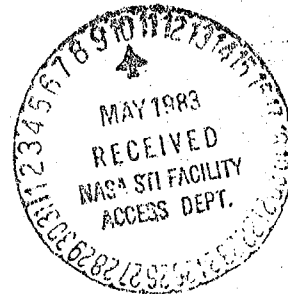


Technical Memorandum 84995

EVALUATION OF THE SYNOPTIC AND MESOSCALE PREDICTIVE CAPABILITIES OF A MESOSCALE ATMOSPHERIC SIMULATION SYSTEM

S. E. Koch, W. C. Skillman, P. J. Kocin, P. J. Wetzel, K. Brill, D. A. Keyser, and M. C. McCumber

MARCH 1983



National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

LIBRARY

NOV 21 1983

LANGLEY RES.
LIBRARY,
HAMPTON, VIRGINIA

TM84995

EVALUATION OF THE SYNOPTIC AND MESOSCALE PREDICTIVE CAPABILITIES
OF A MESOSCALE ATMOSPHERIC SIMULATION SYSTEM

by

Steven E. Koch, William C. Skillman, Paul J. Kocin, Peter J. Wetzel
Goddard Laboratory for Atmospheric Sciences
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

and

Keith Brill, Dennis A. Keyser
General Software Corporation
Landover, MD 20785

and

Michael C. McCumber
Universities Space Research Association
Columbia, MD 21044

March 1983

1183-25317 #

EXECUTIVE SUMMARY

The overall performance characteristics of a limited area, hydrostatic, fine (52 km) mesh, primitive equation, numerical weather prediction model are determined in anticipation of future satellite data assimilations with the model. The synoptic and mesoscale predictive capabilities of version 2.0 of this model, the Mesoscale Atmospheric Simulation System (MASS 2.0), were evaluated at NASA/GLAS, which has provided the support for MASS development by Systems and Applied Sciences Corporation. The authors of this report comprised the independent group of model evaluators. The two-part study is based on a sample of approximately thirty 12h and 24h forecasts of atmospheric flow patterns over the United States during spring and early summer of 1982.

The *synoptic-scale* evaluation results (Part I) benchmark the performance of MASS 2.0 against that of the National Meteorological Center's operational, synoptic-scale weather prediction model, the Limited area Fine Mesh (LFM). The large sample allows for the calculation of statistically significant measures of forecast accuracy and the determination of systematic model errors. The synoptic-scale benchmark is required before unsmoothed mesoscale forecast fields can be seriously considered. The *mesoscale* forecast fields cannot, of course, be objectively verified with routinely collected synoptic data. Instead, the mesoscale fields are (a) subjectively "verified" (in Part II) in a diagnostic sense against a large sample of observed mesoscale convective systems, and (b) used in a real-time experiment in late June to forecast "convective outlook" areas.

The results of the synoptic-scale evaluation demonstrate that MASS 2.0 predicts the lower tropospheric mass fields significantly better than does the

LFM model, but that the LFM does significantly better in the upper troposphere. The greatest improvement by MASS 2.0 over the LFM model forecasts is in the thickness field over the western U.S., where the LFM showed a systematic cold bias in this sample. It is suggested that the higher resolution and improved planetary boundary layer parameterization in the mesoscale model resulted in less eastward phase error of western troughs and more realistic lower tropospheric heating distributions, respectively.

The most important systematic forecast error in MASS 2.0 appears under a specific kind of synoptic flow pattern, in which serious loss of mass occurs throughout the model atmosphere over the eastern United States. It appears that eastern boundary conditions may have been the source for this problem. When such cases are deleted from the whole sample, MASS 2.0 either equals or exceeds LFM performance throughout virtually the entire troposphere. Other significant systematic errors exist in the 24h MASS 2.0 forecasts. These include overforecasts of thickness values and surface cyclone intensity over the Plains states. These *synoptic-scale* errors can be traced to an erroneous CISK-like (Conditional Instability of the Second Kind) process that occurs at the *mesoscale* as a result of the omission of a cumulus parameterization scheme in this version of the model. The last important model error studied is low moisture/instability levels forecast over the southern U.S. This bias seems mostly related to problems in the model moisture initialization.

The mesoscale study demonstrates that MASS 2.0 produces coherent vertical motion fields that offer a significant improvement over the LFM forecasts, in terms of information content and displacement errors. The MASS 2.0 vertical motion fields are combined with other mesoscale forecast fields to produce "predictor variable" fields. These fields are "accurately" related in a

diagnostic sense to the locus of about 50% of 149 strong mesoscale convective systems (MCS) which were observed during the 3-month duration of the experiment. The verification criteria used were 3h/250 km temporal/spatial forecast accuracy. The false alarm rate was very low (about 14%). Thus, the chief problem was *underforecasting*, the leading causes for which were (a) improperly forecast "short wave" disturbances aloft, (b) underforecasts of frontal convergence intensity (mostly related to underforecast surface high pressure intensity resulting from the eastern boundary condition problem), and (c) underforecasts of potential instability (mostly related to the moisture initialization problem). These problems are identifiably situation-dependent.

The limited real-time experiment showed that a forecaster could reduce the size of his convective outlook area by 42% when using MASS 2.0 information in place of the LFM, without suffering a major reduction in the ability to detect ("hit") score. He thereby achieved a modest reduction in false alarms and a 15% increase in forecast skill. MASS 2.0 provided the forecaster with useful information on the depth, strength, and continuity of fields indicative of possible strong convection, and thus helped refine his threshold for action.

These results indicate that the existing version of the model can be used to gain insight into mesoscale processes and to assess the impact of satellite data upon the mesoscale in those cases where the effects of the noted systematic errors at the synoptic scale are minimal. Further model development to remedy these systematic errors is recommended before satellite data insertion in general is attempted.

TABLE OF CONTENTS

	Page
EXECUTIVE SUMMARY	i
FIGURES	v
TABLES	x
SECTION	
1. INTRODUCTION	1
2. DESCRIPTION OF THE MODEL	4
PART I: SYNOPTIC SCALE MODEL EVALUATION	10
3. METHODOLOGY	10
4. RESULTS OF STATISTICAL EVALUATION OF MODEL SYNOPTIC-SCALE PERFORMANCE	20
5. SYSTEMATIC MODEL ERRORS OBSERVED AT SYNOPTIC SCALE	30
6. SOURCES OF IMPORTANT MASS 2.0 SYSTEMATIC ERRORS	40
7. SUMMARY OF PART I: SYNOPTIC-SCALE MODEL EVALUATION RESULTS	55
PART II: EVALUATION OF MASS 2.0 MESOSCALE PREDICTABILITY.	58
8. COHERENT MESOSCALE INFORMATION PROVIDED BY MASS 2.0 IN TWO CASE EXAMPLES	60
9. GENERAL MASS 2.0 MESOSCALE PREDICTIVE CAPABILITIES	74
10. SUMMARY OF PART II: MESOSCALE MODEL EVALUATION RESULTS	95
ACKNOWLEDGEMENTS	99
REFERENCES	100

FIGURES

- Fig. 1: Domain of integration used in MASS 2.0 and smaller region of actual forecast verification. Grid size shown is 52 km true at 90°N in a polar stereographic projection. (p.12)
- Fig. 2: Distribution of eight large boxes and smaller (500 km x 500 km) component boxes used in construction of an error "histo-map". Each large box is composed of nine smaller boxes in which the raw data is entered. Data from alternate small boxes supplies information to two (four) surrounding large boxes in the top and bottom (middle) row for overlap purposes. (p.19)
- Fig. 3: Synoptic-scale forecast skill score comparisons between MASS 2.0 and LFM models based upon entire (nearly 30) case sample. (p.22)
- Fig. 4: Daily LFM-MASS 2.0 24h forecast difference statistics. Positive value represents favorable MASS 2.0 skill score compared to LFM. (p.24)
- Fig. 5: Daily LFM-MASS 2.0 normalized and pressure parameter-averaged 24h forecast difference statistics. (p.26)
- Fig. 6: Fields of 500 mb geopotential and sea level pressure on May 5, 1982. Left figures show initial (1200 GMT) fields, right figures show 2h MASS 2.0 forecast (1400 GMT) fields. (p.27)
- Fig. 7: Synoptic-scale forecast skill score comparisons between MASS 2.0 and LFM models based upon the sample that excludes "BC regime" days. (p.29)

Fig. 8: Error histo-map for 12h anticyclone forecasts. Within each large box (See Fig. 2), histogram plots of amplitude and phase errors and total number of "observed, but no forecast centers" (NFC) and "forecast, but no observed centers" (NOC) are shown. Error interval on amplitude (phase) error histogram is 2 mb (45°). Maximum event frequency on amplitude (phase) error histogram is 8(4). On phase histogram, two bar sizes are used, with the wider bar representing a forecast displacement error larger than 300 km.

Fig. 9: The eight regions within which subjectively determined forecast errors are tabulated.

Fig. 10: 24h forecast and verification analyses valid at 1200 GMT 9 June 1982: (p.43)
(1) 1000-500 mb thickness analysis (heavy lines in m) and thickness "difference map" (light solid (forecast greater than observed values) and dashed (forecast less than observed values) lines, in m), (ΔZ and ΔZ_c , respectively); (b) forecast surface pressure field (P_{SFC}^F) and diagnosed frontal system (solid), and observed locations of surface low pressure systems (P_{SFC}^O) and frontal system (dashed), in mb (07 = 1007 mb); (c) forecast 500 mb geopotential (Z_{500}) (solid, in m) and vorticity (ζ_{500}) (dashed, $10^{-5} s^{-1}$) fields; and (d) forecast of 24h accumulated stable precipitation (P_{ACC}), in mm. Dashed line traces past movement of maxima from its origin at position "X".

Fig. 11: 24h 500 mb geopotential and vorticity (a) verification analysis, (p.45)
(b) filtered MASS 2.0 forecast, and (c) LFM forecast valid for 1200 GMT 9 June 1982. "Short waves" in the vorticity fields are

labelled numerically. Reporting rawinsonde station locations shown in (a). Isolines in same format as in Fig. 10c.

Fig. 12: MASS 2.0 unfiltered forecasts of 500 mb geopotential and vorticity (p.46) (as in Fig. 11) at 1800 GMT 8 June 1982 (6h), 0000 GMT 9 June 1982 (12 h), and 0600 GMT 9 June 1982 (18 h).

Fig. 13: MASS 2.0 unfiltered forecasts of (a) planetary boundary layer depth (p.48) (mb) at 12h (H^{12}); (b) 700 mb omega vertical motion field ($\mu\text{bar s}^{-1}$) at 12 h (ω_{700}), (dashed = upward); and (c) 300 mb wind vectors (m s^{-1}) and divergence (solid positive, dashed negative, intervals of 10^{-5} s^{-1}) at 18h (V_{300} , D_{300}) produced from initial conditions at 1200 GMT 8 June 1982.

Fig. 14: Comparison of 12h forecast of 700 mb vertical motion (upward, dashed (p.62) at intervals of $2 \mu\text{bar s}^{-1}$) by the LFM (left) and MASS 2.0 (center) models. Also shown is verifying GOES-E visible satellite image. Model forecasts verify at 0000 GMT 3 April 1982, satellite image is for 2230 GMT 2 April 1982.

Fig. 15: Comparison of 12h LFM and filtered MASS 2.0 forecasts of 500 mb (p.65) geopotential height (solid lines, CI = 60m) and absolute vorticity (dashed lines, CI = $2 \times 10^{-5} \text{ s}^{-1}$) from initial analysis at 1200 GMT 14 April 1982. Also shown is verification analysis at 0000 GMT 15 April 1982. Box over center portion encloses sub-synoptic area of in-depth study in succeeding figures.

Fig. 16: Sequence of unfiltered MASS 2.0 forecasts of 500 mb geopotential (p.66) height and vorticity on 14 April 1982. Same format as in Fig. 15.

- Fig. 17: Sequence of MASS 2.0 700 mb vertical motion fields and NAFAX Automated Radar Summaries on 14 April 1982. Vertical motions contoured at intervals of $2 \mu\text{bar s}^{-1}$ (dashed = upward). Radar echo tops (underlined) reported in $\text{ft} \times 10^2$. (p.68)
- Fig. 18: Comparison of 12h forecast of 700 mb vertical motion (upward, dashed at intervals of $2 \mu\text{bar s}^{-1}$) by the LFM (left) and MASS 2.0 (right) models, verifying at 0000 GMT 15 April 1982. (p.69)
- Fig. 19: Initial (top) and 24h (bottom) forecasts of 850 mb temperature (left, dashed), geopotential height (left, solid), and wind vectors and isotachs (right, solid). Also shown are two-hourly locations of maximum surface pressure falls, with circled numbers denoting time in GMT. Isotherm interval is 2.5C , geopotential height interval is 30 m , and isotach interval is 5 m s^{-1} . (p.71)
- Fig. 20: Model 12h forecasts verifying at 0000 GMT 15 April 1982 of (a) planetary boundary layer height (in m), (b) dew point temperature ($^{\circ}\text{C}$) at the lowest level ($\sigma = 0.96$) in the model ("surface"), and (c) "surface" wind divergence (10^{-5} s^{-1}) with convergence dashed. (p.73)
- Fig. 21: Constituent fields of model convective predictor variable field WLI and comparison with observed MCS loci at 0200 GMT 15 April 1982: (a) lifted index forecast, (b) 700 mb vertical motion (upwards is dashed, in $\mu\text{bar s}^{-1}$), (c) WLI field produced from overlap of 700 mb $\omega \geq -2 \mu\text{bar s}^{-1}$ and lifted index ≤ 0 areas, with numbered loci of WLI and corresponding MCS (forecast spatial offset shown by (p.79)

line segments), and (d) 0135 GMT radar summary (storm tops in 10^2 ft) showing distribution of the three MCS's.

Fig. 22: LFM- and MASS 2.0-based convective outlooks produced by experimental forecaster (dashed and solid boxes, respectively) on the six days of real-time forecast experiment. Irregularly-shaped MASS 2.0 based areas are the result of combining areas of separate severe weather watch boxes. Also shown are verifying severe weather reports, where H = large hail, T = tornado, W = damaging windstorm, and numbers designate number of such reports within a cluster. Position of locus of convective predictor variable (WLI, FCE, or WA) is shown at 2h intervals by stars (arrows denote increasing time).

TABLES

- Table 1. Comparisons between the MASS 2.0 and LFM model systems.
(p. 8)
- Table 2. Forecast fields subjected to verification.
(p. 13)
- Table 3. Definitions of model evaluation statistics.
(p. 14)
- Table 4. Average 24h LFM and MASS 2.0 forecast skill statistics.
(p. 20)
- Table 5. Distribution of absolute pressure error (mb) for surface cyclones
(p. 21)
predicted at 24h by LFM and MASS 2.0 models.
- Table 6. Number of significant¹ sea level pressure centers missed in 12h and
(p. 30)
24h forecasts by the MASS 2.0 and LFM models.
- Table 7. Summary of regional variations in model systematic errors in fore-
(p. 34)
casts of sea level pressure features and of 500 mb geopotential and
1000-500 mb thickness fields in close proximity to the observed
position of the pressure features (results from error histo-map
analysis). Undiagnosed, uncertain, or small net errors denoted
by "?"
- Table 8. Regional variations in average model amplitude and phase (in
(p. 37)
parentheses) errors based upon subjective analysis of "difference
maps" (forecast field minus observed field)^{1,2} for 500 mb geo-
potential and vorticity fields. Mean geopotential (vorticity)
amplitude errors smaller than $30 \text{ m} (2 \times 10^{-5} \text{ s}^{-1})$ are not shown.
- Table 9. Same as for Table 8, except for 1000-500 mb geopotential and thick-
(p. 38)
ness and mean relative humidity fields.^{1,2} Mean thickness (relative
humidity) errors smaller than 30 m (20%) are not shown.

Table 10. Features of various fields forecast by MASS 2.0 on days when significant positive errors in 24 hr thickness field forecasts occurred in close juxtaposition to significant overforecast cyclone intensity errors.

Table 11. Summary of diagnostic MCS locus study.
(p. 80)

Table 12. Leading causes for missed forecasts of MCS's.
(p. 82)

Table 13. Temporal variation of MCS forecast evaluation statistics.
(p. 85)

Table 14. Statistical skill scores for LFM- and MASS 2.0-based experimental forecaster (BURGESS) convective outlooks and for SELS 0900 GMT convective outlooks made during real-time experiment 21 June - 29 June 1982.

EVALUATION OF THE SYNOPTIC AND MESOSCALE PREDICTIVE CAPABILITIES
OF A MESOSCALE ATMOSPHERIC SIMULATION SYSTEM

1. INTRODUCTION

It is well established that nonlinear processes in mesoscale numerical weather prediction models can generate realistic mesoscale features from initial conditions obtained only from a larger synoptic scale (Perkey, 1976; Warner *et al.*, 1978; Anthes and Keyser, 1979; Maddox *et al.*, 1981; Anthes *et al.*, 1982; Chang *et al.*, 1982; Kaplan, *et al.*, 1982). As an explanation for such success, Anthes and Warner (1978) have suggested that the smaller the scale of the initial state phenomenon, the more the atmosphere "forgets" that initial state and responds instead to the model's nonlinear dynamics and local forcing. There have, however, been some attempts very recently to incorporate meso-alpha scale (Orlanski, 1975) data into the initial model state, with very encouraging results (Tarbell *et al.*, 1981; Kaplan *et al.*, 1982). There is a growing awareness that the initial data fields for mesoscale models should have a horizontal resolution of at least 30 km to properly resolve small-scale systems (National Research Council, 1980), which may soon be available from both satellite and ground-based remote sensors. Yet, it is highly desirable, first of all, to document the performance characteristics of the model at both the synoptic- and meso-scale prior to such data assimilation experiments.

Most reported studies of mesoscale prediction model capabilities have relied upon either (a) in-depth comparisons of model forecast fields with observed verification data or a case study basis, or (b) have shown how the model can realistically simulate generally known aspects of such mesoscale features as mountain-valley breezes, sea breezes, fronts, mesoscale convective complexes, and mountain waves. Rarely have comprehensive studies of overall

model performance based upon a large population been conducted. Anthes and Keyser (1979) reported on the statistical forecast skill scores of the Pennsylvania State University mesoscale model (using a 60 km mesh) and rated the PENN model performance relative to that of the large-scale, operational Fleet Numerical Weather Central model (Kesel and Winninghoff, 1972), based on a 32-case sample. However, they did not report on the regional variation in systematic model errors, so that the nature of the model errors could be better understood. Furthermore, the mesoscale vertical motion fields were not exhaustively examined for degree of coherence.

In the present work, the regionally-varying systematic errors, the coherence of the mesoscale fields, and statistically significant forecast skill scores of the smoothed synoptic fields forecast by a mesoscale modelling system known as the Mesoscale Atmospheric Simulation System (MASS, version 2.0) are described. The general structure of the system has been discussed, with an example of mesoscale forecasts for a severe storm outbreak, by Kaplan *et al.* (1982). The system has been under development at the Systems and Applied Sciences Corporation, with support from the NASA/Goddard Laboratory for Atmospheric Sciences (GLAS). The evaluation was performed at GLAS, and therefore was independent of the modelling development group. It is the objective of this research to evaluate the synoptic and mesoscale predictive capabilities of the model in terms of approximately thirty 12h and 24h forecasts of atmospheric flow patterns over the United States during spring and early summer of 1982. The synoptic-scale performance is benchmarked against that of the National Meteorological Center's large-scale operational Limited Fine Mesh (LFM) model (Gerrity, 1977; Newell and Deaven, 1981). Brief comparisons between the LFM's performance during this experiment and that

reported on by Silberberg and Bosart (1982) from a much larger sample will also be made, in order to see whether our sample is representative of LFM performance in general.

The synoptic-scale evaluation results are given in Part I of this paper. These results serve as a benchmark by which the unsmoothed MASS 2.0 mesoscale field forecasts can be interpreted and seriously considered. Part II consists of (a) a subjective "verification" of the unsmoothed mesoscale forecast fields, by using such fields as "convective predictor variables" which are verified against a large sample of observed mesoscale convective systems, and (b) the results of a limited real-time convective forecast experiment using model derived variable fields to forecast "convective outlook" and "severe weather watch" areas. Sole reliance upon a subjective verification for the mesoscale fields is naturally the consequence of not having a mesoscale observed database by which the forecasts can be objectively verified. Thus, our approach differs from that of Anthes and Keyser (1979) in this regard, as they attempted to verify the unsmoothed mesoscale forecasts of the PENN model against synoptic observations. The consequence of their approach was that predicted field gradients were stronger than could be observed, which reflected to an unknown degree upon the final statistical results.

We note that it is not the prime objective of this study to use the model's forecast fields to further our understanding of those mesoscale processes that are important to the development of severe convection. However, several mesoscale phenomena resolved by MASS 2.0 (mountain waves, dryline bulge, low-level jet, etc.) will be illustrated. Moreover, these results can serve as a basis for future, more exhaustive case studies of individual storm systems.

2. DESCRIPTION OF THE MODEL

A complete description of MASS is given by Kaplan *et al.*, (1982). The most important model characteristics are summarized here and contrasted with those of the LFM. In brief, MASS utilizes higher order accurate finite difference approximations, a mesoscale mesh with nesting capabilities, high vertical resolution, the diabatic primitive equations in hydrostatic form, and a planetary boundary-layer (PBL) parameterization based on surface-outer layer similarity theory matching. Version 2.0 lacks a cumulus parameterization scheme, and differs in a few other respects from the general discussion given in Kaplan, *et al.*, (1982), as will be noted. MASS 2.0 utilizes the high computation speed of the CDC CYBER 203 vector processor machine to produce a 24h simulation in 30 min CPU time.

a. *Model Numerics, Boundary Conditions, and Initialization*

MASS 2.0 uses 6th (2nd) order accurate finite difference formulae to approximate the horizontal (vertical) derivatives. A diffusion filter is applied to remove "noise" at scales of twice the grid spacing. The Matsuno (1966) time-marching technique is used. The grid spacing is 52 km true at 90° north latitude on a polar stereographic projection. The model domain comprises a 142 x 106 matrix, which covers the larger area depicted in Fig. 1. For forecast comparisons with the LFM model, a subset of this area covering mainly the data-rich United States region is used (the interior rectangle in the figure). The model has 14 sigma-p levels.

The boundary conditions are specified from the LFM forecast fields. A linear interpolation from the 12- and 24-hour LFM forecast fields determines values at the mesoscale model boundaries at intermediate times. A boundary

"sponge" zone is employed whereby weighting of the LFM boundary tendency values and MASS 2.0 predicted tendency values is a function of distance away from the boundaries.

The initial data for a model simulation is obtained via telephone link from NASA/Langley Research Center (LaRC) to the Water and Power Resources Service (WPRS) data base in Denver, CO. This data set is the LFM Initial Data Matrix created at Suitland, Maryland on an operational basis, which consists of mandatory level rawinsonde and remotely sensed sounding data. This matrix is then modified using a Barnes (1964) objective analysis technique applied to additional surface and significant level rawinsonde data on pressure surfaces spaced 25 mb apart. Next, the merged data set is vertically interpolated to the 14 sigma surfaces. A static initialization is then applied. Finally, a cubic spline is used to interpolate the data horizontally from the LFM 190 km mesh to the 52 km mesh.

The static initialization technique used is a variational scheme that constrains the resulting integrated mass divergence to be zero in a least squares sense (Sasaki, 1958). In effect, the external gravity wave mode is suppressed, at least initially. However, remaining imbalances between the mass and momentum fields can generate internal gravity waves during model integration, although they are somewhat attenuated by the Matsuno scheme. The nonlinear normal mode initialization scheme mentioned in Kaplan, *et al.* (1982) was not operating by the start of the evaluation experiment, so it was not utilized. Proper specification of the initial divergent wind component in mesoscale models is still a very complex and controversial problem (Tarbell *et al.*, 1981).

The terrain used in MASS 2.0 is currently that also used in the LFM, rather than the mesoscale terrain base mentioned in Kaplan, *et al.* (1982).

b. *Parameterization of Physical Processes*

The LFM and MASS 2.0 both recognize the presence of water either as vapor or liquid, although the condensate is either evaporated or forced to precipitate immediately. If the relative humidity in a given layer exceeds 95 (75) percent in the MASS 2.0 (LFM) model, the excess mixing ratio is converted to precipitation. The quantitative precipitation forecast by MASS 2.0 at any point on the ground is equal to that part of the condensed moisture from the lowest 10 layers that reaches the ground. Diabatic latent heating and evaporation effects associated with stable precipitation processes are considered by both models.

The two models define convective precipitation and accommodate the effects of convection differently. The LFM defines as convective precipitation a local increase in precipitable water if such an increase occurs in a conditionally unstable layer whose relative humidity exceeds 75 percent. In the mesoscale model, parcels which rise to their lifting condensation level are then lifted pseudoadiabatically until their buoyancy vanishes. The LFM reduces the conditional instability of the atmosphere where convective precipitation is forecast by producing diabatic heating (cooling) effects resulting from phase changes; also, the model redistributes momentum vertically at such points. *MASS 2.0 does not specify (or parameterize) the vertical distribution of convective heating from known properties of cumulus clouds. Neither can it properly stabilize the convective atmosphere because it neglects the effects of evaporationally-caused downdrafts on the environment.* This is contrasted with the PENN model, which specifies the heating function from a one-dimensional cloud model averaged over the characteristic lifetime of the cumulus cloud, and

which also specifies the size spectrum of clouds (Anthes and Warner, 1978). More elaborate schemes have recently been developed to attempt to model the evolution of mesoscale convective systems by including the effects of fractional entrainment, vertical wind shear, moist downdrafts, freezing and melting, and momentum transport in cold outflows produced by such systems (Fritsch and Chappell, 1980). Research is presently being conducted into ways to incorporate the Fritsch and Chappell scheme into a future version of MASS. It will be seen that omission of a cumulus parameterization scheme in MASS 2.0 had observable consequences at the synoptic scale.

The LFM and MASS 2.0 models treat the transfer of heat by radiation in a very similar fashion. Also, a surface energy budget approach common to many mesoscale models (Anthes and Warner, 1978) is used by both models. No evaporation of soil moisture was permitted in version 2.0 of MASS, in contrast to Kaplan, et al., (1982), in which ground wetness was simply set at a constant value.

One very critical difference between the physical parameterizations used in the two models lies in the way the PBL is treated. The LFM resolves the PBL with only two levels, and fixes its depth at 50 mb. Simple bulk parameterizations are employed, wherein the surface stress is estimated from drag laws. MASS 2.0 employs a variable-depth PBL that is better resolved in the vertical ($\sigma = 1.0, 0.96, 0.89$ in MASS 2.0 vs. $\sigma = 1.0, 0.95, 0.72$, in the LFM). MASS 2.0 physics are parameterized by use of a generalized similarity theory treatment of heat, moisture, and momentum fluxes (Deardorff, 1972). The effects of stratification and surface layer-PBL coupling are accounted for. The value of such an integrated parameterization approach has been stressed in the review paper by Driedonks and Tennekes (1981). Anthes and Keyser (1979)

demonstrated the importance of using a higher-resolution PBL scheme in models.

A summary of the major differences between the two models appears in Table 1. Referral to the table will help in interpreting the results to be

TABLE 1. COMPARISONS BETWEEN THE MASS 2.0 AND LFM MODEL SYSTEMS.

ASPECT	LFM	MASS 2.0
<u>Model Numerics</u>		
Coordinate system	7 sigma-p levels	14 sigma-p levels
Grid and array sizes	190 km (53x45)	52 km (142x106) at 90°N
Accuracy of horizontal differencing	4th order	6th order
Time integration method	Smoothed leapfrog	Matsuno (Euler-backward)
<u>Model Initialization</u>		
Data base:	Mandatory pressure data (LFM Initial Data Matrix)	Mandatory and significant pressure and surface data
Hydrostatic-superadiabatic checks	Dry convective adjustment	Dry convective adjustment
Wind initialization	Sum of PE-6 forecast divergent and analyzed initial rotational wind components,	Static initialization to remove external gravity waves initially (Sasaki, 1958) and use of Matsuno scheme at each time step
<u>Lateral Boundary Conditions</u>	Time-dependent tendencies obtained from previous global spectral 12h forecast using boundary zone "sponge"	Time-dependent tendencies obtained from current LFM 12-24h forecast using boundary zone "sponge"
<u>Post-Processing</u>	Sigma-to-pressure interpolation, "desloshing" of mass fields, and spatial filtering	Sigma-to-pressure interpolation, Shuman filter to LFM scale and cubic spline from 52 km to 190 km grid for objective evaluation
<u>Physical Parameterizations</u>		
Convective Precipitation	Conditional instability is reduced and momentum is redistributed where convective precip is forecast	Precip may form, but no cumulus parameterization scheme is included (thus improper, if any, stabilization)

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 1. (Continued)

ASPECT	LFM	MASS 2.0
Radiative heat transfer	Surface energy budget Cloud effects incl. Land-ocean albedo difference. Sensitive to water vapor distribution	Essentially same as LFM
Planetary Boundary Layer	One fixed layer depth (50 mb) Use of bulk para- meterization and surface drag formulae No surface evaporation over land Sea surface sensible, latent heat fluxes PBL fluxes coupled to free atmosphere only through dry and moist convection treatments	Variable depth with predictive equations for h for stable and un- stable cases Generalized similarity theory with surface layer matching No surface evaporation over land Turbulent heat, moisture, and momentum fluxes explicitly treated allow PBL-free atmosphere coupling

presented. Notice the "post-processing" aspect in particular. Model predicted fields on sigma surfaces must be vertically interpolated to either isobaric surfaces or to the ground for model evaluation. The "build-down" procedure presents some problems in interpretation. For example, the exact way that lifted index is both defined and calculated may differ slightly between models, although much effort was made to minimize these differences. The examined fields most sensitive to build-down problems are relative humidity, lifted index, and sea level pressure. Also notice that no "desloshing" of MASS 2.0 predicted mass fields was done; however, it will be seen that this omission did not cast an unfair light upon the performance results of the mesoscale model. The filtering procedure is described in section 3b.

PART I: SYNOPTIC-SCALE MODEL EVALUATION

3. METHODOLOGY

a. Case Requirements

The experimental period started April 2, 1982, and ended on July 2, 1982. Thus, a variety of spring and early summer systems was studied. MASS 2.0 24h simulations were attempted every weekday except Fridays during this period (the verification data for Friday runs could not be collected beyond 12h). A 24h simulation case experiment was considered totally successful only if all the following requirements were met:

- (1) The 1200 GMT LFM Initial Data Matrix and auxiliary data base were successfully acquired at LaRC from the WPRS link.
- (2) Both the 12h and 24h LFM forecasts and verification data were successfully acquired at LaRC from WPRS.
- (3) MASS 2.0 simulation was successfully run at LaRC for a 24h period and all derived variable fields plotted.
- (4) Inspection of the initial condition state of the primary meteorological fields by the NASA/GLAS evaluation team showed no major data loss, bias, or inconsistencies.

The second criterion created the most frequent problem, as the inexpensive telephone data link with WPRS proved rather unreliable. Although every attempt was made to acquire a sample of 30 cases, in actuality 28 verifications of 24h forecasts and only 23 verifications of 12h forecasts were conducted. The daily model output was typically available from the CYBER computer by early afternoon and could be sent to GLAS within a day or two, but the evaluation effort required more than a day's work per case.

b. *Objective Statistical Evaluation Procedures*

Because the LFM Data Matrix was the sole source for objective verification of both models, the 52 km mesh MASS 2.0 forecast fields needed to be smoothed prior to statistical verification. Following the build-down from sigma to pressure surfaces, MASS 2.0 output was subjected to: (a) multiple (20) passes of the 3 point Shuman smoother-desmoother filter, to eliminate the $4\Delta x$ wave (where Δx is the LFM grid size of 190.5 km), followed by (b) a cubic spline algorithm to interpolate these smoothed grid point values on the 52 km mesh up to the 190.5 km mesh.

A statistical objective evaluation was performed on the nine 12h and 24h forecast fields given in Table 2. The statistics generated were averaged over the interior domain in Fig. 1 and are:

- (1) Root-mean-square error (RMSE), a measure of average model forecast error;
- (2) The (S1) score measure of errors in horizontal gradients (Teweles and Wobus, 1954);
- (3) The difference between the field means of the forecast and observed variables (BIAS); and
- (4) A spatial correlation matrix scheme that determines mean spatial offset biases (Tarbell *et al.*, 1981). The initial state mean is first subtracted to remove large forecast persistence effects, and then the forecast and observed field mean values are subtracted from the respective field values to account for forecast field bias. The resulting correlation matrix better spotlights the spatial maxima in correlation (CORR).

The mathematical definitions for these four statistics appear in Table 3. The relative performance levels of the LFM and MASS 2.0 models are given in terms of "difference statistics" in this paper (LFM statistic minus MASS 2.0 statistic). Thus, a positive difference will mean MASS 2.0 has outperformed

ORIGINAL PAGE IS
OF POOR QUALITY.

the LFM, with the single exception that the opposite is true for the CORR statistic. In order to permit a degree of generalization, each statistic score for each meteorological variable field was first averaged over all (~ 30) cases, and then daily values of each such statistic were normalized by that average. This enables a study of the temporal variation between various pairs of the four statistics (e.g., RMSE vs. S1) to be made. Finally, a

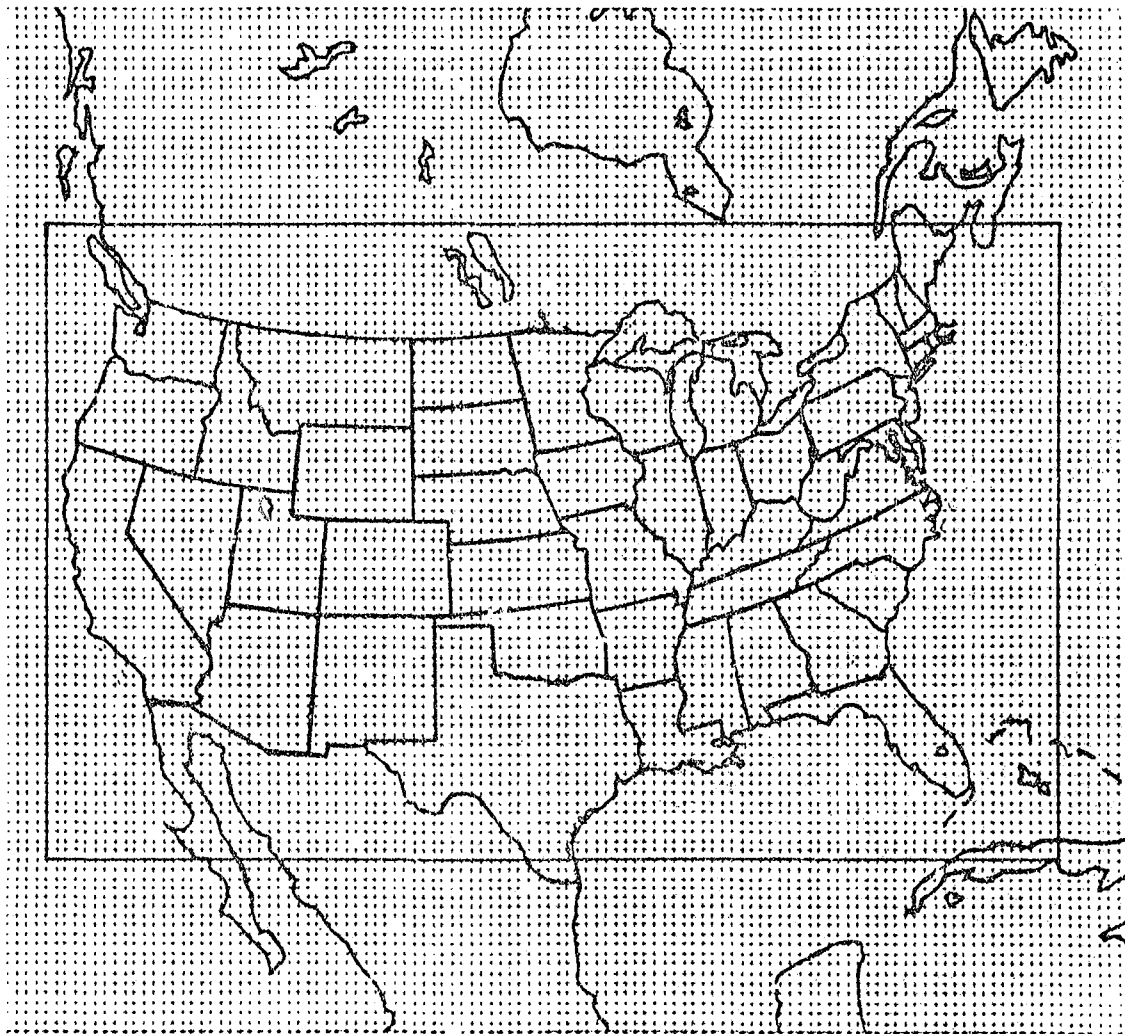


Figure 1. Domain of integration used in MASS 2.0 and smaller region of actual Forecast verification. Grid size shown is 52 km true at 90°N in a polar stereographic projection.

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2. FORECAST FIELDS SUBJECTED TO VERIFICATION.

MASS 2.0 Forecast Field (2-hourly map output)	Subject to Evaluation Team Inspection	Subject to Statistical Verification
500 mb vorticity	yes	no
500 mb geopotential	yes	yes
Sea level pressure	yes	yes
1000-500 mb thickness	yes	yes
SFC-500 mb relative humidity	yes	yes
Lifted index	yes (no LFM)	no
Stable precipitation amount	no	no
850 mb wind vectors	no	yes (speed)
850 mb geopotential	no	yes
850 mb temperature	no	yes
300 mb divergence	no	no
300 mb geopotential	no	yes
300 mb wind vectors	no	yes (speed)
Dewpoint at $\sigma = 0.96$	no	no
Divergence at $\sigma = 0.96$	no	no
Surface temperature	no	no
PBL height	no	no
Pressure tendency	no	no

"pressure parameter-averaged normalized difference statistic" was calculated daily for RMSE and BIAS by averaging each day's normalized difference statistic over the four "pressure" fields of geopotential at the 300, 500, and 850 mb levels and sea level pressure. In this manner, daily normalized model errors seen in a specific statistic like BIAS for four separate mass fields can be simply combined into one value. Variation of the statistics values throughout the course of the experiment are of great interest, because study of temporal variations may reveal transitions in performance levels of the mesoscale model. Recognition of such important transitions can lead the way to more fundamental understanding of model behavior by suggesting subtle systematic errors. Therefore, these results should be used to prioritize future MASS improvements.

TABLE 3. DEFINITIONS OF MODEL EVALUATION STATISTICS.

BASIC STATISTICS	LFM-MASS 2.0 DIFFERENCE STATISTICS	NORMALIZED DIFFERENCE STATISTICS	PRESSURE PARAMETER-AVERAGED NORMALIZED DIFFERENCE STATISTICS
F_i = Forecast Value O_i = Observed Value NG = Number of Grid Points S = Standard Deviation	K = 1, N CASE DAYS L = 1, P PARAMETER FIELDS	$\bar{X}(L) = \frac{\sum_{K=1}^N X(K,L)}{N}$ (avg stat for each parameter)	$\bar{X}(K) = \frac{\sum X^*(K,L)}{M}$ where M = P/2 (pressure (MASS) fields only)
$RMSE = \sqrt{\frac{\sum (F_i - O_i)^2}{NG}}$	$D_R(K,L) = \text{LFM RMSE}(K,L) - \text{MASS RMSE}(K,L)$	$D_R^*(K,L) = \frac{D_R(K,L)}{RMSE(L)}$	$\hat{D}_R(K) = \frac{D_R^*(K,L)}{\bar{D}_R(K)}$
$SI = \frac{100 \sum F_i - O_i }{\sum \max(F_i, O_i)}$	$D_S(K,L) = \text{LFM SI}(K,L) - \text{MASS SI}(K,L)$	$D_S^*(K,L) = D_S(K,L)$	$\hat{D}_S(K) = \frac{D_S^*(K,L)}{D_S(K)}$
CORR = MAX(R_j), where $R_j = \frac{\sum \frac{F_i O_i + j}{NG} - \frac{\sum F_i}{NG} \frac{\sum O_i}{NG}}{S_F S_O}$	$D_C(K,L) = \text{LFM CORR}(K,L) - \text{MASS CORR}(K,L)$	$D_C^*(K,L) = D_C(K,L)$	$\hat{D}_C(K) = \frac{D_C^*(K,L)}{D_C(K)}$
$\text{BIAS} = \bar{F}_i - \bar{O}_i$ $= \frac{\sum F_i}{NG} - \frac{\sum O_i}{NG}$ $= \frac{\sum (F_i - O_i)}{NG}$	$D_B(K,L) = \text{LFM BIAS}(K,L) - \text{MASS BIAS}(K,L)$	$D_B^*(K,L) = \frac{D_B(K,L)}{\text{BIAS}(L)}$	$\hat{D}_B(K) = \frac{D_B^*(K,L)}{\bar{D}_B(K)}$

14

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

The mean of each difference statistic for the entire ~ 30-case experiment was computed. Only those mean difference statistics that are statistically significant at the 99% level or higher will be reported. Significance is determined with the Student's t-test (Panofsky and Brier, 1968). A difference that is significant at the 99% level means the hypothesis that, were an infinite number of forecasts to have been made with each model then the scores would have come out the same (i.e., the differences between the LFM and MASS 2.0 models would be zero), can be rejected at the 99% probability level. The Student's t-score is computed, assuming the population mean difference is zero, from

$$t = \frac{\bar{D}}{s} \sqrt{N-1}, \quad (1)$$

where N is the sample size, D is the mean of the difference statistic scores, and

$$s = \frac{N}{N-1} \sqrt{\frac{\sum D^2}{N} - \bar{D}^2} \quad (2)$$

is the sample standard deviation of the differences.

The actual meaning and utility of these four statistics used to evaluate the model's relative performance should be clarified. The familiar RMSE measures the average error, but unfortunately by itself supplies no information on what part of the error is "systematic" in nature and what portion is random (Wilmott, 1982). The whole field portion of the systematic error is examined with the BIAS, whereas regionally systematic errors are obtained from the subjective analyses described in the next section. Of course, part of the

total error is related to errors in forecast horizontal gradients, as measured by the S1 statistic.

The calculated value of the S1 is very sensitive to the proper specification of the grid spacing used whenever the forecast gradient is uniform and the observed gradient is non-uniform (or vice versa). In particular, the most intense parts of the non-uniform gradients will be underestimated if any grid points are skipped in the S1 formulation. This leads to the result that the calculated S1 is smaller than the true S1 score. The problem remains even though the same formulation is utilized in evaluating the performance of two different models, especially if the models employ different sized grids. We employ a first-order accurate centered finite difference approximation to the horizontal derivative of each variable.

Finally, the CORR value used is the maximum element in the spatial correlation matrix. The observed and forecast fields are shifted until the correlation is maximized at some displacement, and it is the absolute value at that location in the offset matrix that is recorded. CORR is intended to give credit for correct forecasts of the shape and intensity of a scalar variable, but displaced from the observed pattern by some small distance. Wilmott (1982) has found that conventional correlation coefficients can be very misleading, at times being totally unrelated to the size of the difference between forecast and observed variables. Although we account for part of this problem by considering the spatial offset bias, the CORR statistic will be relied upon much less so than the other statistics.

The average relation between the RMSE, BIAS, and S1 scores will also be reported to ascertain which scores give the most consistent impact results. Such relationships have also been studied by Atlas et al. (1982) who find that

the RMSE and S1 scores contradict each other at times. Hence they do not always leave similar impressions about model performance.

c. *Procedures for Determining Systematic Model Errors*

A combination of subjective and quasi-objective approaches to finding systematic model errors was developed for this study. Table 2 lists those six fields for which such errors were determined. The procedure is composed of the following:

- (1) Analysis of time series of the model difference statistics, to search for transitions in model performance levels which can be related to systematically differing model responses to definable synoptic pattern regimes;
- (2) General conclusions drawn from map plots of objectively-determined model amplitude/phase errors, to search for regional variations in systematic model errors, defined objectively within 500 km x 500 km boxes centered over observed cyclone and anticyclone centers;
- (3) General conclusions drawn from map plots of model errors determined from subjective partitioning of error into amplitude and phase errors, irregardless of error location relative to surface pressure centers;
- (4) Synthesis of above results in search of new systematic errors that actually underlie, or are the ultimate cause for, the other multiple errors.

The success of the first approach hinges upon our ability to identify features in the synoptic flow regime common to all cases of poor model performances. Use of the "pressure parameter-averaged normalized difference statistics" makes possible such study.

The second and third approaches both start with the same data base, namely computed "*difference maps*" which show contours of the model (MASS 2.0 or LFM) - predicted variable values minus the observed variable values. The second approach was applied to errors in the forecast fields of 500 mb geopotential,

1000-500 mb thickness, and MSL pressure. This approach is very similar to the method first developed by Leary (1971) and later adapted by Silberberg and Bosart (1982) to the study of systematic LFM2 model errors in forecasts of surface cyclone and related upper air fields. Individual errors in these three fields were tabulated in 500 km x 500 km boxes in which the cyclone was observed at verification time. The average for the entire experiment of all such errors within each box was then entered at the center of the box (see Fig. 2). Groups of nine boxes then supplied information to a 1500 km x 1500 km larger box, with some overlap allowed to provide inter-box continuity. Within each of the eight large boxes the amplitude and phase (displacement) errors are tabulated in histogram form. The total number of observed cyclone centers not forecast and forecast cyclone centers not observed are also computed. Anti-cyclone centers, not studied by Silberberg and Bosart (1982), are also studied by us. Such displays will be referred to as *error "histo-maps"*, and are for the most part objective means of determining the regional variations in model forecast systematic errors.

The third approach is less objective than, but complements, the second approach. In this, the difference maps (which show the distribution of total forecast error) are used to find the amplitude error after subjectively removing the phase error (e.g., in the forecast of a cyclone center or thickness ridge). This method was applied to the 500 mb geopotential, 1000-500 mb thickness, 500 mb absolute vorticity, and mean (sigma layers 1 thru 7) relative humidity field forecasts. This approach has the dual advantages of (1) being able to determine amplitude errors in any field irrespective of whether a surface pressure center is observed, and (2) permitting isolation of displacement errors in such field features as ridges and troughs (such

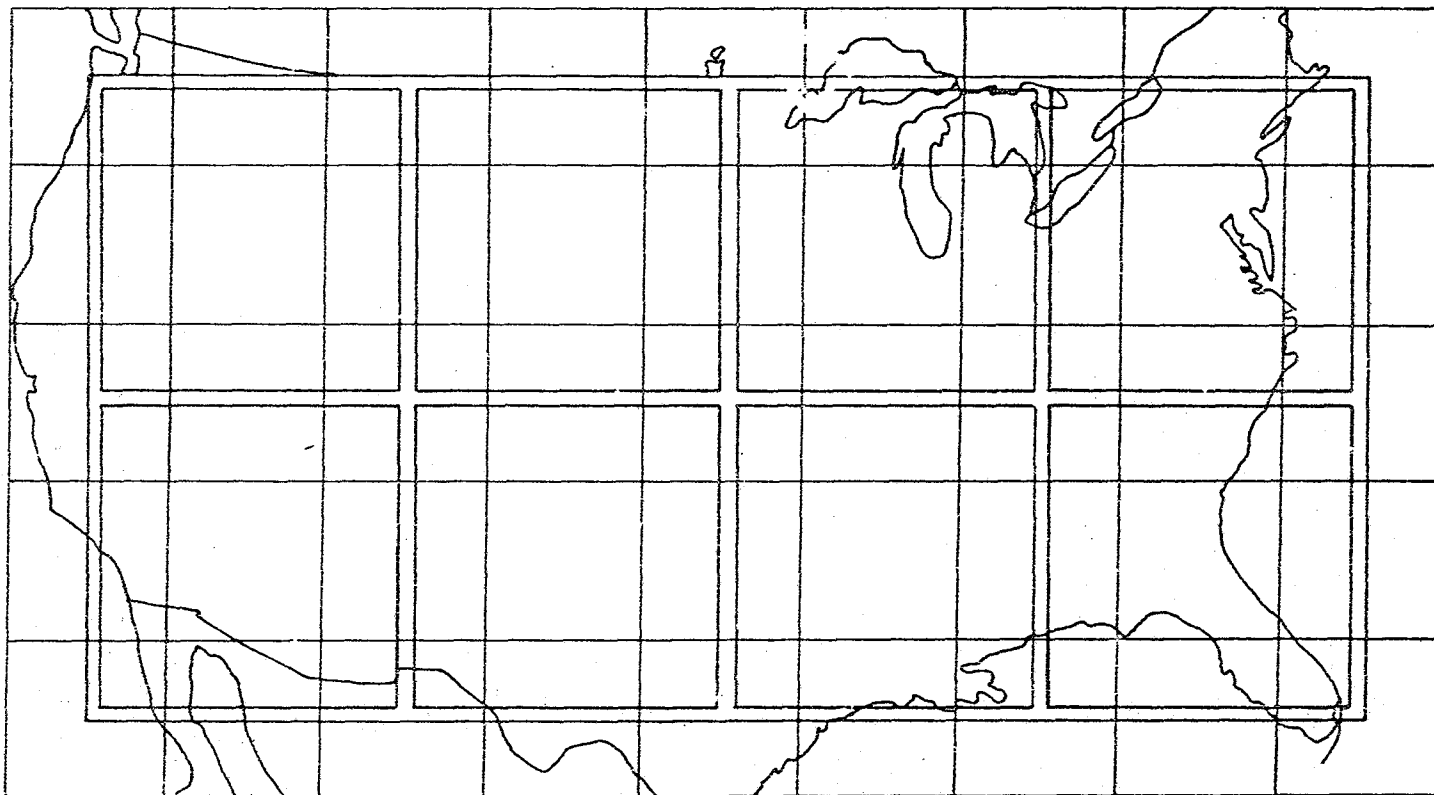
ORIGINAL PAGE IS
OF POOR QUALITY

Figure 2. Distribution of eight large boxes and smaller (500 km x 500 km) component boxes used in construction of an error "histo-map". Each large box is composed of nine smaller boxes in which the raw data is entered. Data from alternate small boxes supplies information to two (four) surrounding large boxes in the top and bottom (middle) row for overlap purposes.

distinction is made only with the surface pressure features in the second approach). These results are then displayed in chart form as a function of geographical region.

The fourth approach was used only to a limited degree in this study, for it was a rather complex and time-consuming endeavor. However, it was used in regard to those systematic model errors that were the most pronounced and frequent, in order that the true underlying causes for such errors could be found.

4. RESULTS OF STATISTICAL EVALUATION OF MODEL SYNOPTIC-SCALE PERFORMANCE

a. Sea Level Pressure and 500 mb Geopotential Forecast Skill Scores

A simple measure of overall model performance is a comparison between (smoothed) MASS 2.0 and LFM average RMSE and S1 scores for 24h forecasts of sea level pressure and 500 mb geopotential (Table 4). The only difference in performance levels between the two models that is statistically significant at

TABLE 4. AVERAGE 24H LFM AND MASS 2.0 FORECAST SKILL STATISTICS.

	Average RMSE		Average S1	
	MASS 2.0	LFM	MASS 2.0	LFM
Sea Level Pressure	3.2 mb	3.0	43.9 mb	43.6
500 mb Geopotential	28.5 gpm	25.2	25.5 gpm	23.9

the 99% probability level or higher (from (1)) is that between the S1 scores for 500 mb geopotential. Table 5 shows that the difference between the two models in terms of both mean and median absolute cyclone pressure errors for 24h forecasts was no larger than 1.0 mb. In general, the MASS 2.0 surface pressure fields are predicted with skill similar to the LFM model forecasts. Thus, the large-scale circulation patterns are predicted about as well by the LFM as by the much finer mesh MASS 2.0 model.

b. *Statistically Significant Model Impacts*

An overall summary of statistically significant differences between the LFM and MASS 2.0 forecast skill scores at both 12h and 24h verification periods is displayed in Fig. 3. Notice that MASS 2.0 outperforms the LFM in forecasts

TABLE 5. DISTRIBUTION OF ABSOLUTE PRESSURE ERROR (MB) FOR SURFACE CYCLONES PREDICTED AT 24H BY LFM AND MASS 2.0 MODELS.

Error interval (mb)	NUMBER OF MODEL CYCLONES	
	LFM (April-July 1982)	MASS 2.0 (April-July 1982)
0-1	17	8
2-3	8	17
4-5	11	4
6-7	0	7
8-9	1	1
10-11	0	0
12-13	0	0
14-15	0	0
over 16	0	0
TOTAL	37	37
Mean absolute error (mb)	2.4	3.2
Median absolute error(mb)	1.3	2.3

ORIGINAL PAGE 19
OF POOR QUALITY

SYNOPTIC-SCALE FORECAST SKILL SCORE
COMPARISONS BETWEEN MASS 2.0 AND LFM MODELS

ENTIRE SAMPLE

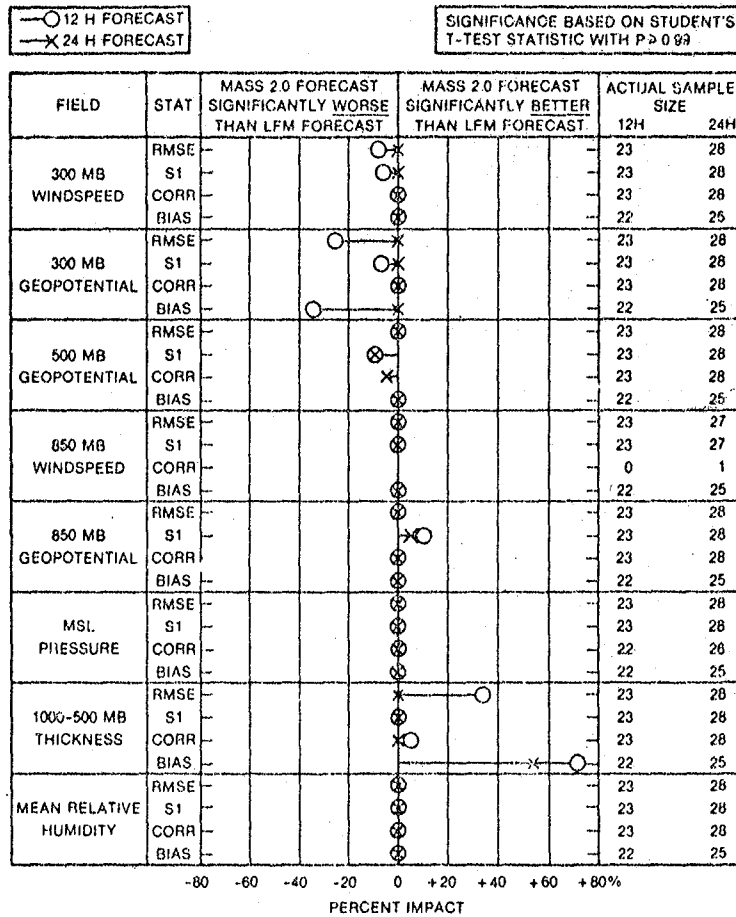


Figure 3. Synoptic-scale forecast skill score comparisons between MASS 2.0 and LFM models based upon entire (nearly 30) case sample.

of the lower tropospheric mass (particularly thickness) fields, but that the LFM is superior in its forecasts of upper tropospheric mass and momentum fields.

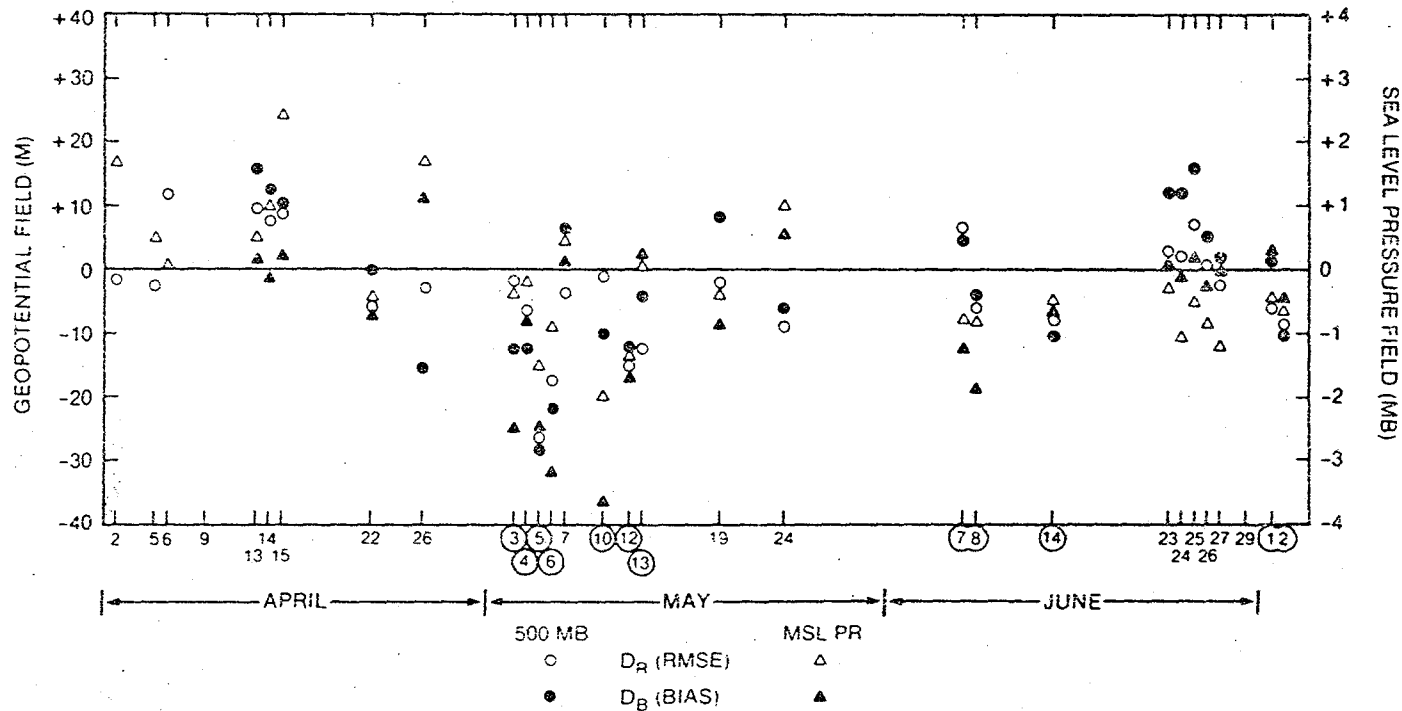
The largest differences either way appear in the RMSE and BIAS statistics, amounting in particular in the 12h thickness bias to more than a 70% positive impact made by MASS 2.0 upon the LFM forecasts.

The frequency of agreement in the sign of the impact of MASS 2.0 between the RMSE, S1, and BIAS statistics for 12h and 24h forecasts of both 500 mb geopotential and sea level pressure was considered. This was done to make a comparison with Atlas et al. (1982) and to justify use of only one or two of the statistics for the purpose of examining the temporal behavior of the model performance statistics. In the Atlas, *op. cit.*, study, the S1 scores and RMSE differences agreed in the sign of the impact 62% of the time; we find such agreement occurred 60% of the time (56 of 94 cases). The lowest frequency of agreement was between S1 and BIAS (39%), whereas the highest was between RMSE and BIAS (70%). Given this result and the aforementioned observation that the RMSE and BIAS scores also exhibited the largest impacts, it was decided to use these statistics to demonstrate time tendencies in model performance.

c. *Transitions in Model Performance Levels*

The scattergram plot in Fig. 4 demonstrates that the RMSE and BIAS statistics for 24h forecasts of 500 mb geopotential and sea level pressure undergo systematic variations in the sign of the impact throughout the course of the experiment. The first 9 cases show a significant MASS 2.0 superiority, whereas many of the cases in May and early June show clear LFM superiority; the remaining cases in late June tend to cluster around no impact either way. *Such an analysis as this is very powerful and informative, because it shows that very misleading results can be perceived when small samples (under 10 in size) composed of individual cases closer than a few days apart are used to make conclusive statements about model performances. In effect, two basic axioms of statistical analysis are then violated. First, the sample must be sufficiently large that its statistical properties approach that of the*

DAILY LFM-MASS 2.0 24H FORECAST DIFFERENCE STATISTICS



(N) DAY WITH PRONOUNCED 500 MB TROUGH OVER WESTERN ATLANTIC AND ANOTHER TROUGH SOMEWHERE ELSE WITHIN THE CONTINENTAL U S

Figure 4. Daily LFM-MASS 2.0 24h forecast difference statistics. Positive value represents favorable MASS 2.0 skill score compared to LFM.

ORIGINAL PAGE IS
OF POOR QUALITY

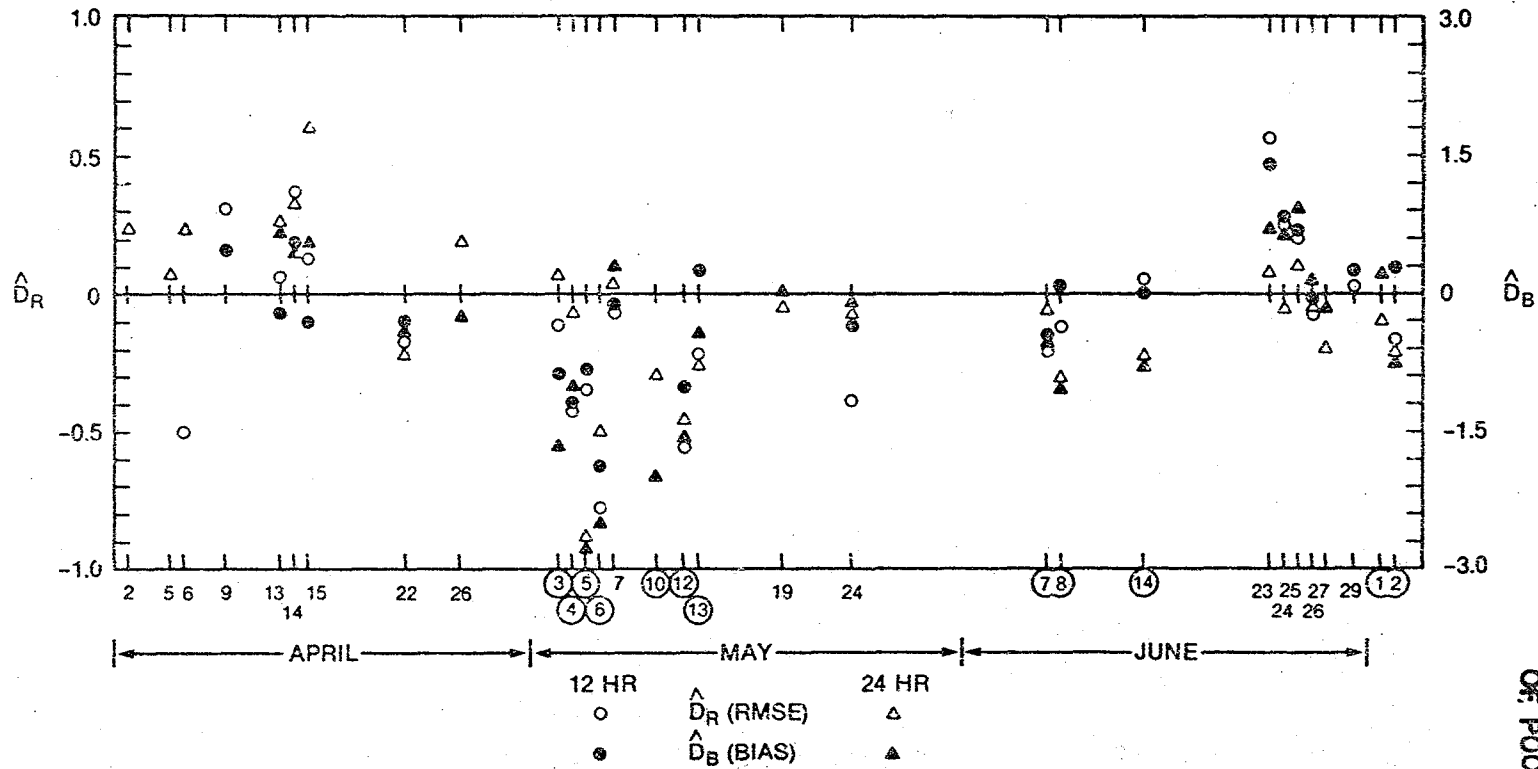
normal Gaussian distribution. Second, the principle of statistical independence of the individual data must not be violated (this problem arises because baroclinic waves normally develop with a 3-4 day period). These very important principles should be considered in any model evaluation.

It might be thought that the choice of only the 24h 500 mb geopotential and sea level pressure results may be unrepresentative of model performance in general. Therefore, the "pressure parameter-averaged normalized difference statistics" (defined in Table 3) for RMSE and BIAS at both 12h and 24h are shown in Fig. 5. The same general temporal trends still appear.

d. *The "BC Regime" Problem*

The dramatic drop in MASS 2.0 performance levels from April to early May was explored more carefully. Fig. 6 shows the early forecast evolution of the 500 mb geopotential and sea-level pressure fields on the day of the worst synoptic-scale performance by the model (May 5, 1982). A single contour and a single isobar are darkened to illustrate the sudden, rapid loss in atmospheric mass (pressure) at both levels in the model and the appearance of mesoscale 500 mb waves over the northern midwest and northeastern states only two hours into the model forecast. This unrealistic behavior suggests an initial model "shock", from which the model never fully recovers. The atmospheric mass fields forecast at 24h still show strong negative BIAS and huge RMSE differences. The other days with large negative difference statistics (Fig. 5) also were characterized by such model "shocks" in that part of the country.

A search was made for the underlying cause(s) of this problem. A particular large-scale flow pattern in the initial 500 mb geopotential field was found to have occurred on those days of poorest MASS 2.0 performance. This pattern, which will be referred to as the "BC (bad case)" regime, is one



Ⓝ DAY WITH PRONOUNCED 500 MB TROUGH OVER WESTERN ATLANTIC AND ANOTHER TROUGH SOMEWHERE ELSE WITHIN THE CONTINENTAL U.S.

Figure 5. Daily LFM-MASS 2.0 normalized and pressure parameter-averaged 24h forecast difference statistics.

ORIGINAL PAGE IS
OF POOR QUALITY

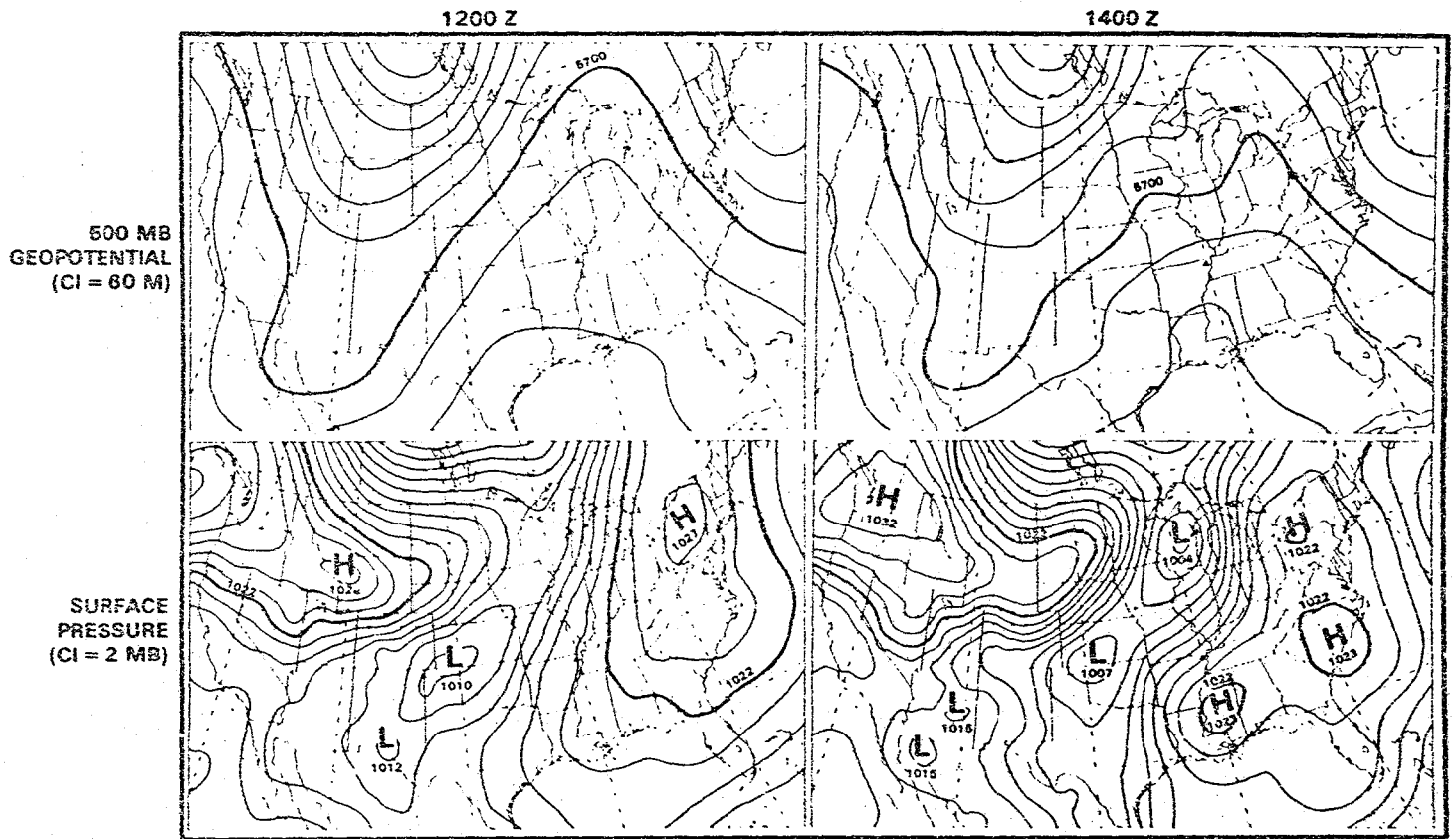


Figure 6. Fields of 500 mb geopotential and sea level pressure on May 5, 1982. Left figures show initial (1200 GMT) fields, right figures show 2h MASS 2.0 forecast (1400 GMT) fields.

27

OFFICE OF THE
 DIRECTOR
 OF THE
 NATIONAL
 WEATHER SERVICE

characterized by pronounced troughs over the far western Atlantic and also somewhere else within the continental United States (just as in Fig. (c)). Those twelve days on which such a regime existed are circled in Figs. 4 and 5. It is clear that these were also the days of poorest model performance (notice, for example, that 7 May stands out from the string of poor performance days in early May, and that the regime did not exist on that day). The presence of the regime did not always result in poor model performance (as on 1 July). Thus, our analysis shows that the model's performance is very sensitive to the nature of the initial flow.

It is unclear why this particular synoptic flow pattern would lead to such serious mass loss problems and yet other patterns do not. Baumhefner and Perkey (1982) also noticed a major loss in the forecast amplitude of synoptic features near those domain boundaries where significant changes were taking place. In particular, their December 10, 1967 case is identical in nature to our "BC regime", with a trough on the eastern boundary, a sharp ridge upstream, and a vigorous trough westward of the ridge. In both their study and ours, the amplitude loss errors had time continuity (persisted) and propagated inward (westward) from the eastern boundary. They suggest that the growth of serious error could be delayed by expanding the boundaries of the limited domain model away from areas of major synoptic activity. They show that a combination of (a) improper specification of the boundary values from the larger-scale model (the LFM in our case) and (b) blending of the nested (MASS 2.0) model predicted values with the specified boundary values within the "boundary sponge zone" explained the forecast errors. Boundary conditions may be the underlying cause of the "BC regime" problem also, since (a) there are strong similarities with the flow features and characteristics of Baumhefner and Perkey's case, and (b)

ORIGINAL PAGE 19
OF POOR QUALITY

there is evidence of a model "shock" phenomenon ^{OF POOR QUALITY} at the domain boundary where considerable synoptic activity is occurring. For these reasons, it is suggested that the method used in MASS 2.0 in determining the lateral boundary values may be the ultimate source for the most significant synoptic-scale errors in the model. It is shown in Fig. 7 that when those days characterized as "BC regime" days are excluded from the whole sample, MASS 2.0 equals or

SYNOPTIC SCALE FORECAST SKILL SCORE
COMPARISONS BETWEEN MASS 2.0 AND LFM MODELS

INCLUDING CASES CHARACTERIZED BY SYNOPTIC REGIME 'BC'

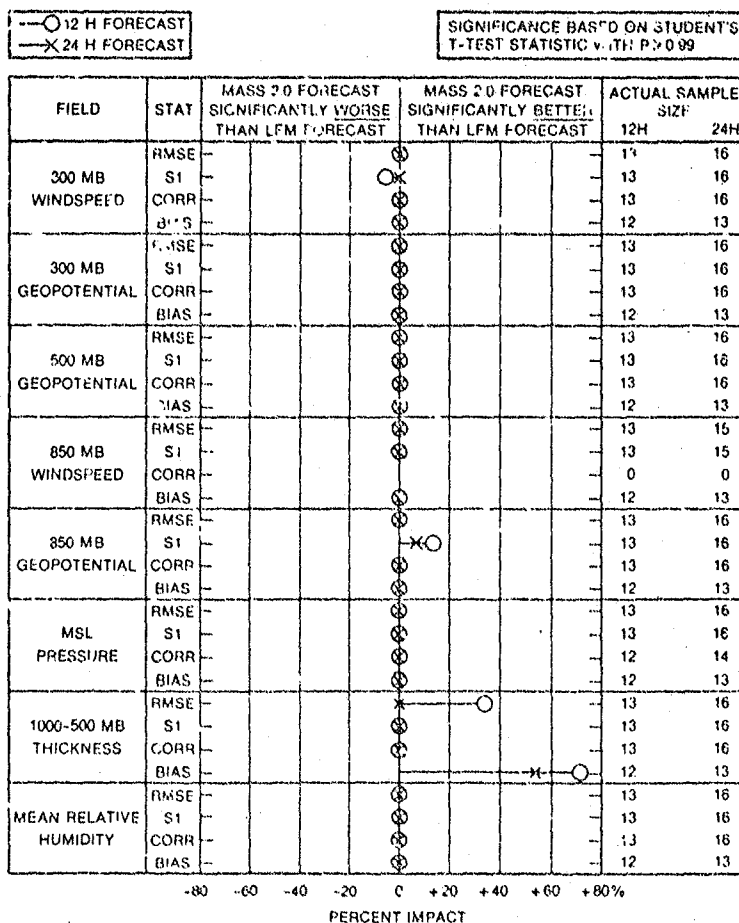


Figure 7. Synoptic-scale forecast skill score comparisons between MASS 2.0 and LFM models based upon the sample that excludes "BC regime" days.

ORIGINAL PAGE IS
OF POOR QUALITY

exceeds the performance of the LFM in forecasts of all but one synoptic-scale field studied.

5. SYSTEMATIC MODEL ERRORS OBSERVED AT SYNOPTIC SCALE

a. Results of Error Histo-Map Analysis

This discussion is organized into 12h and 24h cyclone and anticyclone center forecast errors and associated 500 mb geopotential and 1000-500 mb thickness forecast errors. First, a summary of the total number of sea-level pressure centers missed in the forecasts by MASS 2.0 and the LFM is given (Table 6), as derived from the error histo-maps. It is evident that both models underforecast the occurrence of observed cyclone centers much more than they overforecast them, particularly at 12h. MASS 2.0 suffers less from this

TABLE 6. NUMBER OF SIGNIFICANT¹ SEA LEVEL PRESSURE CENTERS MISSED IN 12H AND 24H FORECASTS BY THE MASS 2.0 AND LFM MODELS.

		PRESSURE CENTERS					
		CYCLONES			ANTICYCLONES		
Model	Error ²	12h	24h	Total	12h	24h	Total
LFM	NOC	1	4	5	7	7	14
MASS 2.0	NOC	1	5	6	6	0	6
LFM	NFC	13	10	23	2	5	7
MASS2.0	NFC	8	9	17	7	4	11

¹Significance determined by presence of at least one closed isobar.

²NOC = forecast center not observed (overforecast occurrence).
NFC = observed center not forecast (underforecast occurrence).

problem than does the LFM. In the case of anticyclone centers, MASS 2.0 underforecasts their occurrence much more so than the LFM, but only at the 12h period. A substantial increase in overforecasts of the number of cyclone centers from 12h to 24h verification period occurs in both models, whereas a dramatic decrease in overforecast number of anticyclone centers occurs with MASS 2.0. The total (12h plus 24h) anticyclone center forecast errors show MASS 2.0 (LFM) displaying a systematic bias towards underforecasting (overforecasting) the number of anticyclone centers, whereas both models show a bias towards underforecasting the number of cyclone centers observed.

1. *Anticyclone Forecasts*

An examination of the regional variation in systematic model errors is made possible by use of the error histo-maps. The error histo-map for 12h anticyclones is presented in Fig. 8 as an illustration of how the method was used to obtain general results. Both the distribution (scatter) and the number of events at the mode in the separate histograms were considered. It is evident that the LFM overforecast surface anticyclone amplitude with eastward displacement bias over the western half of the United States. In addition, the LFM had more of a problem with overforecasting the occurrence of anticyclone centers than with underforecasting them in this region, by a ratio of 7 to 2. Thus, the earlier result that the LFM generated unobserved anticyclone centers more frequently than it missed observed centers, is clarified here in two ways. First, the anticyclone problem was restricted to the western region. Secondly, the problem appeared in terms of overforecasting both the occurrence and the amplitude of anticyclone centers. MASS 2.0 did not exhibit such a nature (an explanation for this difference is offered in section 6). However, MASS 2.0

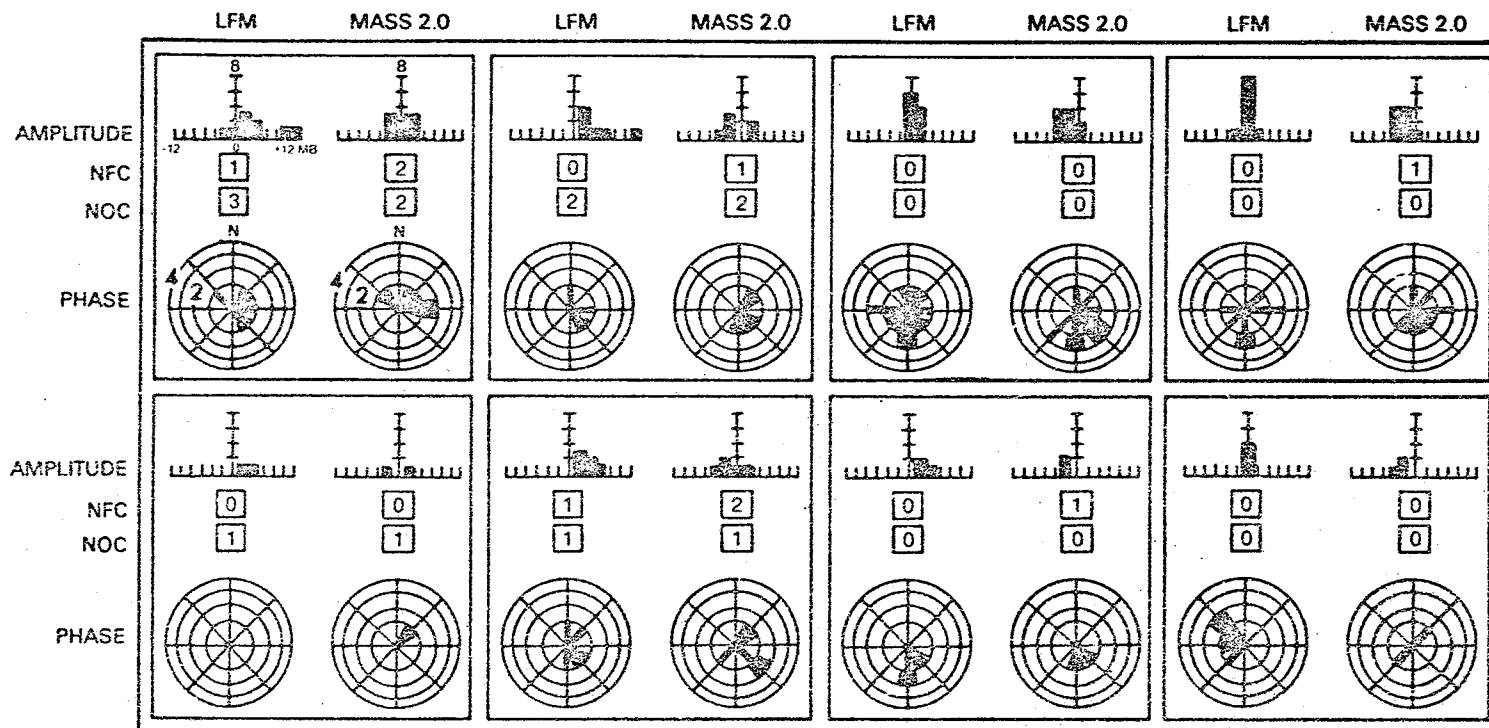


Figure 8. Error histo-map for 12h anticyclone forecasts. Within each large box (See Fig. 2), histogram plots of amplitude and phase errors and total number of "observed, but no forecast centers" (NFC) and "forecast, but no observed centers" (NOC) are shown. Error interval on amplitude (phase) error histogram is 2 mb (45°). Maximum event frequency on amplitude (phase) error histogram is 8(4). On phase histogram, two bar sizes are used, with the wider bar representing a forecast displacement error larger than 300 km.

underforecasted anticyclone amplitude and noticeably displaced anticyclones too far southeastward over the northeastern and north central states.

It is of interest to compare these 12h results to those at 24h and to examine the relationship of the surface anticyclone prediction errors to errors in the predicted 500 mb geopotential and 1000-500 mb thickness fields. Recall that the geopotential and thickness fields were examined within 500 km x 500 km square boxes centered over the observed surface pressure center positions. These errors represent the combined effects of amplitude and displacement errors. No distinction is made between these effects, as was done with the surface error fields. Such a distinction will be made later in the discussions of errors in predicted upper-level fields as determined from the subjective partitioning method.

Table 7 summarizes the 12h and 24h systematic forecast errors determined by this "error histo-map" method. The respective LFM and MASS 2.0 anticyclone 24h forecast errors are pretty similar to the 12h results just discussed. No amplitude errors occurred in the LFM forecasts of 500 mb geopotential over the anticyclones. The combination of LFM overforecast of surface anticyclone pressure and a lack of corresponding 500 mb geopotential error resulted in strongly underforecasted thickness fields, mainly over the western United States. In contrast, the MASS 2.0 12h and 24h forecasts of the 500 mb geopotential field did display systematic underforecast errors, but only over the northeastern and north central U.S. The combination of MASS 2.0 underforecasts in both the surface pressure and the 500 mb geopotential values resulted in no net systematic thickness errors there. Reasons for why the two models behaved so differently are discussed later.

TABLE 7. SUMMARY OF REGIONAL VARIATIONS IN MODEL SYSTEMATIC ERRORS IN FORECASTS OF SEA LEVEL PRESSURE FEATURES AND OF 500 MB GEOPOTENTIAL AND 1000-500 MB THICKNESS FIELDS IN CLOSE PROXIMITY TO THE OBSERVED POSITION OF THE PRESSURE FEATURES (RESULTS FROM ERROR HISTO-MAP ANALYSIS). UNDIAGNOSED, UNCERTAIN, OR SMALL NET ERRORS DENOTED BY "?"

OBSERVED SEA LEVEL PRESSURE FEATURE	FORECAST FIELD	FORECAST PERIOD	MODEL			
			LFM		MASS 2.0	
			ERROR TYPE	AREA OF UNITED STATES	ERROR TYPE	AREA OF UNITED STATES
			AMP ¹ /PHASE ²		AMP ¹ /PHASE ²	
ANTICYCLONE CENTERS	SEA LEVEL PRESSURE	12	A ⁺ /E	WEST 1/2	A ⁻ /SE	NE/N CENT
		24	A ⁺ /SE	WEST/CENT	A ⁻ /S	EAST 1/2
	500 MB GEOPOTENTIAL	12	??	ALL	A ⁻ /?	NE/N CENT
		24	??	ALL	A ⁻ /?	NE/N CENT
	1000-500 MB THICKNESS	12	A ⁻ /?	WEST/NORTH	??	ALL
		24	A ⁻ /?	WEST	??	ALL
CYCLONE CENTERS	SEA LEVEL PRESSURE	12	?/NE	WEST 1/2	??	ALL
		24	?/S	CENTRAL	A ⁻ /NE	PLAINS
	500 MB GEOPOTENTIAL	12	A ⁻ /?	WEST 1/2	??	ALL
		24	A ⁻ /?	WEST 1/2	??	ALL
	1000-500 MB THICKNESS	12	A ⁻ /?	WEST 1/2	??	ALL
		24	A ⁻ /?	WEST 1/2	A ⁺ /?	PLAINS

¹Amplitude Errors: A⁻ = too low (forecast < observed), A⁺ = too high

²Phase errors refer to forecast directional errors (e.g., S = forecast feature southward of observed)

2. Cyclone Forecasts

In the case of LFM cyclone forecasts, most of the significant errors in surface, 500 mb geopotential, and thickness fields again occurred over the western half of the United States with the noticeable exception of southward displacement errors in 24h cyclone predictions over the central region (Table 7). However, systematic amplitude errors in cyclones were not observed as they were with anticyclones. Negative biases in thickness fields over the West still result, but for a different set of reasons, namely, the combination of underforecast 500 mb geopotential with (in 12 hour only) northeastward cyclone phase errors. This displacement error was associated with the model tendency to bring surface anticyclones inland too rapidly and too strongly. Silberberg and Bosart (1982) also found 24h LFM thicknesses over observed cyclones to be too low (cold) and for model springtime cyclones to move too fast over the western United States.

Cyclones were in general predicted with very little systematic error by MASS 2.0 in 12h, but were overforecast (forecast pressure lower than observed pressure) and displaced systematically northeast of the observed location over the Plains states in 24h. Since the forecast 500 mb geopotential fields did not display any definite amplitude bias, the 24h cyclone error was associated with a noticeable positive bias in 24h thickness forecasts over the Plains states.

b. *Subjectively Determined Systematic Model Forecast Errors*

In the second approach used to obtain systematic model errors, the "difference maps" were again used as input, but phase errors were removed

subjectively prior to the computation of forecast amplitude errors. The results of the analysis of the 500 mb geopotential and vorticity fields are given in Table 8, and the results of the analysis of the 1000-500 mb thickness and layer-mean relative humidity fields are given in Table 9. The individual errors were entered according to the state in which each error occurred, and were then tabulated by regions (Fig. 9) roughly equivalent to the regional definitions used in the error histo-map approach (Fig. 2).

1. 500 mb Geopotential and Vorticity

The geopotential errors are classified according to whether the error occurred closest to an observed ridge or trough. The only region where the geopotential was systematically too high in either model was over the southern plains, where MASS 2.0 overforecasted ridge strength. Troughs and ridges were both underforecasted over the northeastern states by MASS 2.0, in agreement with earlier results. Both models underpredicted 500 mb ridges across most of the northern tier of states, and the LFM also underpredicted troughs over the southwest (in 12h). The most frequent type of error with vorticity maxima was in location, rather than amplitude. These features in MASS 2.0 were displaced south or southeast of their observed positions in several regions, whereas LFM features were predicted too far east (too fast), particularly in the western half of the nation. Thus, MASS 2.0 did not display the systematic LFM eastward phase speed error in forecasts of vorticity maxima, but rather tended to "dig" such features too far southward.

TABLE 8. REGIONAL VARIATIONS IN AVERAGE MODEL AMPLITUDE AND PHASE (IN PARENTHESES) ERRORS BASED UPON SUBJECTIVE ANALYSIS OF "DIFFERENCE MAPS" (FORECAST FIELD MINUS OBSERVED FIELD)^{1,2} FOR 500 MB GEOPOTENTIAL AND VORTICITY FIELDS. MEAN GEOPOTENTIAL (VORTICITY) AMPLITUDE ERRORS SMALLER THAN 30 m ($2 \times 10^{-5} s^{-1}$) ARE NOT SHOWN.

FORECAST PERIOD (HRS)	MODEL	FEATURE	REGION							
			1	2	3	4	5	6	7	8
<u>500 MB GEOPOTENTIAL</u>										
12	LFM	RIDGES	-43	*	-34	*	*	-31	*	*
12	MASS 2.0	RIDGES	-33	*	*	*	-37	*	*	*
24	LFM	RIDGES	*	*	-41	*	-30	*	*	*
24	MASS 2.0	RIDGES	*	*	-37(E)	+32	-49	*	-50	-40
12	LFM	TROUGHS	*	-44	*	*	*	*	*	*
12	MASS 2.0	TROUGHS	*	*	*	*	*	*	-33	*
24	LFM	TROUGHS	*	*	*	*	*	*	*	*
24	MASS 2.0	TROUGHS	*	(E)	*	*	*	*	-68	*
<u>500 MB VORTICITY</u>										
12	LFM	MAXIMA	(ESE)	(ENE)	*	(W)	*	*	*	*
12	MASS 2.0	MAXIMA	*	(SE)	*	(S)	*	*	*	*
24	LFM	MAXIMA	-4	*	(ENE)	*	(E)	*	*	*
24	MASS 2.0	MAXIMA	-4	*	*	*	(SE)	(S)	*	*

¹Mean statistics shown only for a region where at least 4 error observations are of the same sign (or same direction) and no more than 33% of the error observations are of the opposite sign (or deviate by more than $\pm 45^\circ$ from the majority average direction).

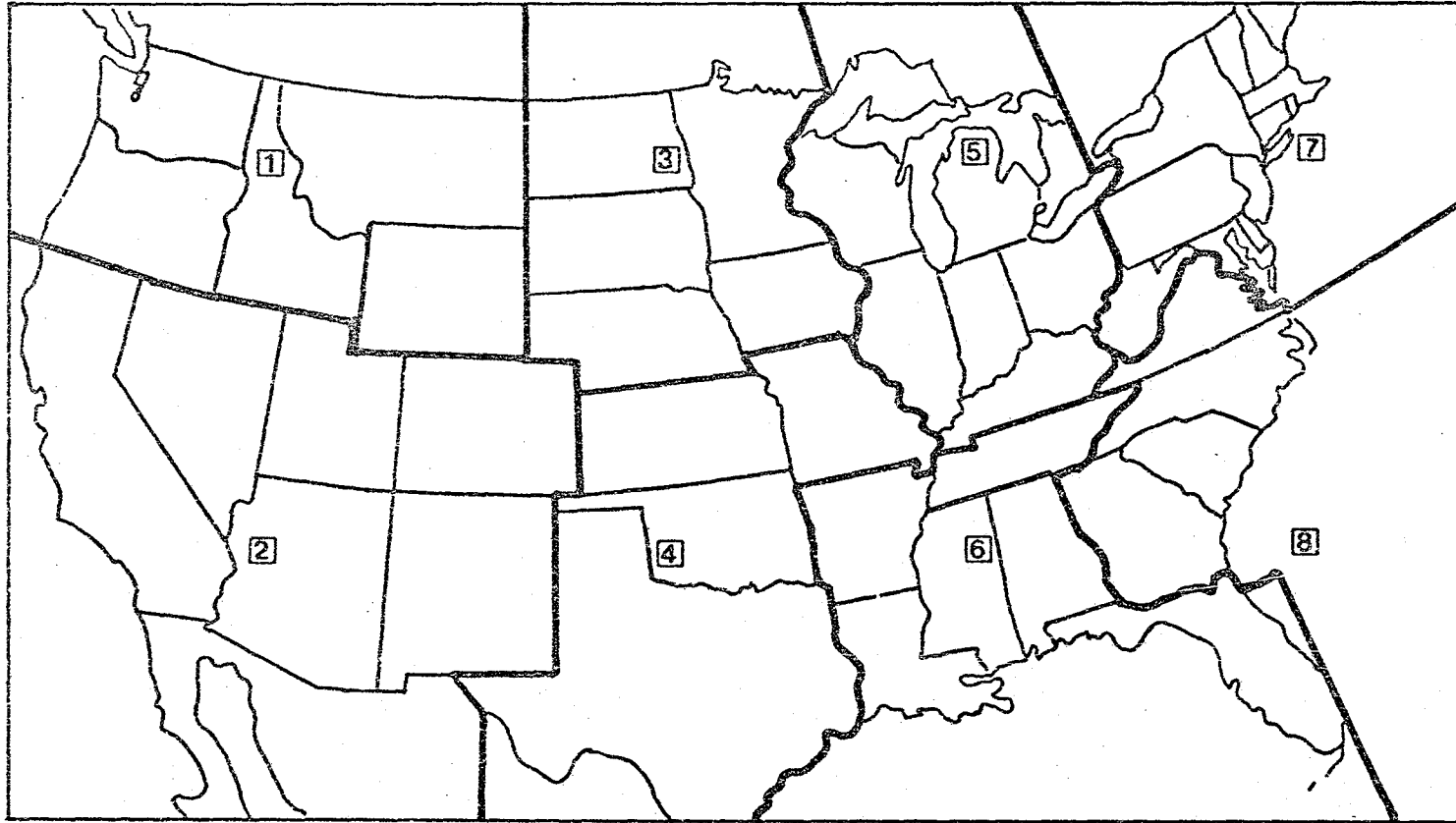
²Mean amplitude error is defined as the average of all ϵ observed, where ϵ is the maximum error value within an area of error on a difference map, after subjective removal of the phase error contribution to that error value.

TABLE 9. SAME AS FOR TABLE 8, EXCEPT FOR 1000-500 MB GEOPOTENTIAL AND THICKNESS AND MEAN RELATIVE HUMIDITY FIELDS.^{1,2} MEAN THICKNESS (RELATIVE HUMIDITY) ERRORS SMALLER THAN 30 M (20%) ARE NOT SHOWN.

FORECAST PERIOD (HRS)	MODEL	FEATURE	REGION							
			1	2	3	4	5	6	7	8
<u>1000-500 MB THICKNESS</u>										
12	LFM	RIDGES	-44	-57	*	-44	40	*	*	-37
12	MASS 2.0	RIDGES	*	-35	*	*	*	*	*	*
24	LFM	RIDGES	-47	-39	*	*	*	*	*	*
24	MASS 2.0	RIDGES	*	*	+57	+40	*	*	*	*
12	LFM	TROUGHS	-73	-69	-43	-37	*	*	*	*
12	MASS 2.0	TROUGHS	*	*	*	*	*	*	*	*
24	LFM	TROUGHS	-85	-71(E)	*	*	*	*	*	*
24	MASS 2.0	TROUGHS	*	*	*	*	*	*	*	*
<u>MEAN RELATIVE HUMIDITY</u>										
12	LFM	MAXIMA	*	-20(E)	*	*	(SE)	-22	*	-26
12	MASS 2.0	MAXIMA	(S)	*	*	-27(S)	(E)	-31	*	*
24	LFM	MAXIMA	*	*	(SE)	-20(NE)	*	*	*	*
24	MASS 2.0	MAXIMA	*	*	*	(N)	*	*	*	*
12	LFM	MINIMA	*	*	*	*	*	*	*	*
12	MASS 2.0	MINIMA	*	*	*	*	*	*	*	*
24	LFM	MINIMA	*	*	*	*	*	*	*	*
24	MASS 2.0	MINIMA	*	*	*	-20	*	*	*	*

¹Mean statistics shown only for a region where at least 4 error observations are of the same sign (or same direction) and no more than 33% of the error observations are of the opposite sign (or deviate by more than $\pm 45^\circ$ from the majority average direction).

²Mean amplitude error is defined as the average of all observed, where is the maximum error value within an area of error on a difference map, after subjective removal of the phase error contribution to that error value.



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 9. The eight regions within which subjectively determined forecast errors are tabulated.

2. 1000-500 mb Thickness

The greatest difference between the two models in terms of systematic errors was in forecasts of the thickness field. The LFM displayed large mean errors, all of which were systematically too low (cold). This was particularly true over the western half of the U.S. The errors were less widespread in 24h, but were still about as large in magnitude. Also, 24h thickness troughs in the southwestern region were frequently east of their observed locations. These LFM errors were also seen in the error histo-map results (not shown). On the other hand, thickness errors rarely appeared in MASS 2.0; the only interesting errors are over-forecasts of ridges over the Plains states in 24h. This result is consistent with earlier results (Table 7). Recall that 500 mb geopotential ridges and surface cyclones were also over-forecast at 24h over the Plains states (Tables 8 and 7, respectively).

3. Mean Relative Humidity

Table 9 shows that both models suffer from negative biases in forecast relative humidity maxima, particularly so at 12h. In particular, humidity maxima are forecast too low by MASS 2.0 at 12h over the southern states (regions 4 and 6), whereas humidity minima are underforecast at 24h over the Southern Plains region.

6. SOURCES OF IMPORTANT MASS 2.0 SYSTEMATIC ERRORS

Explanations for the following systematic forecast errors in MASS 2.0 were sought:

(E1.) Surface anticyclone intensity is seriously underforecast, 500 mb

heights (troughs and ridges) are moderately underforecast, and there is a strong tendency to displace surface anticyclones too far south or southeast over the northeastern and north central states. The surface pressure amplitude error expands to include all the East in 24h. No net thickness bias results.

(E2.) 1000-500 mb thickness is overforecast, surface cyclones are too deep (overforecast) and are displaced northeast of their observed locations, and 500 mb ridges are somewhat overforecast over the Plains states at the 24h verification period.

(E3.) Magnitudes of relative humidity maxima are underforecast over the southern states primarily at 12h. Humidity minima are underforecast at 24h, but over a smaller region.

Before examining further these individual errors, an explanation is provided for the dramatic improvement made by MASS 2.0 on the systematic thickness errors in the LFM forecasts. Recall that the greatest positive impact made at the synoptic-scale by MASS 2.0 on the LFM forecast fields is the dramatic reduction in negative thickness biases over the western half of the nation throughout the 24h verification period. The erroneously cold thicknesses were the net result of (1) overpredicted, eastward displaced (too fast) surface anticyclones, (2) underpredicted 500 mb trough and ridge amplitude with eastward phase errors (vorticity maxima too fast), and (3) eastward displaced surface cyclones, frequently forecast too weak. In other words, upper-level troughs and associated surface ridging were brought eastward into the interior West too rapidly behind underforecast cyclones, producing an atmosphere that was too cold. Apparently, the increased vertical resolution and more sophisticated planetary boundary layer parameterization in MASS 2.0 led to more realistic lower tropospheric heating rates. Moreover, the

increased horizontal resolution and higher order numerics seemingly resulted in smaller eastward phase errors.

The three major systematic errors in MASS 2.0 are now examined in more depth. Underlying causes for these errors are sought.

a. *Weak Surface Anticyclones in the East (E1)*

The source of the weak, southerly displaced anticyclones and erroneously low 500 mb heights was easily found once the "BC regime" had been isolated. Fig. 6 showed an actual example of such errors which occurred on the day of most devastating BC regime effects. A study of those days on which such systematic errors were greatest revealed that in every case a BC regime was present in the initial state. Thus, it is suggested that improper specification and/or formulation of the eastern boundary conditions may have been the underlying cause for this very important systematic error at the synoptic scale.

b. *Overforecast Thickness Values and Surface Cyclone Intensity over the Plains States at 24h (E2)*

More analysis and synthesis was needed to unravel the ultimate cause of the seemingly related thickness and surface cyclone errors. A typical case is presented in Fig. 10. A 24h forecast error of +88 m in the thickness field occurred in southwestern Iowa in this case, about midway between the observed thickness ridge and trough. The error had the effect of displacing the forecast thickness ridge from its observed location over Lake Michigan westward to Iowa. From hydrostatic considerations, the erroneously high mean temperature in the air column over Iowa apparently resulted in a displacement

ORIGINAL PAGE IS
OF POOR QUALITY

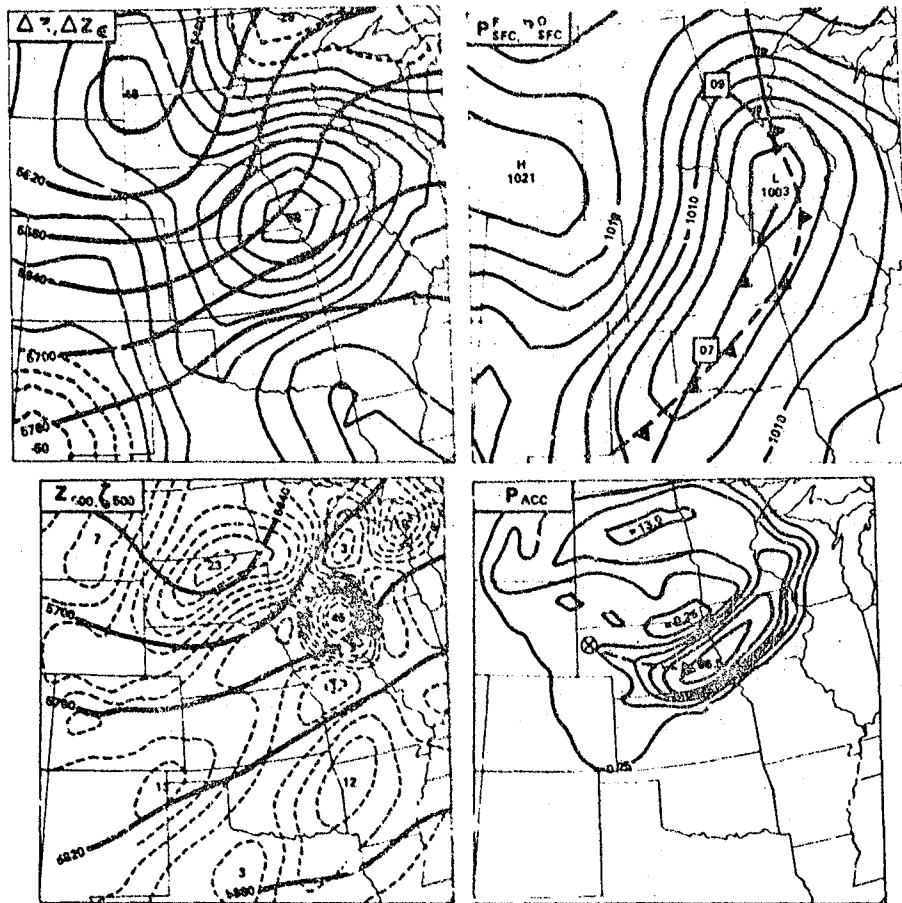


Figure 10. 24h forecast and verification analyses valid at 1200 GMT 9 June 1982: (1) 1000-500 mb thickness analysis (heavy lines in m) and thickness "difference map" (light solid (forecast greater than observed values) and dashed (forecast less than observed values) lines, in m), (ΔZ and ΔZ_e , respectively) / (b) forecast surface pressure field (P_{SFC}^F) and diagnosed frontal system (solid), and observed locations of surface low pressure systems (P_{SFC}^O) and frontal system (dashed), in mb ($07 = 1007$ mb); (c) forecast 500 mb geopotential (Z_{500}) (solid, in m) and vorticity (ζ_{500}) (dashed, $10^{-5} s^{-1}$) fields/ and (d) forecast of 24h accumulated stable precipitation (P_{ACC}), in mm. Dashed line traces past movement of maximum from its origin at position "X".

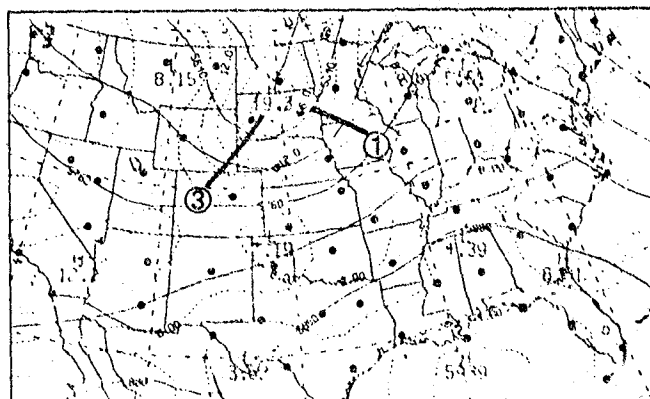
of the forecast surface cyclone toward the forecast position of the thickness ridge and away from the observed cyclone location in western Minnesota. Also note that the forecast cyclone is 6 mb deeper than observed and that the forecast cold front does not display as prominent an eastward bulge as was

observed. The anomalously warm air forecast in Iowa held back the eastward progression of the cold front and hydrostatically resulted in surface pressure that was too low.

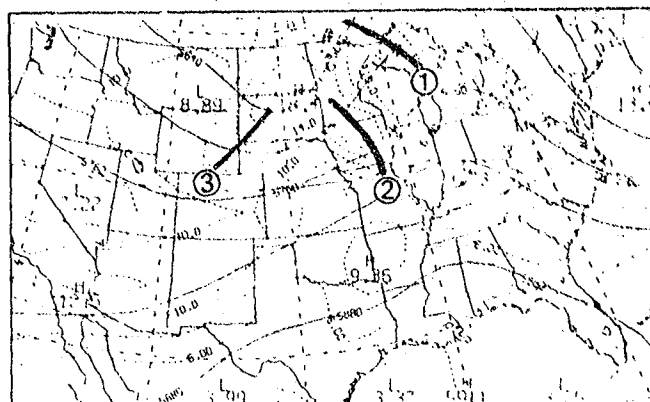
The 500 mb geopotential forecast shows a very curious small-scale feature possessing extremely high vorticity located very close to the maximum thickness error location. The high vorticity anomaly is surrounded by several small centers of very low absolute vorticity. This anomaly serves as a clue to the overforecast thickness values and surface cyclone intensity. Another clue is offered by the nearby huge amount of 24h total stable precipitation predicted by the model (Fig. 10d).

The vorticity anomaly is investigated further first. A comparison between the verification analysis at 12Z 9 June 1982, the smoothed MASS 2.0 24h forecast, and the LFM 24h forecast of the 500 mb geopotential and vorticity fields appears in Fig. 11. Both models forecast wave #3 well. However, MASS 2.0 shows two waves over the upper Mississippi River region, whereas the LFM and the observations both show only one. Even so, the LFM forecast of wave #1 has a significant phase error. The question arises as to whether MASS 2.0 forecast correctly the position of wave #1 (i.e., that perhaps designators 1 and 2 should be switched). Examination of the two-hourly MASS 2.0 map output shows clearly that this is not the case, as explained immediately below. It is important to note that the existence of wave #2 cannot be verified with the operational rawinsonde network (Fig. 11a) since there is not a single station in the entire state of Iowa. Thus, to what extent wave #2 is real cannot be assessed.

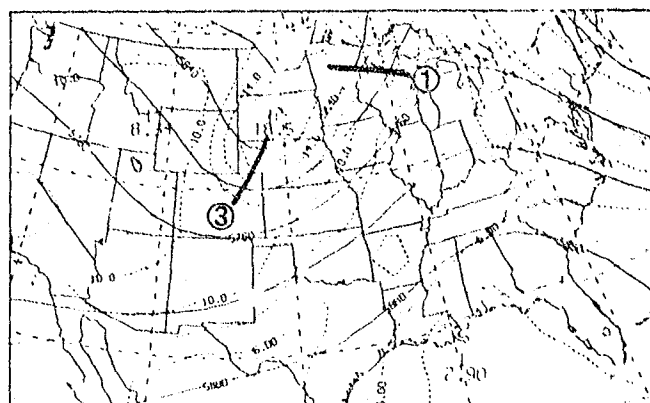
The origin and evolution of waves #1 and #2 in the MASS 2.0 simulation are seen in Fig. 12. Both waves intensify during their northeasterly trek across



VERIFICATION ANALYSIS AT 12Z JUNE 9, 1982



24 HR FILTERED MASS FORECAST



24 HR LFM FORECAST

Figure 11. 24h 500 mb geopotential and vorticity (a) verification analysis, (b) filtered MASS 2.0 forecast, and (c) LFM forecast valid for 1200 GMT 9 June 1982. "Short waves" in the vorticity fields are labelled numerically. Reporting rawinsonde station locations shown in (a). Isolines in same format as in Fig. 10c.

ORIGINAL PAGE IS
OF POOR QUALITY

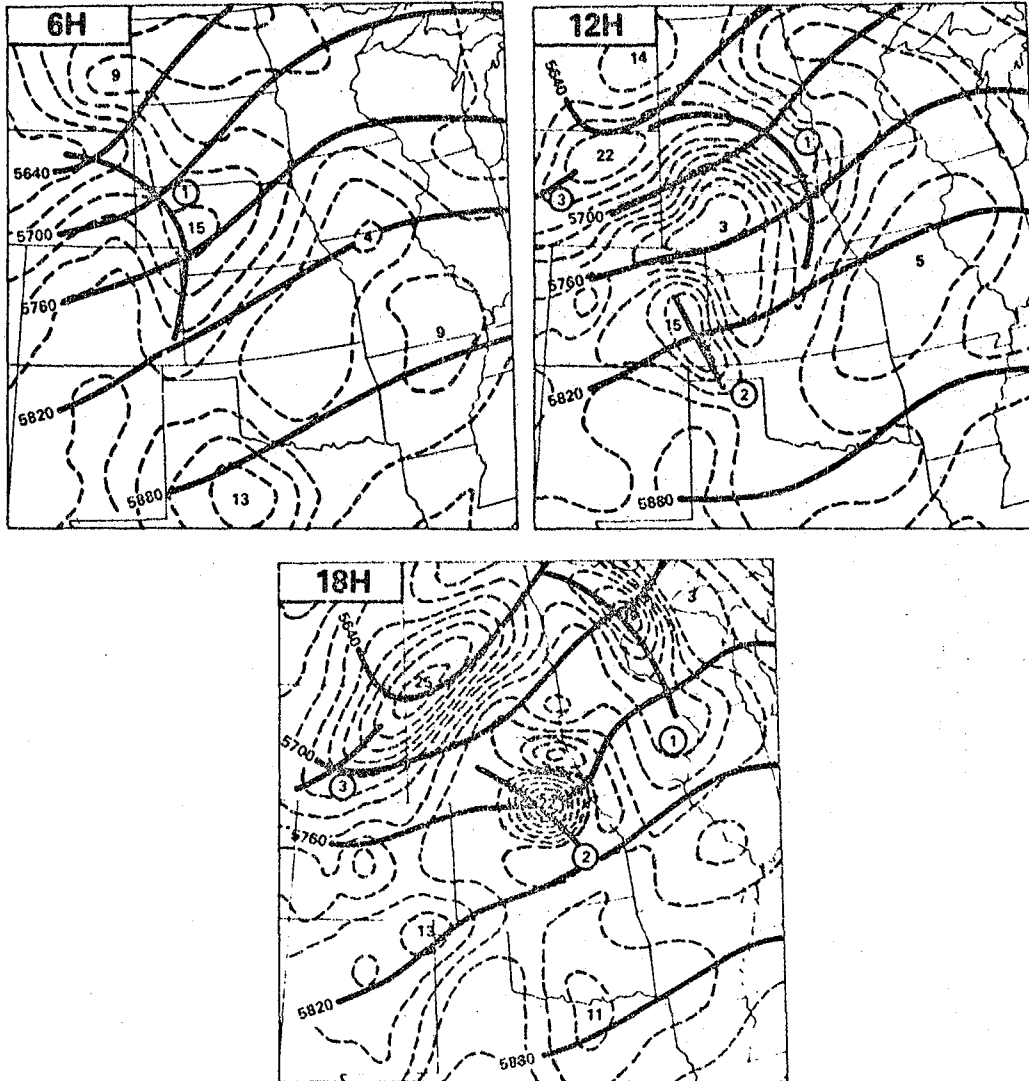


Figure 12. MASS 2.0 unfiltered forecasts of 500 mb geopotential and vorticity (as in Fig. 11) at 1800 GMT 8 June 1982 (6h), 0000 GMT 9 June 1982 (12 h), and 0600 GMT 9 June 1982 (18 h).

the northern plains states. However, wave #2 is non-existent during the morning, forms in the lee of the Colorado Rockies by 12h, and afterward develops extremely rapidly into the vorticity anomaly. It is also evident that a region of low vorticity just to its northeast is identifiable by 12h and that

this feature has by 18h almost become dynamically unstable (absolute vorticity negative). The lobe of low vorticity can be traced back to central Colorado at 6h. There are strong indications that the low and high vorticity centers are coupled after 12h. The appearance of the vorticity maximum in the lee of the mountains at the time of diurnal temperature maximum and the close association between the two vorticity centers thereafter suggest that the vorticity couplet was produced as the strong flow encountered the heated mountain range barrier. The thermal plume (or elevated PBL) over the mountain range apparently acted as an obstacle to this nearly normal flow, thereby redirecting most of the air around the range and producing the vorticity couplet. Downstream propagation of the couplet is consistent with the notion that the plume and redirected flow were advected downstream.

Additional evidence in support of this obstacle flow mechanism is provided in Fig. 13. The maximum PBL development anywhere within the United States is predicted to occur over the Colorado Rockies, directly upstream of the vorticity couplet location (which first appeared only 3h earlier). A large increase in surface temperature there of 19C from 1200 GMT 8 June to 0000 GMT 9 June accompanied the rise of the PBL over the mountains. Notice in Fig. 13b that the center of strongest upward motion is located directly beneath the local region of strong positive vorticity advection associated with the vorticity couplet (Fig. 12). Also observe that subsidence occurs at and to the lee of the highest mountains.

These kinds of features are similar to those shown by Mahrer and Pielke (1977) in a simulation of flow over the Sacramento mountain range in New Mexico. The model used was a hydrostatic, three-dimensional, primitive equation model, which included a detailed PBL parameterization similar to that

ORIGINAL PAGE IS
OF POOR QUALITY

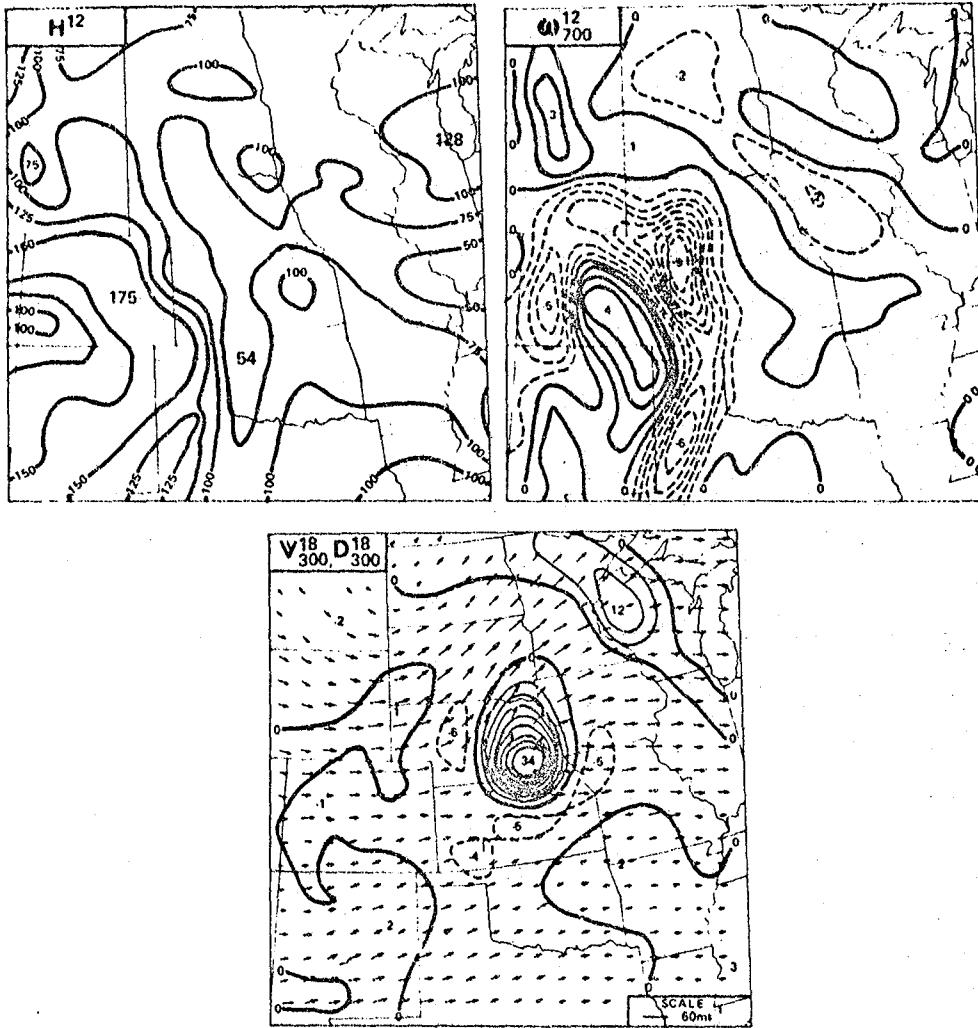


Figure 13. MASS 2.0 unfiltered forecasts of (a) planetary boundary layer depth (mb) at 12h (H^{12}); (b) 700 mb omega vertical motion field ($\mu\text{bar s}^{-1}$) at 12 h (ω_{700}), (dashed = upward); and (c) 300 mb wind vectors (m s^{-1}) and divergence (solid positive, dashed negative, intervals of 10^{-5} s^{-1}) at 18h (V_{300} , D_{300}) produced from initial conditions at 1200 GMT 8 June 1982.

used in MASS 2.0. Their model included an additional expression for the amount of solar radiation received on a slanted surface, so that the eastern slopes of the mountains would warm earlier. Their simulation was designed for a smaller scale study (a 5 km grid was utilized). In their simulation, the impinging air

mainly deflected around the mountain range rather than flowing over it. The effect of the differential heating was to create the familiar mountain-valley breeze. The ambient flow then advected the thermal plumes downwind from the highest mountains. Hydrostatic pressure fall beneath the advected plumes was the cause for isalobarically- forced convergence (upward motion) areas directly downwind of the highest mountain peaks.¹ Warner *et al.* (1978) have also generated such thermally driven circulations downwind of the Appalachian mountain chain in the PENN model. In all cases, the deflection of the air flow produced anticyclonic (cyclonic) vorticity on the left (right)-hand side of the mountain range, looking downwind, and retarded horizontal motion in the downwind convergence areas.

Thus, the vorticity couplet on 8 June 1982 was produced in the lee of the mountains as the likely result of a thermally forced barrier flow. The vorticity couplet was advected downstream, but would not have intensified as such had this been the only process operating. The reason for this intensification is suggested by the fact that MASS 2.0 predicted a large amount of 24h accumulated precipitation in close proximity to the vorticity couplet (Fig. 10d). Notice that the precipitation maximum in eastern Nebraska originally appeared in northwestern Nebraska (at 12h), near to the center of maximum upward motion (Fig. 13b). In this region, the forecast lifted index had fallen to -6. The lifting associated with the mountain barrier flow was able to

¹At the scale of the vorticity couplet in the MASS 2.0 model, quasi-geostrophic vorticity advection and lower tropospheric thermal advection Laplacian effects are important in forcing upward motion. However, isalobaric effects were also evident in the model output downwind of the mountains.

release the potential instability present there. However, because MASS 2.0 lacks a cumulus parameterization scheme, convective overturning could not occur. Hence, the model atmosphere could not properly stabilize as it does in nature. Instead, the latent heat released during the condensation process could only feedback in a positive sense to further drive the mesoscale vertical circulation. In this feedback process, the latent heat liberated in the air column hydrostatically leads to falling surface pressure. Isalobarically-enhanced low-level convergence then further enhances precipitation formation. This process has been described as Conditional Instability of the Second Kind (Charney and Eliassen, 1964), or CISK. By 18h the upper-level wind field has been blistered by the presence of the ever-growing "hot core" (Fig. 13c); note in particular the jet to the north of the divergent outflow core (an increase of 15 m s^{-1} in 6h) and the retarded obstacle-like motion at the core. These features are reminiscent of those produced by mesoscale convective complexes (MCC), which have been shown by Maddox *et al.* (1981) in numerical experiments to be capable of significantly perturbing the synoptic-scale flow fields in a similar manner.

In summary, it appears that the thickness and surface cyclone prediction errors in this one documented case arose from the following sequence of events:

- (1) Under the ideal conditions of strong normal flow to the Rocky Mountain chain and strong surface heating, rapid development of the PBL led to a pronounced barrier flow regime around the heated mountains (vorticity couplet).

- (2) The thermally driven mesoscale circulation was advected downwind, resulting in an area of strong upward motion in the lee of the highest barrier.

- (3) Since the air was convectively unstable, the mesoscale lifting generated stable precipitation in the model. Lack of a cumulus parameterization resulted in a CISK-like process, whereby the circulation rapidly intensified, forcing even greater release of latent heat, which forced

stronger upward motion. This process was not totally runaway, as maximum intensification typically occurred during the 12-18h forecast period.

(4) This "CISK-wave" (not to be confused with wave-CISK (Lindzen, 1974)) appeared as a small-scale vorticity maximum in the 24h synoptic-scale forecasts. Other synoptic-scale fields were also perturbed by this mesoscale instability: (a) thickness was too high over the CISK-wave due to the excessive heating in the CISK process, (b) the surface cyclone was too deep and erroneously displaced toward the thickness error "bullseye" as a hydrostatic consequence to the heating and as a dynamic response to the intense small-scale divergence aloft resulting from the CISK process, and (c) an upper-level jet formed to the north of the CISK-wave where the divergent outflow converged with the ambient flow, similar to MCC-like phenomena.

Five other cases were found wherein (1) an overforecast at 24h of thickness by at least 40m occurred over regions 3 or 4 of Fig. 9, and (2) predicted surface cyclones within the same region were overforecast by at least 2 mb. Recall that these were the errors to be explained in Tables 7, 8, and 9. Relevant synoptic-scale features forecast by MASS 2.0 on the six days are summarized in Table 10. In each of the cases:

(1) A 500 mb vorticity couplet formed near the time of peak surface heating and near the location of maximum PBL development.

(2) The wind direction upstream of the vorticity couplet was strong and nearly normal to the mountain range.

(3) Strong heating and destabilization occurred locally.

(4) Stable precipitation formed at or after the time of vorticity couplet (CISK-wave) appearance.

(5) Precipitation rate maximized at night (after 18h) and amounted to at least 12 mm by 24h.

(6) A strong mesoscale divergence anomaly developed in the 300 mb wind fields, which in most cases was strongest during the 10-18h interval.

(7) Synoptic-scale thickness and surface cyclone errors were directly associated with the CISK-wave. It is also of interest to note that in every case except 8 June, strong convection actually did develop in close spatial and temporal proximity to the predicted 700 mb upward motion center associated with the CISK-wave. *This suggests that the early evolution of the thermally driven wave was somewhat realistic. However, because the observed convection usually weakened much earlier than did the predicted forcing, the later evolution was*

TABLE 10. FEATURES OF VARIOUS FIELDS FORECAST BY MASS 2.0 ON DAYS WHEN SIGNIFICANT POSITIVE ERRORS IN 24 HR THICKNESS FIELD FORECASTS OCCURRED IN CLOSE JUXTAPOSITION TO SIGNIFICANT OVERFORECAST CYCLONE INTENSITY ERRORS.

Field Feature	5 May	10 May	19 May	8 June	14 June	23 June
<u>24 hr Thickness Error</u>						
Location	N TX	NE	SW NE	SW IA	N KS	W OK
Total (amp+phase) error (m)	+52	+63	+94	+88	.56	+48
<u>500 mb Flow</u>						
Vorticity Couplet (vc) init. location	NE NM	W NE	N CO	SW NE	CO	NE NM
Time of vc formation (t_{vc})	21Z	22Z	20Z	21Z	18Z	18Z
Initial wind dir upstream of vc	SW	SW	SW	SW	SW	NW
<u>24 hr Surface Cyclone</u>						
Forecast cyclone location	N TX	W KS	NW KS	IA	SW IA	NW TX
Amplitude (mb)/phase (km) errors	800SW/-3	200NE/-4	400N/-7	500SE/-6	400SW/-2	None Obs/-3
<u>Local QPF Maximum</u>						
24 hr total (mm)/location of max	25.4/NW TX	12.7/WY	106.7/WY	66.0/SE NE	43.2/W KS	17.8/NW TX
Time of QPF formation(max QPF rate)	21 (06) Z	00 (02) Z	22 (10) Z	00 (08) Z	20 (04) Z	02 (04) Z
<u>PBL Features</u>						
Max PBL ht/location of max	2.7km/NE NM	3.1km/E CO	322mb/SE CO	175mb/E CO	198mb/W KS	146mb/W TX
12 h sfc temp change (°C)/LI (t_{vc})	14/-4	11/-3	19/-3	19/-6	16/-6	18/-5
<u>300 mb Divergence/700 mb Up motion</u>						
Location of initial formation	NW TX/NW TX	E CO/W NB	SE WY/W NE	W NB/W NE	NE CO/NE CO	CO/NE NM
Time of first appearance	18Z/22Z	20Z/18Z	18Z/18Z	02Z/18Z	18Z/16Z	20Z/18Z
Time of greatest intensity	06Z/00Z	22Z/02Z	04Z/08Z	06Z/06Z	23Z/22Z	04Z/04Z
<u>Observed Convective System</u>						
Initial location/time of formation	NW TX/19Z	NE/20Z	E CO/20Z	E NE/14Z	W KS/19Z	NW TX/20Z
Time of max intensity	23Z	22Z	22-02Z	14-06Z	22Z	02Z

52

ORIGINAL PAGE 13
OF POOR QUALITY

poorly predicted as the CISK process took over. The mesoscale predictability aspects of the problem are treated further in Part II of this paper.

c. *Low Relative Humidity Values Over Southern States (E3)*

This last error to be discussed is significant because (as will be shown in Part II) it was a major cause of underforecasts of convection. The lifted index fields are later used as convective predictor variables. Since the lifted index is partly determined by the vertical profile of moisture, underforecasts of lifted index were investigated along with underforecasts of relative humidity.

Subjective examination of underforecast relative humidity (RH) and lifted index (LI) patterns over the Southern Plains (SP) and South Central (SC) regions revealed that such patterns frequently could be traced backward in time to inadequacies in the respective initial fields. In particular, 40% of the situations where RH was underforecast by at least 20% occurred in local areas where there was a negative bias of at least 5% in the RH initialization. Likewise, 69% of the situations where unstable LI values were underforecast by at least 4 occurred in local areas where there was a negative bias of at least 2 in the LI initialization.² It was not true that poor initialization inevitably led to poor forecasts, however, as such was the case only about 25% of the time.

²A total of 21 (13) situations of significantly underforecasted RH (LI) occurred over the SC and SP regions. A total of 33 (36) cases of significant RH (LI) negative bias in the initialization over the SC and SP regions were observed.

The relationships between forecast and initialization errors varied monthly. In April, 83% of the underforecast RH cases could be tied to poor RH initialization. In contrast, 86% of the underforecast LI cases were related to poor LI initialization problems in June and July. Thus, poor initialization of the RH field was the overwhelming cause of underforecasts of RH during the spring, whereas poor initialization of the LI field was the overwhelming cause of underforecasts of LI during the summer months.

The forecast and initialization problems also varied regionally. In the SC (SP) area, 63 (23)% of the forecast RH errors were related to poor initialization. No regional dependency was found with the LI errors. It is also interesting that 38 (62)% of the forecast RH errors (not necessarily related to initialization errors) occurred in the SC (SP) area; again, forecast LI errors showed no such regional dependency. These results indicate that in the SC region where forecast RH errors were relatively infrequent, poor initialization was the dominant cause for the appearance of a forecast error (indeed, 66% of the total RH initialization errors occurred in the SC region). The much more frequent appearance of RH forecast errors over the SP states could not be explained so easily as an initialization-related problem. It is reasonable to assume that correction of the initialization and CISK-wave problems would lead to much less frequent RH (and LI) underforecast problems in both areas. Of course, other factors should be considered before a full explanation of the moisture underforecast problem can be realized. It is quite possible that lack of soil moisture effects in MASS 2.0 (see Table 1) significantly contributed to the forecast errors, for example.

7. SUMMARY OF PART I: SYNOPTIC-SCALE MODEL EVALUATION RESULTS

It has been shown that the large-scale circulation patterns at the surface are predicted by MASS 2.0 with a level of accuracy comparable to the LFM. Statistical evaluation of a nearly 30 case sample showed that the lower tropospheric mass fields were forecast significantly better by MASS 2.0 than by the LFM, but that the LFM outperformed MASS 2.0 in forecasts of upper tropospheric mass and momentum fields. Trends in the RMSE and BIAS statistics were used to identify the most serious systematic error in the MASS 2.0 forecasts, namely the loss of mass (anticyclone and 500 mb geopotential amplitude) over the eastern United States. Problems in specifying and/or formulating the eastern boundary conditions at times seemed to be the source of this error. The problem was seen to be situation-dependent, being serious only when a certain kind of synoptic airflow pattern (the "BC (bad case) regime") existed. After deleting BC regime days from the total sample, MASS 2.0 equaled or exceeded the performance of the LFM in synoptic-scale forecasts of nearly all fields studied at the 99% level of significance.

The LFM consistently underforecast 1000-500 mb thicknesses over the western half of the U.S. The systematically too cold LFM atmosphere was due to a combination of a positive bias in surface pressure and eastward phase errors in upper-level troughs and associated surface ridging. The greatest improvement made by MASS 2.0 over the LFM at the synoptic-scale was in its forecasts of the thickness field. It is suggested that increase in horizontal and vertical resolution, better numerics, and a more sophisticated planetary boundary layer parameterization scheme in MASS 2.0 led to the improved forecasts.

Three serious systematic errors in the MASS 2.0 forecasts were identified

and their causes investigated. The most serious error has already been mentioned, namely the loss of mass under the "BC" regime.

The second serious error is the interrelated overforecast of thickness values and surface cyclone intensity over the Plains states at the 24h verification period. In each case, these errors were linked to the omission of a cumulus parameterization scheme in the model. This can be briefly explained as follows. First, under conditions of strong normal flow to the Rocky Mountains and strong solar heating, a pronounced barrier-like flow around the heated obstacle would develop. The thermal plume and associated vorticity perturbation would be advected downstream, resulting in a mesoscale region of strong upward motion in the lee of the mountains. Precipitation developed when the atmosphere there was convectively unstable. Second, an unrealistic amount of latent heat would be released in a CISK-like process (Charney and Eliassen, 1964) as surface pressures fell and convergence intensified. Convective stabilization should have occurred as the result of convective overturning. However, this was impossible because of the lack of a cumulus parameterization scheme. Finally, the liberated heat resulted in a mesoscale region of overforecast thickness values and (from hydrostatic considerations) surface cyclones too deep and displaced erroneously towards the location of maximum thickness error. The early evolution of the thermally driven (mountain-valley) circulation is seemingly quite realistic, because in all but one case, strong convection actually did develop in close spatial and temporal proximity to the core of strong upward motion within the circulation. This indicates that barrier flow around the Rocky Mountains may actually be an important source mechanism for severe convective development.

The third systematic error studied was the frequent underforecast of

relative humidity and unstable lifted index values over the southern states. A study of all underforecast cases indicated that errors in the initialization of the moisture/instability fields was a leading cause of the underforecasts. The obvious question that needs to be addressed is whether this is only a perceived error: Is there really a problem in the initial representation of the moisture field, or rather is it the method that was used to determine relative humidity and lifted index from the basic model-predicted variables? If the error is really an initialization shortcoming, then these results suggest the real need for a better moisture data base over the northern Gulf of Mexico region than presently exists. Perhaps VAS "split window" data (Chesters, et al., 1982) can provide for this apparent need. Perkey (1976) found that when satellite cloud observations were used to enhance the initial moisture state in a mesoscale model, the forecast of convective precipitation was improved. This suggests that short-range moisture and precipitation forecasts are sensitive to moisture features that are below the scale resolvable by the current rawinsonde network.

The identification of important model systematic errors is extremely useful for guiding future MASS model development. The model can only be used intelligently for diagnostic, prognostic, and initialization purposes when its systematic character is kept in mind. Mesoscale features in the model, which are examined next, can be placed into the proper context now that the larger-scale nature of the model is understood.

PART II: EVALUATION OF MASS 2.0 MESOSCALE PREDICTABILITY

It is not possible to objectively verify the unfiltered, mesoscale MASS 2.0-predicted fields with the routinely collected upper air data. A more subjective approach is used to evaluate the mesoscale predictive capabilities of the model. First, examples of various mesoscale fields from 2 of the 30 cases will be presented to illustrate how coherent and useful the forecast fields are. Next, the basic forecast variables are combined in various ways to produce "convective predictor variables", which are then related to observed strong convection using two approaches. The first approach consists of a *diagnostic* study of the temporal continuity in the predictor variable fields and how well they relate to strong mesoscale convective systems (MCS) observed by radar. In this approach, the locus of predictor variables are verified against the locus of each of the 149 MCS's observed throughout the course of the entire 30 case sample. The second approach is *predictive* in nature, as an experienced severe storm forecaster used the MASS 2.0 fields to forecast severe convection in a nine day "real-time" operational setting. The forecaster issued "convective outlooks" using the LFM and MASS 2.0 forecast fields separately. Also, he issued "severe weather watch" boxes from MASS 2.0 information only. His forecasts were statistically verified against the SELS (Severe Environmental Local Storms) logs maintained at the National Severe Storms Forecast Center (NSSFC) in Kansas City, and were objectively compared to the outlooks issued by NSSFC.

The diagnostic MCS locus study and the prognostic experiment constitute a rather novel basis for evaluating the mesoscale predictability of the model. Regarding the diagnostic approach, it may be wondered why the loci of the

predicted and observed events were studied, rather than their areal distributions. An attempt was made to quantitatively compare the areal distribution of the convective predictor variables against the areal distribution of the observed MCS's. This method assumes knowledge of the threshold values for the predictor variable, which then permits an objective scoring of the forecasts using the Critical Success Index (CSI), defined by Donaldson et al. (1975). An attempt was made to use this method for about a third of the sample before it was decided that insurmountable problems prevented any firm conclusions from being reached. There are three problems with this approach:

(1) We did not know a posteriori from a large historical sample of past independent model runs which set of convective predictor variables was optimum for the MASS 2.0 model. Charba (1979) and Reap and Foster (1979) discuss the method whereby multiple screening regression is used to find the optimal set.

(2) The optimal threshold values for each of the individual convective predictor variables were likewise unavailable. For example, if lifted index is considered, is $LI < -2$ or $LI < -4$ a better threshold to use, based upon a past independent sample?

(3) Very minor forecast errors in the position of sharp moisture gradients (drylines) or in the timing of the onset of upward motion (error under 2h) resulted in very poor CSI values, yet there was obviously much useful predictive information present in the model fields. As an example, the 2 April 1982 case was one in which a one grid point (52 km) offset between upward motion and potential instability ($LI \leq 0$) resulted in a calculated $CSI = 0.0$. However, qualitatively a forecaster could have used the coherent predicted mesoscale fields with a high degree of success in issuing severe weather watches on this day (as will be shown in section 8a).

For these reasons, it was decided to abandon the areal forecast method in favor of the locus method. The diagnosed causes for both overprediction and underprediction of MCS's using the diagnostic MCS locus method were determined. Those particular aspects of the model-predicted fields that enabled the

real-time forecaster to improve upon his LFM forecasts, and conversely, those which were a detriment to him, were also examined. These two approaches are discussed in more detail following the presentation of the two case examples.

8. COHERENT MESOSCALE INFORMATION PROVIDED BY MASS 2.0 IN TWO CASE EXAMPLES

The MCS's were identified from the NAFAX Automated Radar Summary. The threshold value for an MCS used was the VIP3 intensity level (after Reap and Foster, 1979). At times, the visible and infrared GOES satellite images assisted in the determination of which radar cells constituted which system. In particular, identification of thunderstorm outflow boundaries, sea breezes, and other organizing mesoscale circulations was made possible with the satellite imagery.

The coherent structure prevalent in many of the MASS 2.0 unfiltered mesoscale forecast fields is demonstrated with two case examples. Comparison is made in both cases between the LFM and MASS 2.0 vertical motion fields. The first case (2 April 1982) is characteristic of the relative performance of the two models in very well organized, large-scale severe weather outbreaks. The second case (14 April 1982) involves a local (meso-alpha scale) outbreak of transient severe storms over Texas. This case is studied in greater depth than the first case to demonstrate the mesoscale phenomena that could be resolved in MASS 2.0.

a. 2 April 1982 Case

On this day, the central part of the United States suffered under an onslaught of 55 tornadoes. This outbreak was not only the most destructive and

widespread of 1982, but it was the biggest outbreak since April 3, 1974, when 148 tornadoes occurred (Ferguson et al., 1983).

The MASS 2.0 forecasts related remarkably well to the several MCS's observed that day. A short wave initially over northern Texas at 1200 GMT 2 April 1982 maintained continuity in the MASS 2.0 forecasts and could be related throughout the day to a severe storm complex as it progressed from Texas into northeastern Arkansas by 2300 GMT. A second MCS took the form of a solid squall-line by 1900 GMT along a well-predicted dryline convergence zone throughout central Kansas and Oklahoma. The third MCS identified on this day developed as a severe storm cluster around 1900 GMT in southeastern South Dakota just northeast of the strong surface low.

The LFM and MASS 2.0 vertical motion fields forecast for 0000 GMT 3 April 1982 are shown in Fig. 14 along with the GOES-E visible satellite image taken 90 min before. The MASS 2.0 upward motion patterns accurately (± 250 km and ± 2 h) depict the positions of: (1) the MCS over northeastern Arkansas (upward motion center over southeastern Arkansas) which apparently was triggered by the Texas short wave; (2) the squall-line stretching from Iowa, through Missouri, western Arkansas, and into extreme southeastern Oklahoma, which formed earlier within the dryline convergence zone; and (3) the strong MCS over extreme southwestern Minnesota associated with the circulation about the surface low in eastern Nebraska. Notice also that the model predicts strong upward motion over Illinois. Moderate convection is occurring at the time over eastern Illinois, distinctly separate from the Iowa squall-line. Weaker instability (forecast and observed) in Illinois could not support the existence of strong thunderstorms, which the upward motion field would seem to imply if considered in isolation. Even weaker convection seems associated with the band of



Figure 14. Comparison of 12h forecast of 700 mb vertical motion (upward, dashed at intervals of $2 \mu\text{bar s}^{-1}$) by the IFM (left) and MASS 2.0 (center) models. Also shown is verifying GOES-E visible satellite image. Model forecasts verify at 0000 GMT 3 April 1982, satellite image is for 2230 GMT 2 April 1982.

moderate upward motion over Lake Superior. Finally, we note that in western Kansas light showers are occurring in apparent association with the weak upward motion forecast by the model over the Kansas-Oklahoma border. The shape of the forecast upward motion field bears an outstanding resemblance to the observed pattern of convection. Equally as important is that the forecast subsidence zone that stretches in a curved fashion (with imbedded mesoscale maxima) from eastern Colorado-New Mexico, through most of Oklahoma, and into extreme eastern Kansas and southwestern Iowa is virtually identical in shape to the observed duststreaks in the satellite images. The presence of such duststreaks is used to infer the strong downward turbulent transport of jet stream momentum to the surface layer (e.g., Koch and McCarthy, 1982).

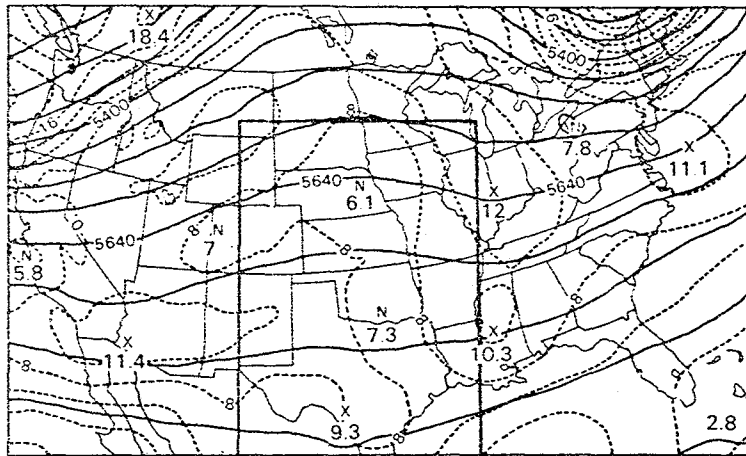
The highly informative and temporally coherent mesoscale aspects of the MASS 2.0 forecast were also compared to the corresponding LFM vertical motion forecast. Of course, the LFM cannot be expected to resolve such mesoscale systems. Indeed it only shows a large-scale dipole of vertical motion, indicating that an outbreak of strong convection would be likely somewhere over the central United States. Notice also that the LFM has a significant eastward phase error in its forecast upward motion center (if one were to grossly smooth out the MASS 2.0 upward motions, a maximum about 600 km west of the LFM center would likely result).

Thus even in such a large-scale, well-organized severe weather outbreak case as this, MASS 2.0 forecast fields show significantly more useful information for forecasting strong MCS's than does the LFM. Such comparisons between the two models were not made on a case-by-case basis, although an occasional check (such as in the 14 April 1982 case) was made.

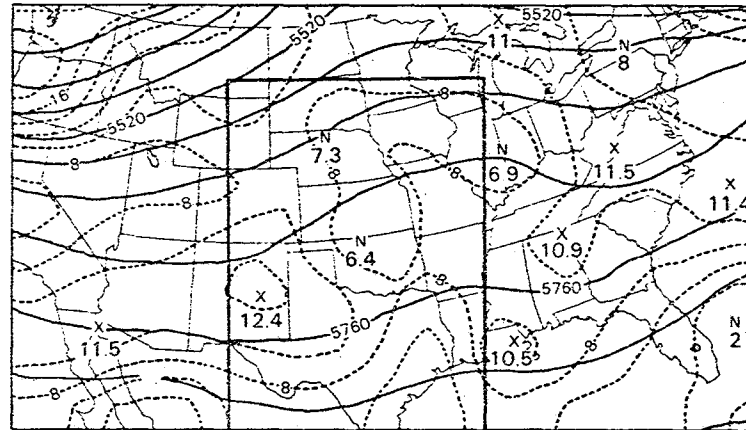
b. 14 April 1982 Case

Both the previous case and the 14 April case were ones in which the synoptic-scale performance by MASS 2.0 was superior to that of the LFM (Figs. 4 and 5), and therefore ones in which there is a sound justification for looking at the unfiltered mesoscale fields. This justification is all the more important in the 14 April case, because the apparent forcing of the Texas MCS is a very small-scale feature in the almost zonal 500 mb flow (Fig. 15). During the period from 1200 GMT 14 April to 0000 GMT 15 April, a weak perturbation in this flow moves eastward from western Arizona to western Texas. An associated vorticity maximum of $11.4 \times 10^{-5} \text{ s}^{-1}$ is found at 12Z in western Arizona. This same feature is observed in western Texas at 00Z ($12.4 \times 10^{-5} \text{ s}^{-1}$). This feature is lacking in the LFM forecast, but appears as a 10.4 center in the *filtered* MASS 2.0 forecast. Also of interest is the observed axis of weak vorticity in eastern Texas-Oklahoma at 00Z. Notice that the LFM forecast does not clearly show this feature either. MASS 2.0 shows the feature about 400 km west of the observed axis. In effect, the observed distance between the vorticity maximum and minimum reduces to less than half its initial value by 00Z; MASS 2.0 overforecasts the reduction somewhat.

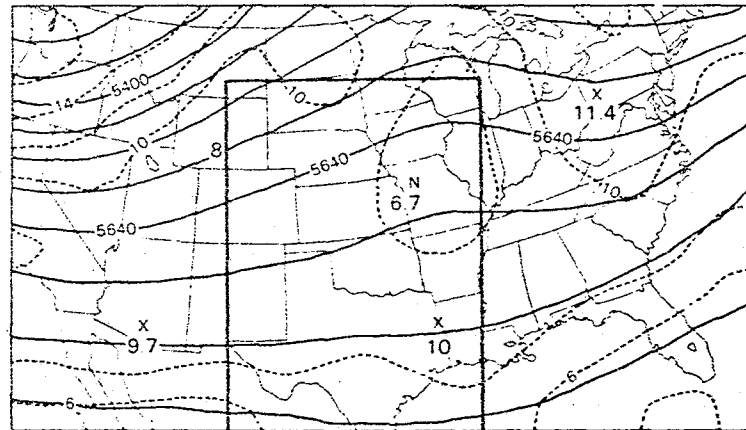
The sequence of *unfiltered* MASS 2.0 500 mb forecasts at 6h, 10h, and 12h into the forecast period (fig. 16) shows the evolution of these features. The vorticity maximum quickly translates to the New Mexico-Texas state border in only 6h and then remains essentially anchored there as it spreads latitudinally. During the same period the axis of weak vorticity drifts slowly eastward. Meanwhile, a new feature appears suddenly at 2200 GMT just east of the stationary vorticity maximum. This axis of very low vorticity stretches all along the High Plains from South Dakota to west Texas and is associated



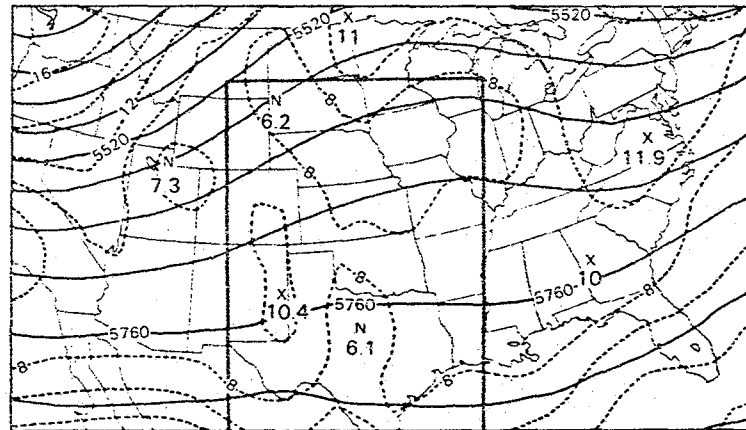
INITIAL ANALYSIS AT 1200Z APRIL 14, 1982



VERIFICATION ANALYSIS AT 00Z APRIL 15, 1982

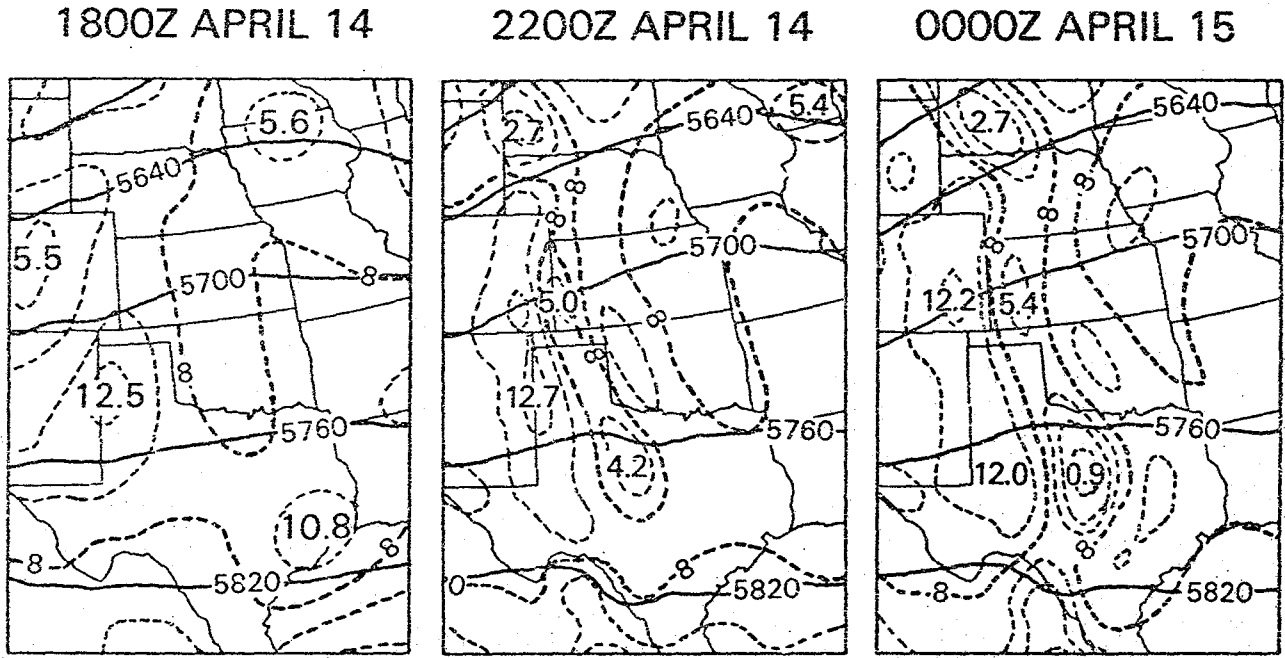


12 H LFM FORECAST FOR 00Z APRIL 15, 1982



12 H FILTERED MASS 2.0 FORECAST FOR 00Z APRIL 15, 1982

Figure 15. Comparison of 12h LFM and filtered MASS 2.0 forecasts of 500 mb geopotential height (solid lines, CI = 60m) and absolute vorticity (dashed lines, CI = $2 \times 10^{-5} s^{-1}$) from initial analysis at 1200 GMT 14 April 1982. Also shown is verification analysis at 0000 GMT 15 April 1982. Box over center portion encloses sub-synoptic area of in-depth study in succeeding figures.



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 16. Sequence of unfiltered MASS 2.0 forecasts of 500 mb geopotential height and vorticity on 14 April 1982. Same format as in Fig. 15.

with a small meso-alpha (200 km half-wavelength) wave in the geopotential field, which is most pronounced in west Texas. Notice that this meso-alpha ridge is a new feature developed by MASS 2.0.

The sequence of 700 mb vertical motion forecasts (Fig. 17) at the same intervals as in Fig. 16 shows that an organized zone of moderately strong upward motion is found already at 18Z. However, the magnitudes of the individual maxima more than double by the time of appearance of the meso-alpha feature. The accompanying radar data show that thunderstorms began forming in Texas by about 21Z at the exact location of strongest predicted mesoscale forcing. These storms grow to severe limits by 0135Z 15 April, and new storms form in western Kansas and central Nebraska between 2335Z and 0135Z (see Fig. 21d). Three separate MCS's can be identified: the Texas complex, the western Kansas-central Nebraska complex, and a weakening system which propagates from eastern Nebraska at 1735Z to southern Minnesota by 2135Z. The model upward motion fields show three separate maxima that agree in position rather well with these three systems. Two of the three maxima are associated with the meso-alpha feature at 500 mb.

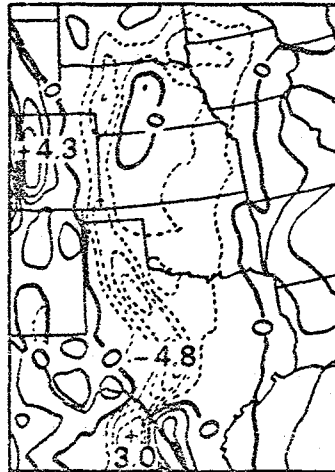
The LFM vertical motion forecast field is devoid of any useful detail, and just as in the 2 April 1982 case, has an eastward phase error (Fig. 18). Since the observed storms formed further west near the LFM's $\omega = 0$ isoline, this error is considered significant. Recall that the LFM tended to displace surface pressure and 500 mb vorticity features too far eastward over the western half of the country (Tables 7 and 8).

1. *Meso-Alpha Scale Mountain Wave*

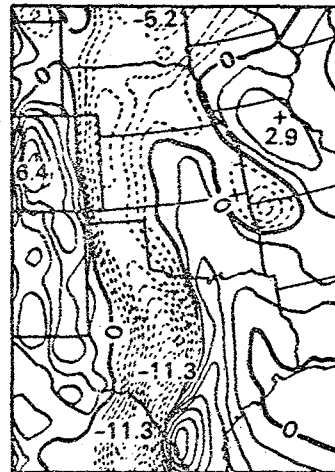
The excellent comparison between the observed MCS's and the predicted

700 MB
OMEGA
FORECASTS

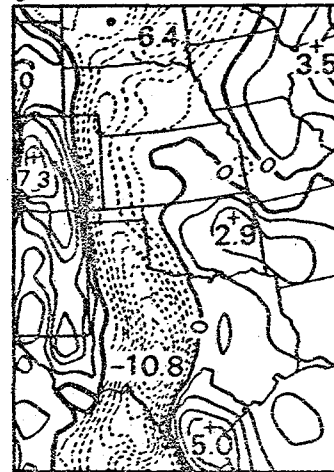
1800Z APRIL 14



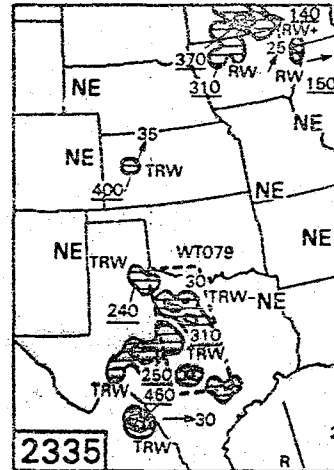
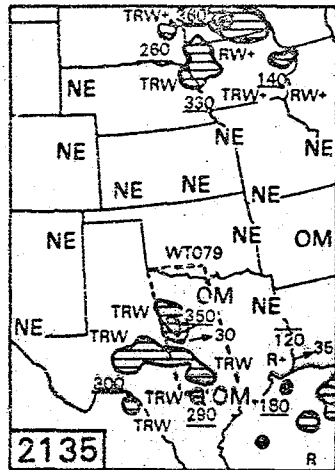
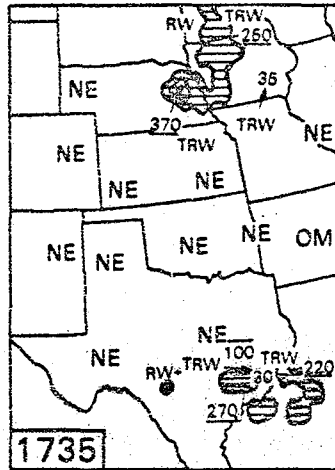
2200Z APRIL 14



0000Z APRIL 15



RADAR
SUMMARIES



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 17. Sequence of MASS 2.0 700 mb vertical motion fields and NAFAX Automated Radar Summaries on 14 April 1982. Vertical motions contoured at intervals of $2 \mu\text{bar s}^{-1}$ (dashed = upward). Radar echo tops (underlined) reported in $\text{ft} \times 10^2$.



Figure 16. Comparison of 12h forecast of 700 mb vertical motion (upward, dashed at intervals of 2 mb/s) by the IEM (left) and NARR 2.0 (right) models, verified at 0000 GMT 15 April 1982.

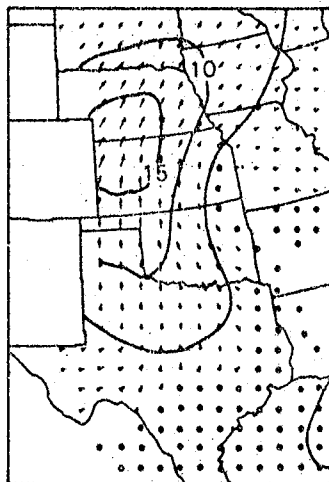
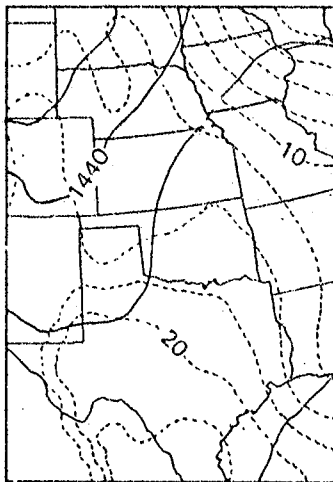
ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL AGE 19
OF POOL QUALITY

upward motion fields on 14 April 1982 gives strong indication that the meso-alpha 500 mb ridge that suddenly appears just before 2200 GMT is to some extent real. The possible origin for, and nature of, this feature was examined. Its stationary presence in the immediate lee of the northern Mexican Plateau suggests the possibility of its being a large mountain wave. The necessary conditions for large-scale mountain waves were present, namely strong tropospheric winds blowing perpendicular to the large mountain barrier and low static stability below mountain tops. These conditions were most evident over Texas, where the barrier half-width is quite large.³ A very similar feature has been observed by Anthes et al. (1982) in the PENN model simulation of the 10-11 April 1979 situation over west Texas and shown to be in good agreement with the theoretical results of Klemp and Lilly (1975) for a ramp-shaped mountain (i.e., plateau idealization). In the simulations from both the PENN and MASS 2.0 models, a vertical motion couplet is observed, with the upward branch of the circulation located immediately to the lee of the barrier. Other similarities are found in the lower troposphere (at the 850 mb level), including (a) a strong increase in westerly momentum beneath the downward branch of the standing mountain wave, and the development at the leading edge of the downward branch of the wave circulation of: (b) a pronounced potential temperature maximum, and (c) a lee trough. All of these features are apparent in the 00Z 15 April analyses at the 850 mb level in Fig. 19. Notice the strong warming, height (pressure) falls, and increase in westerly momentum that occurred in west Texas during the preceding 12h.

³The mountain wave concept may also be applicable over eastern Colorado, because the half-wavelength (the distance between vorticity maximum and minimum in Fig. 16b) decreases northward. The effective half-width of the mountain barrier is also much smaller in Colorado.

1200Z
APRIL 14,
1982



0000Z
APRIL 15,
1982

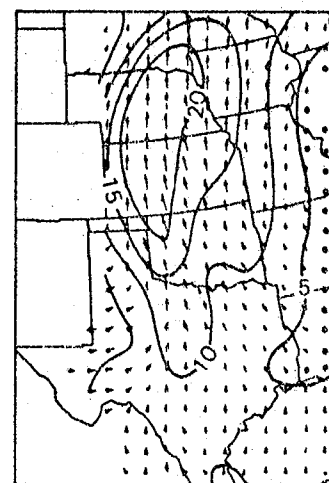
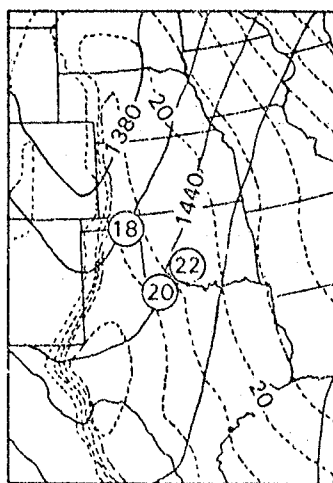


Figure 19. Initial (top) and 24h (bottom) forecasts of 850 mb temperature (left, dashed), geopotential height (left, solid), and wind vectors and isotachs (right, solid). Also shown are two-hourly locations of maximum surface pressure falls, with circled numbers denoting time in GMT. Isotherm interval is 2.5C, geopotential height interval is 30 m, and isotach interval is 5 m s⁻¹.

2. Low-Level Jet

Significant increase in the speed and backing in the direction of the 850 mb jet winds occurred during the daytime, particularly over western Nebraska and Kansas (Fig. 19). Notice that during the same period, the height of the 850 mb surface fell by as much as 60m in extreme eastern Colorado and

southeastern Wyoming. Thus, the changes in the lower tropospheric winds seem to be largely a response of the atmosphere to a strong isallobaric effect. The strong PBL heating over the High Plains of Texas and Oklahoma and the resulting warm advection maximized from western Kansas to western Nebraska during the day must have contributed strongly to the drop in 850 mb heights in the region.

3. *Westerly Momentum Surge and Dryline Bulge*

The pressure falls at the surface occurred southward of the 850 mb isallobaric fall center, being largest over the eastern Texas Panhandle and southwestern Oklahoma (Fig. 19). A surge of westerly momentum at 850 mb had developed by 12h into the model forecast in southwestern Texas. This is only about 400 km south of the afternoon surface isallobaric fall center. At the head of this westerly momentum surge, the model formed a very pronounced eastward bulge in the surface dryline (Fig. 20b). Convergence became strongest at this dryline bulge during this time (Fig. 20c). Recall that it is precisely at this location that the Texas MCS first began forming (Fig. 17).

These inter-relationships between a surface isallobaric fall center, the maximum westerly surge in the lower troposphere, the development of a dryline bulge, and the preferential formation of severe thunderstorms at the bulge has been observed by McCarthy and Koch (1982) in a separate case study. In both that case and this one, there is a suggestion that the rapid deepening of the PBL west of the dryline permitted higher westerly momentum to be transported downward to the very low troposphere (the development of the mountain wave would reinforce this transport process). The local increase of westerly momentum would force the dryline further eastward at the bulge. Although the highest PBL development occurred throughout a large area from Mexico to northern Colorado (fig. 20a), the presence of an upper-level jet streak over

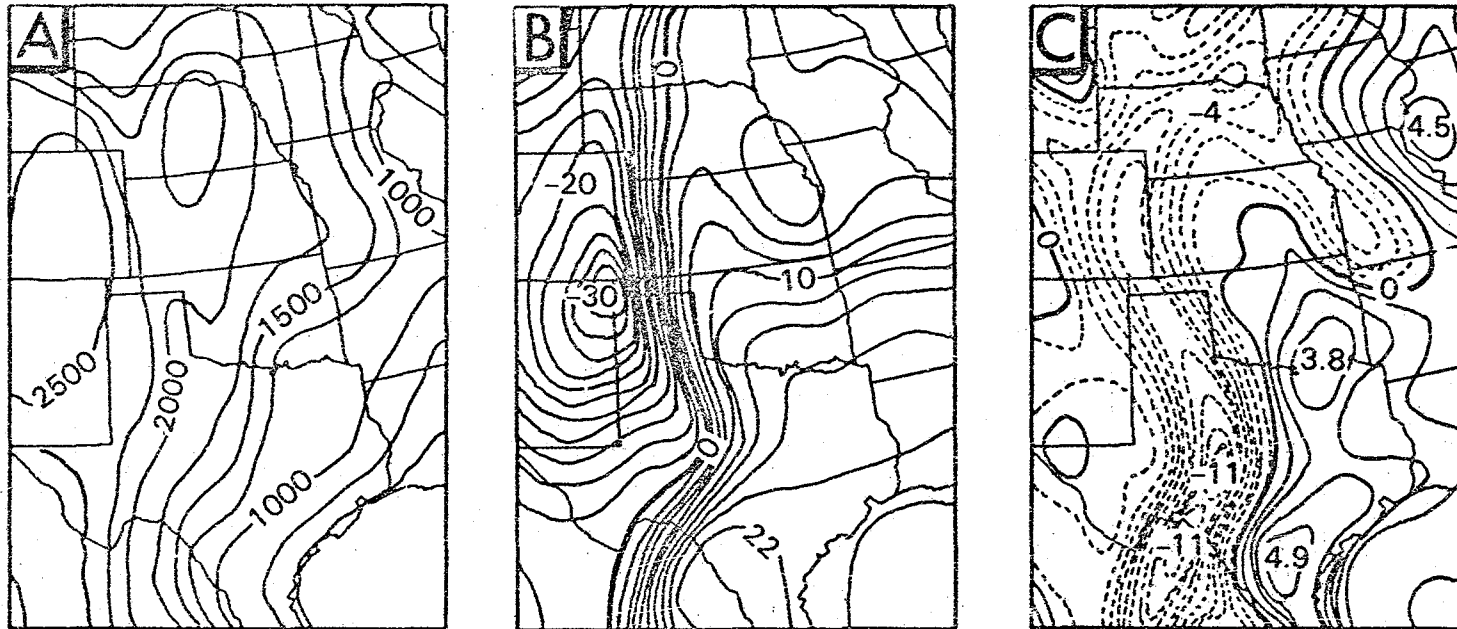


Figure 20. Model 12h forecasts verifying at 0000 GMT 15 April 1982 of (a) planetary boundary layer height (in m), (b) dew point temperature ($^{\circ}\text{C}$) at the lowest level ($\sigma = 0.96$) in the model ("surface"), and (c) "surface" wind divergence (10^{-5} s^{-1}) with convergence dashed.

ORIGINAL PAGE IS
OF POOR QUALITY

southeastern New Mexico at the time (not shown) supplied the best source for the westerly momentum. *This examination suggests that the sophisticated PBL parameterization and high resolution grid used in MASS 2.0 were crucial factors in the model which permitted it to predict an apparent mountain wave, the intensification of the southerly low-level jet, the dryline bulge, and the strong lifting apparently necessary for the formation of the dryline storms.*

9. GENERAL MASS 2.0 MESOSCALE PREDICTIVE CAPABILITIES

a. *Diagnostic MCS Locus Approach*

1. *Methodology*

It was decided to evaluate MASS 2.0 in terms of its ability to forecast the correct location and time of observed (a) MCS formation and (b) MCS maximum intensity. Primary emphasis is placed on MCS genesis because it was not expected that the model would handle well the further evolution of deep convection without having a convective parameterization scheme. The locus of a convective predictor variable field verifies against the locus of the radar-observed MCS only if (a) the field exhibits acceptable continuity (a subjective judgment), (b) the forecast genesis (or maximum intensity) time is within ± 3 h of the observed time, and (c) the forecast position is no more than 250 km beyond the observed MCS location at the time of forecast genesis (or maximum intensity). The forecast verification criterion for accurate timing results from the fact that the MASS 2.0 derived variable fields were mapped at 2h intervals. The spatial accuracy criterion has a less objective basis, but from our experience seems to represent the minimum scale of spatial information of the forecast fields.

The choice of convective predictor variable fields used was guided by the

ORIGINAL PAGE IS
OF POOR QUALITY

results of various short-range objective forecast methods reported upon in the literature. These schemes combine variables derived from conventional data with those derived from the LFM model and/or climatology of severe weather events (Charba, 1979; Wilson and Turner, 1982). The most successful variables in such schemes fall under the general categories of some kind of stability index, some measure of the lower tropospheric moisture convergence, and an indicator of the horizontal gradient of moist static stability.

These three facets are incorporated into a planetary boundary layer (PBL) variable abbreviated FCE for Flux Convergence of available moist static Energy. Mathematically, FCE is expressed as

$$FCE = c_p \bar{\nabla} \cdot (\bar{\mathbf{V}}_{PBL} \partial \theta_e / \partial z) \quad (3)$$

$$= c_p \bar{\mathbf{V}}_{PBL} \cdot \bar{\nabla} (\partial \theta_e / \partial z) + c_p (\partial \theta_e / \partial z) \bar{\nabla} \cdot \bar{\mathbf{V}}_{PBL} \quad (4)$$

where $\bar{\mathbf{V}}_{PBL}$ is the PBL layer-averaged horizontal wind vector and $\partial \theta_e / \partial z$ is the equivalent potential temperature differential over the $\sigma = 0.89$ to 0.96 layer. The FCE index represents the amount of forecast convective instability, which can be realized through boundary-layer lifting processes forecast by the model (Koch and McCarthy, 1982). Low-level wind convergence within a potentially unstable air mass and horizontal advection of moist static instability are the two component processes that can contribute to a net positive FCE effect. An $FCE \geq 2 \text{ W kg}^{-1} \text{ m}^{-1}$ is the forecast threshold value.

The second convective predictor variable used is abbreviated WLI for forecast 700 mb upward (w) motion of at least 2 cm s^{-1} within an air mass whose Lifted Index is negative. In practice, the two component fields of the WLI

ORIGINAL PAGE IS
OF POOR QUALITY

index were studied to find the overlap region between $-LI > 0$ and $-w \geq 2$. The WLI index describes forecast regions where the atmosphere is conditionally and potentially unstable, and where there is significant mid-tropospheric lifting which can make the potential instability available. Matthews and Silverman (1980) have shown that greater mesoscale lifting in a potentially unstable atmosphere produces deeper convective cloud development in the Kreitzberg and Perkey (1976) one dimensional MESOCU cloud model. The WLI index is somewhat similar to the second term in equation (4), with the exception that it indicates the strength of the lifting at a higher level in the model atmosphere (thus, only lifting throughout a deeper layer will be detected by the WLI index).

The third and last convective predictor variable used is abbreviated WA for forecast 850 mb Warm Advection which is considered if and only if the area of most pronounced WA is located downwind of a potentially unstable ($LI < 0$) region. In effect, the WA index was used only as a last resort if the other two indices failed to relate well diagnostically to the observed thunderstorms, because it could only be used in a qualitative sense in this experiment. Maddox and Doswell (1982) have suggested that the WA index may serve as an aid to severe weather forecasters (apparently overlooked in the past). The WA index is based upon one of the two terms in the quasi-geostrophic omega (vertical motion) equation, namely the horizontal Laplacian of the thickness advection. In the absence of the other term (differential vorticity advection), quasi-geostrophic rising motion will generally occur in association with warm advection regions, thus near favorable thermal boundaries in the lower troposphere (Holton, 1972). It is necessary to have unstable air upstream of this localized lifting region to have convective potential. Of

ORIGINAL PAGE IS
OF POOR QUALITY

course, MASS 2.0 vertical motion fields would be the result of not just this one quasi-geostrophic effect, but also the result of differential vorticity advection and non-quasi-geostrophic effects. The idea is to isolate one physical mechanism, as the others (perhaps forecast incorrectly) could counteract the effect.

Although all three of these predictor variables have sound dynamical bases, none of them explicitly and unambiguously indicates that the model is predicting deep convection at any given grid point. It is easy to lose sight of what model dynamics are actually occurring when sole reliance is placed upon one or two predictor variable fields. For example, all three variables relied upon here "overpredict" thunderstorm occurrence in cases where strong lower tropospheric temperature inversions (termed "lids" by Carlson *et al.*, (1980)) prohibit the release of potential instability. Thus, even though the model-predicted FCE_m show a region of very large positive values indicating the tendency of PBL energy convergence to deepen and destabilize the PBL, the model may also be predicting a strong "lid". Unfortunately, we have no knowledge of actual model-predicted "lid" strength from the model output fields which were examined.⁴ *Thus, it is important to realize that the three predictor variables used during the course of our evaluation to indicate the location, timing, and intensity of model-predicted strong thunderstorms are presumptive*

⁴Perhaps the best and least ambiguous indicator of where and when a mesoscale model is predicting the onset of deep convection is some direct measure of the positive buoyant energy at any grid point where the model predicts that a surface air parcel has attained its level of free convection. The amount of such energy that exists above this level, which can be interpreted as the positive area on a thermodynamic diagram, is then "available" to the air parcel. It is only after at least some air parcels within a grid column have buoyant energy made available to them that both the Kreitzberg and Perkey (1976) and the Fritsch and Chappell (1980) models produce convection.

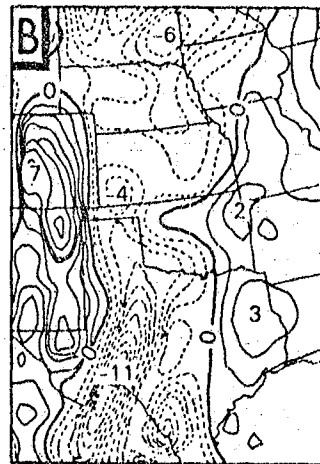
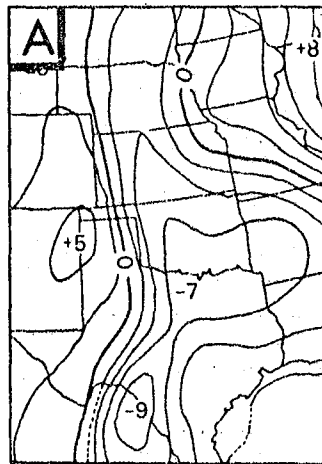
in nature, i.e., only indirect indications rather than definite, direct measures of deep convection forecast by the model.

2. Results of Diagnostic MCS Locus Study

The forecast fields from the 14 April 1982 case are used to demonstrate how the WLI convective predictor variable field was determined and how both its locus and that of the verifying MCS were obtained and related. Figs. 21a and b show the two constituent fields of WLI, namely lifted index (LI) and 700 mb vertical motion (ω). At this time, there is a large unstable region with imbedded maxima over extreme northern Texas (LI = -7) and the Rio Grande River (LI = -9). The strongest upward motion is occurring over southwestern Texas ($\omega = -11$), and smaller maxima are also evident in southwestern Kansas ($\omega = -4$) and southeastern South Dakota ($\omega = -6$). Since most of the Great Plains is convectively unstable, more emphasis is placed on the three upward motion centers than on the lifted index minima. Thus, the predictor field shows three loci with two of them located at two of the ω maxima, and the third displaced slightly away from the Texas ω maximum and towards the average of the two LI maxima. Fig. 21c depicts these three WLI loci, the field coverage of the WLI field, and the loci of the three MCS's described earlier (and seen in Fig. 21d). Each MCS relates well to a nearby predictor variable locus; in fact, the 250 km/2h forecast verification criteria were satisfied for each of the MCS's both at their predicted time of formation and at their predicted time of maximum intensity.

This procedure was followed using the WLI, FCE, and WA predictor variables in a total of 149 cases of observed MCS's throughout the course of the 30-day experiment. The results in Table 11 show that about 50% of the observed MCS

ORIGINAL PAGE IS
OF POOR QUALITY



(N) LOCUS OF MCS #N
(N) PREDICTOR FIELD LOCUS

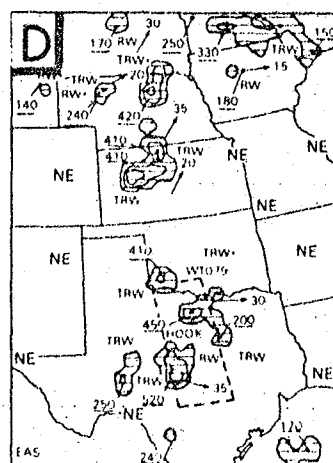
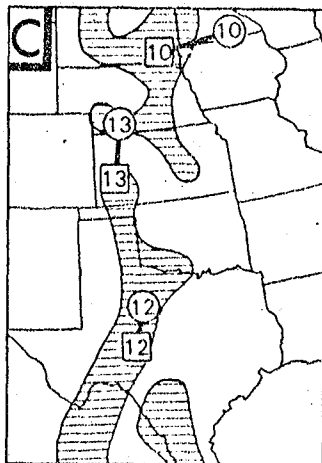


Figure 21. Constituent fields of model convective predictor variable field WLI and comparison with observed MCS loci at 0200 GMT 15 April 1982: (a) lifted index forecast, (b) 700 mb vertical motion (upwards is dashed, in wbar s^{-1}), (c) WLI field produced from overlap of 700 mb -2 wbar s^{-1} and lifted index ± 0 areas, with numbered loci of WLI and corresponding MCS (forecast spatial offset shown by line segments), and (d) 0135 GMT radar summary (storm tops in 10^2 ft) showing distribution of the three MCS's.

TABLE 11. SUMMARY OF DIAGNOSTIC MCS LOCUS STUDY.

Symbol	Evaluation Statistic	Time of Evaluation	
		At Forecast MCS Genesis	At Forecast MCS Maximum Intensity
x	TOTAL MCS CORRECTLY PREDICTED	75	76
y	TOTAL OBSERVED MCS NOT PREDICTED	74	71
z	TOTAL PREDICTED MCS NOT OBSERVED	11	13
POD	$x/(x + y)$	0.50	0.52
FAR	$z/(x + z)$	0.13	0.15
CSI	$x/(x + y + z)$	0.47	0.48
BIAS	$(x + z)/(x + y)$	0.58	0.61

loci were predicted accurately by the MASS 2.0 mesoscale convective predictor variable fields, and that very few systems were overpredicted (about 14%). The most successful predictor variable fields were WLI and FCE, which correctly predicted MCS's 39 (42%) of the time, respectively, whereas WA proved useful as a standby only 4% of the time. The net result of a moderately high POD and a very low FAR is a respectable CSI of 0.47 - 0.48 (these quantities are defined in section 9b).

The reason why the FAR is so low must be fully appreciated. This locus method really boils down to a "point forecast". Verification of forecasts of areal distribution would result in a much higher FAR, even if the optimum threshold values for the optimum set of convective predictor variables were known. These results should be recalled when the results of the real-time forecast experiment are presented, in which it will be seen that the FAR value there of 0.58 represents a more typical value both for objective techniques that make areal forecasts (Charba, 1979) and for SELS (Pearson and Weiss, 1979; Weiss, 1977). *What the statistics in Table 11 really show is that MASS 2.0 mesoscale forecasts yield useful, coherent fields that can be related about 50% of the time to observed convective systems, and that the noise level of the forecasts (FAR) is generally very low. The POD value (0.50) is very respectable, as compared to that reported by SELS for its convective outlooks.*

The reason for both the underpredictions (y) and the infrequent overpredictions (z) are categorized by importance in Table 12. Incorrectly forecast "short wave" disturbances in the upper- and mid-troposphere were the leading cause of both over- and under-predictions. The underlying causes for these bad forecasts were not studied.

The next most important causes for convection *overpredictions* were the

TABLE 12. LEADING CAUSES FOR MISSED FORECASTS OF MCS's.

ORDER OF IMPORTANCE	DESCRIPTIVE CAUSE	NUMBER OF OCCURRENCES	
		TIME OF FORECAST MCS GENESIS	TIME OF FORECAST MCS MAXIMUM INTENSITY
OVERPREDICTIONS			
1	"SHORT WAVE" POORLY FORECAST	3	5
2	STRONG DRYLINE CONVERGENCE ("LID"?)	3	3
3	OVERFORECAST INSTABILITY	3	3
4	PROBLEMATIC CISK-WAVE	<u>2</u>	<u>2</u>
	TOTAL	11	13
UNDERPREDICTIONS			
1	"SHORT WAVE" POORLY FORECAST	18	20
2	FRONTAL CONVERGENCE TOO WEAK	18	14
3	UNDERFORECAST INSTABILITY	13	12
4	OUTFLOW BOUNDARIES (MCC) NOT FORECAST	9	7
5	PROBLEMATIC FLORIDA SEA BREEZE	4	4
6	FRONTAL BOUNDARY POSITIONED WRONGLY	4	3
7	OTHER ERRORS	<u>13</u>	<u>15</u>
	TOTAL	79	75

"over-predictions" of the magnitudes for dryline convergence and instability. As discussed earlier, the prediction of strong dryline convergence over the High Plains is not necessarily erroneous, because we had no way of knowing whether the model was also producing a capping inversion ("lid"). Overforecast instability is not considered a major problem, because the relative frequency of *underforecast* instability is four times as large. The least important cause for forecasts of MCS's not observed was the presence of model-predicted CISK-waves resulting from the lack of a cumulus parameterization scheme (diagnosed error E2 in section 6).

The second most important cause for convection *underpredictions* was weaker-than-observed frontal intensity (low-level convergence). This problem and the problem of incorrectly forecast positions of frontal boundaries (cause #6) were both the result, in most instances, of underforecast surface high pressure amplitude over the eastern part of the country. Recall that MASS 2.0 displayed a serious systematic bias there in its synoptic-scale forecasts of high pressure amplitude and position (diagnosed error E1 in section 6). The third most important source for missed MCS forecasts was the systematic model tendency to underforecast convective instability over the southern states. Recall that poor moisture initialization was the most important underlying cause of instability underforecasts (diagnosed error E3 in section 6). *Thus, we see that major synoptic-scale errors E1 and E3 had a significant impact on the mesoscale predictability of the model (causing reduction in the probability of detection of convective systems). The lack of a cumulus parameterization scheme further caused several underforecasts of MCS's as thunderstorm outflow*

*boundaries, mesoscale convective complexes, and Florida sea breeze thunderstorms could not form in the model.*⁵

Lastly, we briefly examined the temporal variations in MASS 2.0 mesoscale predictability, in order to address two important questions. The first question is "Can a mesoscale prediction model maintain a nearly constant level of predictability from spring to summer?" It is well known, for example, that SELS severe storm predictability (Pearson and Weiss, 1979) and LFM predictability drop dramatically from spring to summer as synoptic-scale systems weaken considerably. The second question is "Did the mesoscale information in the forecast fields suffer significantly during the model's period of poorest synoptic-scale performance?" This question is concerned with the fact that MASS 2.0 synoptic-scale forecasts suffered during "BC regime days" particularly prevalent during May (section 4c).

In Table 13, the temporal variation in MASS 2.0 mesoscale performance is simply divided into three time intervals, each composed of 10 cases, to smooth out daily fluctuations and to confine most of the "BC regime" cases to one interval. Note that an increasing number of MCS's were observed from April to July, a very typical climatological trend. The model's performance as measured by the CSI score shows a pronounced drop in predictability after early June. This drop is the result of a reduction in POD (see equation (8)), as the FAR does not increase (in fact, it drops slightly). A drop in POD means that fewer observed MCS's were "detected" by the model.

⁵This is not to say that Florida sea breezes did not form in the model. On the contrary, there was evidence that this phenomenon was frequently forecast by MASS 2.0.

C-2

TABLE 13. TEMPORAL VARIATION OF MCS FORECAST EVALUATION STATISTICS.

CASE PERIOD	CASE NUMBER	% "BC" CASES IN PERIOD	TIME OF FORECAST MCS GENESIS			TIME OF FORECAST MAXIMUM MCS INTENSITY				
			MCS OCCURRENCES (x+y)	FAR	POD	CSI	MCS OCCURRENCES (x+y)	FAR	POD	CSI
2 APR-3 MAY	1-10	10	32	0.10	0.56	0.53	32	0.10	0.56	0.53
4 MAY-7 JUN	11-20	70	46	0.14	0.67	0.60	44	0.21	0.59	0.51
8 JUN-2 JUL	21-30	40	71	0.13	0.37	0.35	71	0.11	0.45	0.43

85

ORIGINAL PAGE IS
OF POOR QUALITY

Referring to Table 12 again, we see that there are many possible sources for such underpredictions. In order that the leading causes for this modest drop in model performance level can be determined, we first of all recognize that causes #2 and #6 in Table 12 ("BC regime"-related) are not likely to be prominent. This is because Table 13 shows that model performance did not suffer most during the (second) period of most frequent "BC regimes". In fact, the mesoscale forecast fields displayed the highest POD and CSI skill score during the period of worst synoptic-scale performance by the model! The reason for this is that the "BC regime" (trough over western Atlantic Ocean and a pronounced ridge over the eastern U.S.) is not one which is highly conducive to the actual occurrence of deep convection in the east where the systematic errors were pronounced.

Careful inspection of the individual cases constituting the 8 June-2 July interval showed that the entire spectrum of causes shown in Table 12 contributed to the increased number of underpredictions during that period. There are two interesting points, however. First, there were very few cases related to underforecast instability during this period. On the other hand, nearly all of the underforecast problems related to the lack of a cumulus parameterization scheme (causes #4 and #5 in Table 12) occurred during this period. *Thus, it is unquestionably true that continuation of high levels of mesoscale predictability from spring into summer hinges upon the level of sophistication of the parameterized physics in the moist convective adjustment scheme of the mesoscale model.*

b. *Prognostic Real-Time Experiment Approach*

1. *Methodology*

Originally, the hope was to have a team of experienced, recognized severe weather forecasters make real-time forecasts at LaRC with the MASS 2.0 model output in an operational-like setting. However, only one such forecaster could be obtained for the experiment because of scheduling conflicts and other problems. Two means were found for controlling the deleterious influence of personal bias and limited experience factors, which would otherwise obscure the meaningful results of the experiment. First, the forecaster was asked to make two separate forecasts, one based first upon LFM output and then one based upon MASS 2.0 output. The forecaster was also permitted to see and to analyze conventional weather data up to, but not after, 1200 EDT. Any differences between the two forecasts are then due essentially to an impact made by MASS 2.0 upon the forecaster's earlier LFM-based forecast. Second, the two model-based forecasts were compared objectively (in terms of the CSI score explained below) to those issued by NSSFC. This procedure is used as a benchmark by which the effects of the forecaster's personal bias and his limited experience could be measured.

The NSSFC forecasters have experience with a wide range of severe storm-producing phenomena over the entire country, whereas the forecaster in the model experiment had much more limited experience. Mr. Don Burgess, who has for several years acted as a chief forecaster for the various spring field programs of the National Severe Storms Laboratory, participated as the sole "operational forecaster" in the experiment. Mr. Burgess has naturally been more concerned with severe weather situations in the Southern Plains than elsewhere, a factor which did reduce his effectiveness elsewhere.

The experiment was conducted during the period 21 June 1982 to 29 June 1982 at NASA/LaRC. The forecaster first used the LFM (plus limited conventional data) to issue a severe storm "convective outlook" (Pearson and Weiss, 1979) for the day by 1500 EDT. Then MASS 2.0 was used to produce another convective outlook, but sometimes not until 1200 EDT the next day because of computer related problems at LaRC. However, Burgess never had knowledge of any weather data or events after 1200 EDT on the forecast day. Both outlooks were valid from 1800 GMT on the forecast day to 1200 GMT the following day. In actuality, the forecaster issued "severe weather watch boxes" (Pearson and Weiss, 1979) with MASS 2.0 output and then constructed a convective outlook area from the perimeter enclosing adjoining watch boxes. Watch/outlook areas which were spatially separated on any given day were counted as separate areas.

Some of the case days could not be included in the statistical sample. One of Burgess' most favored predictor variable fields was the surface moisture convergence, which was not output on 21 June; this day was excluded because of the handicap. The 22 June model run did not use the customary 1200 GMT database as input because of problems with the WPRS datalink. Rather, the database from 0000 GMT the evening before was used; so, this day was also excluded from the sample. Finally, the 28 June run was discarded because the significant level rawinsonde data were not available for model initialization. It is thought that use of this data gives the MASS 2.0 system a decided advantage over the LFM. Consequently, only six days remained in the sample. Yet, this supplies enough cases to give us a feeling for whether MASS 2.0 had a pronounced impact upon the forecaster's ability to predict accurately.

Altogether, the forecaster issued 11 LFM- and 13 MASS 2.0-based convective outlooks during the six days.

Only the verification results for the convective outlook will be reported here, because it is unfair to judge the forecaster's ability to accurately make severe weather watch boxes when he was removed from all conventional weather, satellite, and radar data input after 1200 EDT (a decided advantage for SELS). Burgess worked under three important handicaps in issuing convective outlooks. First, the 700 mb data observed at 1200 GMT on each forecast day frequently was not received, so that he had little knowledge of existing or forming capping inversions. Second, he had no information on capping inversions predicted by either model. Third, he had little practical experience with severe weather forecasting over the eastern United States.

The actual verification of the Burgess LFM- and MASS 2.0-based and NSSFC (SELS) convective outlooks was performed by NSSFC using an existing computer program and the SELS severe weather report logs. All outlooks were in the "slight risk" category. The modified critical success index (CSI) method described by Weiss *et al* (1980), which accounts for both the coverage (actual and forecast) and the areal distribution of verifying reports within the false alarm ratio (FAR) calculation, was employed. The CSI is defined as the ratio of successful severe storm predictions to the total number of observed severe storm events and false alarms:

$$CSI = x/(x + y + z), \quad (5)$$

where

x = severe storms correctly predicted ("hits")

y = severe storms not predicted ("misses")

z = non-severe storms predicted to be severe ("false alarms"),

are the three groups. The "hits" and "false alarms" are measured in terms of percentages with the probability of detection (POD) and false alarm ratio (FAR) scores, respectively:

$$\text{POD} = x/(x + y), \quad (6)$$

and

$$\text{FAR} = z/(x + z). \quad (7)$$

The most successful CSI of 1.0 is attained when $\text{POD} = 1.0$ and $\text{FAR} = 0.0$, since the three variables are inter-related as

$$\text{CSI} = [(\text{POD})^{-1} + (1-\text{FAR})^{-1} - 1]^{-1}. \quad (8)$$

The modified CSI incorporates event density (clustering of reports characteristic of severe storm events) and accounts for the problem of CSI insensitivity to the POD value when FAR exceeds 90% (Weiss et al, 1980).

Finally,

$$\text{BIAS} = (x + z)/(x + y) \quad (9)$$

is calculated to reveal any systematic overprediction ($\text{BIAS} > 1$) or underprediction ($\text{BIAS} < 1$) of severe storm events. A generally recognized dilemma in severe storm forecasting is that an outlook area must be large enough to keep the POD high, but small enough to keep the FAR tolerably low and the BIAS close to unity. Severe storm events are classified as either hail, tornado, or windstorm as defined in Weiss et al. (1980).

2. Results of Limited Real-Time Forecast Experiment

The lone forecaster in this experiment came away with the strong perception that "MASS 2.0 output was highly accurate in forecasting areas of strong mesoscale thunderstorm development", and "a much better convective storm forecasting tool than is the LFM".⁶ These personal feelings are borne out to some degree by the objective scoring statistics displayed in Table 14. The Burgess-LFM and SELS forecasts are compared, first of all, to establish a fundamental performance level suitable for Burgess. Overall, Burgess and SELS forecasted at a comparable level of skill, with Burgess being more successful with "hits" (POD). It is very important for establishing a base line for MASS 2.0 impact that the Burgess-LFM and SELS outlooks were of quite similar size. When Burgess used the MASS 2.0 output, he was able to produce outlooks that were a mere 58% of the size of his LFM-based outlooks, without suffering a major reduction in his ability to detect (POD). He thereby achieved a significant reduction in his false alarms (FAR) and a 15% increase in skill (CSI).

The size of this sample is naturally too small to allow us to make quantitative conclusions that are statistically significant. Moreover, SELS' performance during this experiment was slightly below normal for June, notably in the FAR (average SELS scores for the month of June from 1973-1976 as reported by Weiss (1977) are: POD = 0.41, FAR = 0.50, CSI = 0.25). Thus, no significance should be attached to the apparent result that Burgess outscored SELS by 32% when using MASS 2.0 output. However, the results do indicate that the mesoscale model could be of assistance to a severe storms forecaster who

⁶Personal communication, Mr. Don Burgess (NOAA/ERL/NSSL), August 1982.

ORIGINAL PAGE 13
OF POOR QUALITY

TABLE 14. STATISTICAL SKILL SCORES FOR LFM AND MASS 2.0-BASED EXPERIMENTAL FORECASTER (BURGESS) CONVECTIVE OUTLOOKS AND FOR SELS 0900 GMT CONVECTIVE OUTLOOKS MADE DURING REAL-TIME EXPERIMENT 21 JUNE - 29 JUNE, 1982.

PARAMETER	CONVECTIVE OUTLOOK FORECAST		
	BURGESS- LFM	BURGESS- MASS 2.0	SELS- 0900 GMT
NUMBER OF OUTLOOK AREAS	11	13	?
AVERAGE OUTLOOK AREA (km ²)	633,579	367,953	731,723
POD	0.57	0.52	0.44
FAR	0.67	0.58	0.69
CSI	0.25	0.29	0.22

must presently rely upon surface fields and satellite/radar signatures to detect early signs of mesoscale organization. The model may provide the forecaster with information on the likely depth, strength, and future continuity of the fields, and thus help to determine an appropriate threshold for forecaster action.

This point can be demonstrated by comparing Burgess' LFM- and MASS 2.0-based convective outlooks on a couple of the days to see what kind of impact the two-hourly, mesoscale information had on his judgmental decisions. The separate outlooks and accompanying verifying severe weather reports on all six days of the experiment are shown in Fig. 22. The 23 June case shows a typical reduction in the outlook area, with correspondingly lower FAR, when using MASS 2.0 in place of the LFM. In this case, Burgess placed an LFM-based outlook too far south in Texas because he saw a diffluent exit region of a

ORIGINAL PAGE IS
OF POOR QUALITY

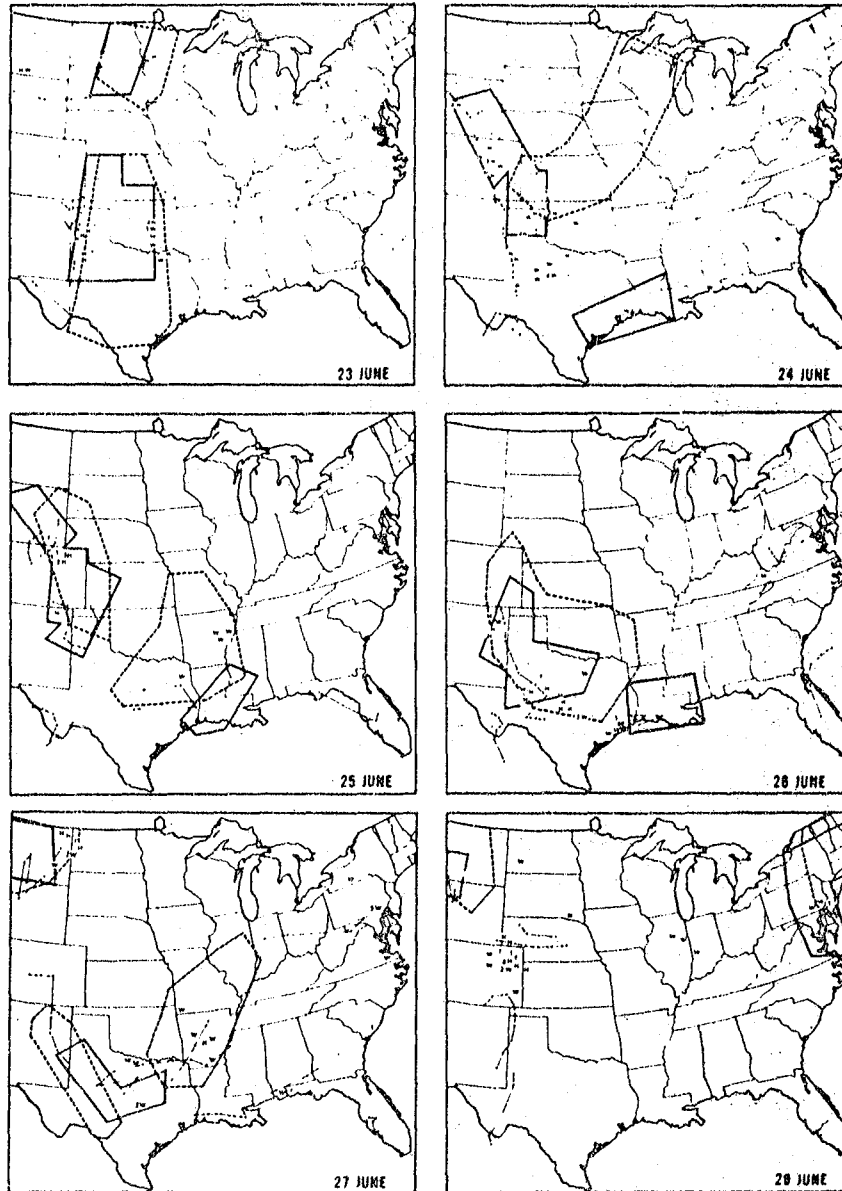
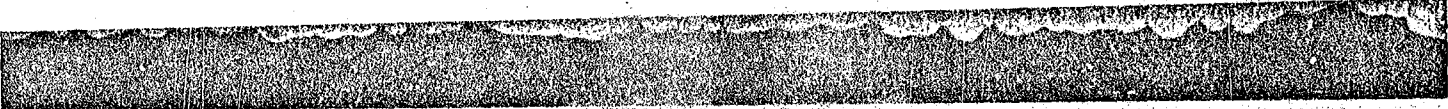


Figure 22. LFM- and MASS 2.0-based convective outlooks produced by experimental forecaster (dashed and solid boxes, respectively) on the six days of real-time forecast experiment. Irregularly-shaped MASS 2.0 based areas are the result of combining areas of separate severe weather watch boxes. Also shown are verifying severe weather reports, where H = large hail, T = tornado, W = damaging windstorm, and numbers designate number of such reports within a cluster. Position of locus of convective predictor variable (WLI, FCE, or WA) is shown at 2h intervals by stars (arrows denote increasing time).



200 mb jet positioned over a very unstable air mass there. With MASS 2.0 information, he was able to trim off a large part of this area in Texas as the model developed maximum instability and dryline convergence further northward over the Texas Panhandle. Likewise on 24 June, MASS 2.0 information enabled him to successfully eliminate the large region from central Kansas to Wisconsin which did not verify, as the mesoscale model failed to develop sufficient instability there that was anticipated with the LFM output and early morning observations.

MASS 2.0-based outlooks were generally overforecast along the Gulf coast where the dynamics were very weak and instability persistently high. We also note that there was occasionally a problem with lack of sufficient information to the forecaster, which necessitated his reliance upon perceived similarities with the past day's synoptic situation. On 29 June, Burgess did not outlook Colorado because he suspected the existence of a strong mid-level capping inversion which had existed in that area the day before. The 700 mb chart was not received by him that day, so using feedback from conditions the day before, he overlooked strong MASS 2.0 indications of convective development. Notice that the convective predictor variable (MCS locus) method used by the lead author gave good indications of potential there (Fig. 22). Notice that with few such exceptions, there was good agreement between the objective MCS locus method and the experimental forecaster's best judgments as to where strong mesoscale convection was being forecast by the model. The good agreement between the two independent measures of model mesoscale performance gives credence to the validity of the two separate approaches for measuring that performance.

10. SUMMARY OF PART II: MESOSCALE MODEL EVALUATION RESULTS

A comprehensive study of the mesoscale predictability of the model was performed by subjectively verifying the model forecasts with observed convection. The results of this study were supplemented with findings from a limited real-time, "operational-like" convective forecast experiment with the model. Also, two case examples were presented as illustration of the fact that unfiltered mesoscale fields forecast by MASS 2.0 possess a high degree of coherent information.

a. Case Examples

In the one case example, a large-scale, well-organized, severe weather outbreak case from early springtime was forecast very well by the model. The corresponding LFM forecast showed only a large-scale dipole in the vertical motions with a large eastward phase error. In contrast, MASS 2.0 predicted several temporally coherent upward motion maxima that were accurately related to several observed mesoconvective systems (MCS's).

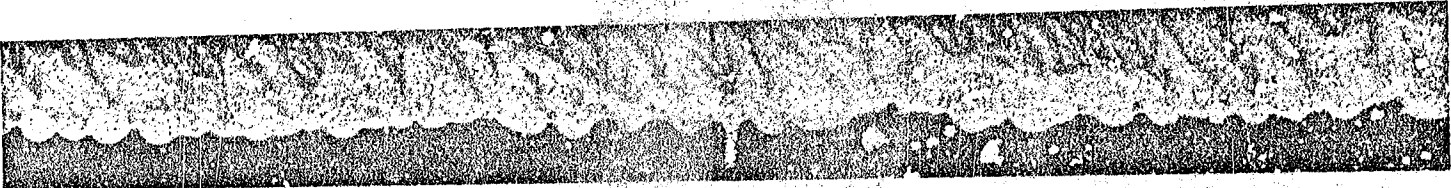
The differences between the two models are even more pronounced when isolated outbreak cases are considered. In the second case example, MASS 2.0 produced what appeared to be a meso-alpha scale mountain wave in west Texas. The vertical circulation associated with this wave, coupled with the vertical circulation about a dryline bulge was successfully related to the transient outbreak of severe storms there. The sophisticated planetary boundary layer parameterization and high resolution grid in MASS 2.0 were apparently crucial to model development of the mountain wave, as well as to the intensification of the low level jet, the formation of the dryline bulge, and the production of the lifting mechanisms that apparently resulted in severe storm formation.

ORIGINAL PAGE IS
OF POOR QUALITY

b. *Comprehensive Model Verification against Observed Convection*

The comprehensive study of MASS 2.0 mesoscale predictability was based on a method whereby the loci of convective predictor variable fields, derived from the model forecasts, were related diagnostically to the loci of 149 individual MCS's observed by radar. This approach was used to assess mesoscale predictability because it is not possible to objectively verify the forecasts with upper air observations collected routinely at the synoptic scale. Three convective predictor variables were defined, all of which basically included measures of lifting and air mass convective instability. However, none of the variables explicitly and unambiguously indicated when and where the model was predicting deep convection, being instead presumptive in nature. In spite of this drawback, use of these variables allowed for successful diagnosis of 50% of the observed MCS's within 250 km and \pm 3h of their observed formation. False alarms were very low (14%). Although it is true that areal distribution of forecast and observed events was not considered, these results strongly indicate that *MASS 2.0 mesoscale forecasts yielded coherent fields useful for forecasting convective weather and for diagnosis of mesoscale processes relevant to the initiation of strong convection.*

"Underforecasts" of convection were much more frequent than "overforecasts". The leading causes for both problems were analyzed. The most frequent causes for underpredictions were poorly forecast upper tropospheric "short wave" disturbances, underforecast potential instability, poor forecasts of frontal convergence (too weak, mispositioned), and a host of phenomena related to the lack of a cumulus parameterization scheme in the model. The instability and convergence problems are mainly related to systematic synoptic-scale errors discussed in Part I, namely poor moisture initialization



and situation-dependent problems blamed on the eastern boundary conditions, respectively. However, the model displayed its highest skill score during the period of worst synoptic-scale performance by the model, showing that useful mesoscale forecasts in the western half of the domain could be attained despite synoptic-scale problems to the east. This identification of the leading causes for missed forecasts of convection suggests limits to the current mesoscale predictability of the model, which should prove helpful in future MASS development.

c. *Convective Forecast Experiment*

The limited forecast experiment was conducted during nine days in late June, under a flow regime dominated by High Plains severe activity. A single forecaster made two 18h convective outlooks, one based upon the LFM output, the other based upon MASS 2.0 output, both with the help of analyzed observations that were obtained only prior to making the forecast. The forecaster worked under the handicaps of never having any idea of whether MASS 2.0 was predicting strong capping inversions. Also, he frequently had no idea of whether capping inversions actually existed at forecasting time, and had limited prior experience in forecasting severe weather outside of the Southern Plains.

Despite these handicaps, the forecaster was able to reduce the size of the LFM-based outlooks by 42% by using MASS 2.0 information without suffering a major reduction in his ability to detect ("hit"). The mesoscale model provided useful information to him about evolving fields of potential instability and the depth and coherence of lifting mechanisms, not obtainable from the LFM (nor from conventional surface observations). These findings suggest that the model

could be used in an intelligent way to refine a forecaster's threshold for action.

It was noted that there was, in general, good agreement between the objective MCS locus method and the experimental forecaster's best judgments as to where strong mesoscale convection was being forecast by the model. This agreement gives credence to the validity of the two independent approaches for measuring the mesoscale performance of the model.

It is our general conclusion that MASS 2.0 can provide very informative, coherent, and meaningful fields that can be used for a variety of purposes. This model has great potential as both a research tool for diagnostic case study purposes and as a forecasting aid, but only in those situations where the effects of model systematic errors are minimal. The limits to the model's mesoscale predictability were strongly controlled by the characteristics of the large-scale circulation (e.g., the "BC regime"), by the particular type of mesoscale regime (e.g., convective vs. non-convective), and by inadequacies in the initialization (e.g., in the moisture field). We therefore recommend that experiments aimed at assimilating satellite data like VAS into the model take these model characteristics into account, and that every attempt be made in the meantime to remedy the systematic model deficiencies before assimilation in general is attempted.

ACKNOWLEDGEMENTS

Dr. Louis Uccellini (NASA/GLAS) was the initiator of this project, and provided both the needed support for its successful completion and many constructive ideas. We also thank Dr. James Dodge at NASA Headquarters for his continued support necessary to develop and test the MASS model. Norman Crabill (NASA/LaRC) provided the necessary computer support. We are grateful to the SASC modelling group, headed by Dr. Michael Kaplan, for their compliance with our requirements. Dr. Joseph Schaefer (NOAA/NSSFC) was kind enough to process the forecasts issued during the real-time experiment. We gratefully acknowledge the contributions made by Mr. Donald Burgess (NOAA/ERL/NSSL) both in his role as the experimental forecaster and for the well-written summary of his fruitful experiences with the model.

REFERENCES

- Anthes, R. A., Y-H Kuo, S. G. Benjamin, and Y-F Li, 1982: The evolution of the mesoscale environment of severe local storms: preliminary modeling results. *Mon. Wea. Rev.*, **110**, 1187-1213.
- _____, and D. Keyser, 1979: Tests of a fine-mesh model over Europe and the United States. *Mon. Wea. Rev.*, **107**, 963-984.
- _____, and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev.*, **106**, 1045-1078.
- Atlas, R., M. Ghil, and M. Halem, 1982: The effect of model resolution and satellite sounding data on GLAS model forecasts. *Mon. Wea. Rev.*, **110**, 662-662.
- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396-409.
- Baumhefner, D. P., and D. J. Perkey, 1982: Evaluation of lateral boundary errors in a limited-domain model. *Tellus*, **34**, 409-428.
- Carlson, T. N., R. A. Anthes, M. Schwartz, S. G. Benjamin, and D. G. Baldwin, 1980: Analysis and prediction of severe storms environment. *Bull. Amer. Meteor. Soc.*, **61**, 1018-1032.
- Chang, C. B., D. J. Perkey, and C. W. Kreitzberg, 1982: A numerical case study of the effects of latent heating on a developing wave cyclone. *J. Atmos. Sci.*, **39**, 1555-1570.
- Charba, J. P., 1979: Two to six hour severe local storm probabilities: An operational forecasting system. *Mon. Wea. Rev.*, **107**, 268-282.
- Charney, J. G., and A. Eliassen, 1964: On the growth of the hurricane depression. *J. Atmos. Sci.*, **21**, 68-75.
- Chesters, D., L. W. Uccellini, A. Mostek, and W. Robinson: Low-level water vapor fields from VISSR Atmospheric Sounder (VAS) "Split-Window" channels at 11 and 12 microns. *NASA Tech. Memo. 83951*, 33 pp. (Accepted for publication in *J. of Climate and Appl. Met.*).
- Donaldson, R. J., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. *Proc. 9th Conf. on Severe Local Storms*, Norman, AMS, 321-326.
- Deardorff, J. W., 1972: Parameterization of the planetary boundary layer for use in general circulation models. *Mon. Wea. Rev.*, **100**, 93-106.

- Driedonks, A. G. M., and H. Tennekes, 1981: Parameterization of the atmospheric boundary layer in large-scale models. *Bull. Amer. Meteor. Soc.*, 62, 594-597.
- Ferguson, E. W., J. T. Schaefer, S. J. Weiss, and L. F. Wilson, 1983: The year of the tornado. *Weatherwise*, 36, No. 1, 18-27.
- Fritsch, J. M., and C. F. Chappell, 1980: Numerical prediction of convectively driven mesoscale pressure systems. Part I: convective parameterization. *Mon. Wea. Rev.*, 37, 1722-1733.
- Gerrity, J. F., 1977: The LFM model-1976: A documentation. *NOAA Tech. Memo NWS-NMC-60*, 68 pp.
- Holton, J. R., 1972: *An Introduction to Dynamic Meteorology*. Academic Press, New York, 319 pp.
- Kaplan, M. L., J. W. Zack, V. C. Wong, and J. J. Tuccillo, 1982: Initial results from a mesoscale atmospheric simulation system and comparisons with the AVE-SESAME I data set. *Mon. Wea. Rev.*, 110, 1564-1590.
- Kesel, P. G., and F. J. Winninghoff, 1972: The Fleet Numerical Weather Central operational primitive-equation model. *Mon. Wea. Rev.*, 100, 360-373.
- Klemp, J. B., and D. K. Lilly, 1975: The dynamics of wave-induced downslope winds. *J. Atmos. Sci.*, 32, 320-339.
- Koch, S. E., and J. McCarthy, 1982: The evolution of an Oklahoma dryline Part II: Boundary-layer forcing of mesoconvective systems. *J. Atmos. Sci.*, 39, 237-257.
- Kreitzberg, C. W., and D. J. Perkey, 1976: Release of potential instability: Part I, a sequential plume model within a hydrostatic primitive equation model. *J. Atmos. Sci.*, 33, 456-457.
- Leary, C., 1971: Systematic errors in operational National Meteorological Center primitive-equation surface prognosis. *Mon. Wea. Rev.*, 99, 409-413.
- Lindzen, R. S., 1974: Wave-CISK in the tropics. *J. Atmos. Sci.*, 31, 156-179.
- Maddox, R. A., D. J. Perkey, and J. M. Fritsch, 1981: Evolution of upper tropospheric features during the development of a mesoscale convective complex. *J. Atmos. Sci.*, 38, 1664-1674.
- Maddox, R. A., and C. A. Doswell III, 1982: An examination of jet stream configurations, 500 mb vorticity advection, and low-level thermal advection patterns during extended periods of intense convection. *Mon. Wea. Rev.*, 110, 184-197.
- Mahrer, Y., and R. A. Pielke, 1977: A numerical study of the airflow over irregular terrain. *Beiträge zur Physik der Atmosphäre*, 50, 98-113.

- Matsuno, T., 1966: A finite difference scheme for time integrations of oscillatory equations with second-order accuracy and sharp cutoff for high frequencies. *J. Meteor. Soc. Japan, Ser 2, 44*, 145-157.
- Matthews, D. A., and B. A. Silverman, 1980: Sensitivity of convective cloud growth to mesoscale lifting: a numerical analysis of mesoscale convective triggering. *Mon. Wea. Rev.*, *108*, 1056-1064.
- McCarthy, J., and S. E. Koch, 1982: The evolution of an Oklahoma dryline. Part I: A meso- and subsynoptic-scale analysis. *J. Atmos. Sci.*, *32*, 225-236.
- Newell, J. E., and D. G. Deaven, 1981: The LFM-II model--1980. *NOAA Tech. Memo, NWS-NMC-66*, 20 pp.
- National Research Council, 1980: Atmospheric precipitation: prediction and research problems. Committee on Atmospheric Sciences, National Academy of Sciences, Washington, DC, 63 pp.
- Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, *56*, 527-530.
- Panofsky, H. A., and G. W. Brier, 1968: *Some Applications of Statistics to Meteorology*. Penn State University Press, University Park, PA, 224 pp.
- Pearson, A., and S. J. Weiss, 1979: Some trends in forecast skill at the National Severe Storms Forecast Center. *Bull. Amer. Meteor. Soc.*, *60*, 319-326.
- Perkey, D. J., 1976: A description and preliminary results from a fine-mesh model for forecasting quantitative precipitation. *Mon. Wea. Rev.*, *104*, 1513-1525.
- Reap, R. M., and D. S. Foster, 1979: Automated 12-36 hour probability forecasts of thunderstorms and severe local storms. *J. Appl. Meteor.*, *18*, 1304-1315.
- Sasaki, Y., 1958: An objective analysis based on the variational method. *J. Meteor. Soc. Japan*, *36*, 77-88.
- Silberberg, S. R., and L. F. Bosart, 1982: An analysis of systematic cyclone errors in the NMC LFM-II model during the 1978-79 cool season. *Mon. Wea. Rev.*, *100*, 254-271.
- Tarbell, T. C., T. T. Warner, and R. A. Anthes, 1981: An example of the initialization of the divergent wind component in a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, *109*, 77-95.
- Teweles, S., Jr., and H. B. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.*, *10*, 455-463.

- Warner, T. T., R. A. Anthes, and A. L. McNab, 1978: Numerical simulations with a three-dimensional mesoscale model. *Mon. Wea. Rev.*, 106, 1079-1099.
- Weiss, S. J., D. L. Kelly, and J. T. Schaefer, 1980: New objective verification techniques at the National Severe Storms Forecast Center. *Preprints 8th Conf. on Weather Forecasting and Analysis, Denver, Amer. Meteor. Soc.*, 412-419.
- Weiss, S. J., 1977: Objective verification of the severe weather outlook at the National Severe Storms Forecast Center. *Preprints, 10th Conf. on Severe Local Storms, Omaha, Amer. Meteor. Soc.*, 395-402.
- Willmott, C. J., 1982: Some comments on the evaluation of model performance. *Bull. Amer. Meteor. Soc.*, 63, 1309-1313.
- Wilson, G. S., and R. E. Turner, 1982: Objective predictions of thunderstorm location and severity for aviation. *Journal of Aircraft*, 19, 330-338.

END

DATE

FILMED

JUL 27 1983

LANGLEY RESEARCH CENTER



3 1176 00507 9943