Skill of NMC Operational Dynamical Models in Prediction of Explosive Cyclogenesis

FREDERICK SANDERS

Marblehead, MA 01945

(Manuscript received 2 June 1987, in final form 9 September 1987)

ABSTRACT

The skill of the Nested-Grid Model (NGM) and the global spectral model (GLBL) at the National Meteorological Center in the prediction of explosive cyclogenesis was evaluated for the period 1 September 1986–30 April 1987. Manual analyses covering the eastern North Pacific, North America and the North Atlantic eastward to 20°W were used as ground truth. The criterion for a bomb event in the analyses of the forecasts was a deepening of the center of at least 24 mb at 60°N, normalized geostrophically at other latitudes, in a period of 24 h, beginning at 0000 or 1200 UTC.

Both models displayed skill out to 48 h for the NGM and 60 h for the GLBL. The NGM performed notably better in the innermost fine-grid area than in the surrounding area of overlap with a more coarse grid. For the GLBL in the Atlantic and North America, similar skill was seen through 36 h; skill was very small in the Pacific region. Predicted 12-h deepening beginning 12 h after initialization was compared with analyzed deepening for both models. Correlations ranged from 0.72 for the NGM in the inner grid over the Atlantic and North America to 0.03 in the Pacific. The GLBL values were intermediate, again better in the Atlantic than in the Pacific. All samples showed an average shortfall of predicted deepening from 12–24 h after initialization, ranging from 1 mb for the inner NGM grid to 7 mb for the overlap area, with the GLBL intermediate; again, it was much better in the Atlantic than in the Pacific.

Growth of skill over the past few years is attributable to improved analyses, better model resolution and better treatment of boundary-layer fluxes. Initial data limitations are now the most important factor, both in models and in verifying analyses. These results alter the nature of the problem of research on explosive cyclogenesis from one of discovering a missing ingredient to one of improving the performance and extending the range of predictability.

1. Introduction

Routine inspection in fall 1986 of the prognostic charts received on digital facsimile (DIFAX) from the Nested-Grid Model (NGM) of the National Meteorological Center (NMC) indicated that explosive cyclogenesis was predicted to occur on a number of occasions, apparently much more readily than in prognoses derived from the Limited-area Fine-Mesh Model (LFM) (Sanders, 1986b). This might not have been surprising, since more or less successful research prediction of "bombs" has been reported by Anthes et al. (1983), Chen et al. (1983), Liou and Elsberry (1987), Uccellini et al. (1987), Orlansky and Katzfey (1987), Douglas and Warner (1987) and Nuss and Anthes (1987). It is surprising, however, to encounter evidence that operational prediction, with all its disadvantages relative to a research environment, might be so close to the research state of the art.

Operational numerical underprediction of oceanic cyclogenesis has been noted for many years (Leary, 1971; Sanders and Gyakum, 1980; Silberberg and Bosart, 1982). Since this current change in operational model performance involves a phenomenon which has recently attracted much attention in the research community (western North Atlantic bombs are the focus of the proposed ERICA Project as described by Abbey et al., 1987), a systematic examination of operational forecasting of explosive cyclogenesis during the 1986/87 cold season was undertaken. This included study of the performance of the Global Spectral Model (GLBL) (in whatever version was received on DIFAX) as well, since bombs seemed to be produced by this model also.

2. The data

The area examined was the Northern Hemisphere from longitude 170°W eastward to 10°W for the GLBL forecasts and the DIFAX display area (see Fig. 1) for the NGM. In the latter case, it was noted whether the storm was in the area of the C-grid or in the surrounding ring in which meshed calculations on the C- and B-grids are performed, since cyclogenesis seemed to be suppressed here (Hoke, 1987). On 25 February 1987, the C-grid was enlarged by NMC to cover nearly the entire display area in Fig. 1, a change which was taken into account in the results to be presented.

The definition of a bomb was that used by Sanders and Gyakum (1980) and refers to 24-h deepening. The rate necessary to give a critical value of 1 Bergeron is given as a function of latitude in Table 1.

Seven 24-h forecast periods were examined, named

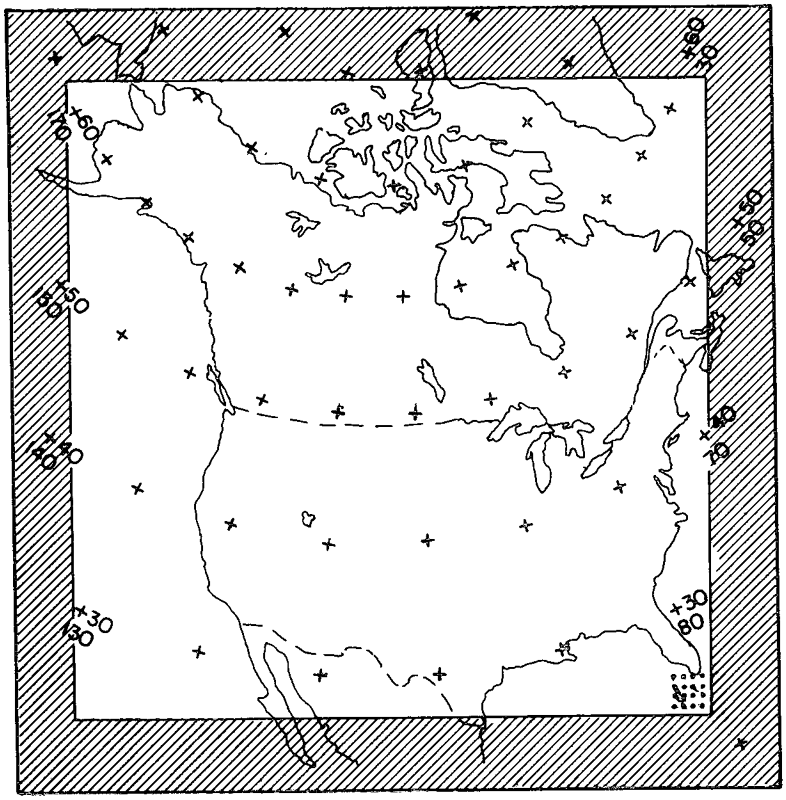


FIG. 1. Verification area for the NGM forecasts, with sample of the C-grid in lower right-hand corner. The hatched area denotes overlap of the C-grid with the coarser B-grid.

for the interval from the initial time to the middle of the 24-h period. Since the NGM prognostic charts were available at 12-h intervals to 48 h and the GLBL output extended to 60 h (although over only the NGM area at this last time), a total of three NGM and four GLBL forecast periods were used. For example, the 24-h NGM period represents the time series of forecasts from 12 to 36 h after the initial time, initialized each 12 h. Verification data were obtained from the NMC preliminary analyses received on DIFAX. (The behavior of the real atmosphere, of course, is unknown.)

Analyses and forecasts were examined at 12-h intervals. For forecasts initialized at 0000 and 1200 UTC, the occurrence of a bomb in successive overlapping 24-h periods was noted, characterized by the latitude, longitude, and time of the midpoint and by the 24-h deepening in terms of the Bergeron value. In the NMC analyses, deepening was determined in overlapping 24-h periods beginning at 0000 and 1200 UTC. In numerous instances no pressure minimum was present at the onset of intense deepening. In almost all of these, however, the position of a maximum Laplacian of pressure (or maximum geostrophic vorticity) could be found in an appropriate location and was used instead. On rare occasion, the pressure at a trough line had to be used, at a point estimated by backwards extrapolation of the subsequent track of the low pressure center.

Additionally, for 12–24 h after initialization for each forecast cycle, the deepening of the forecast cyclone corresponding to an analyzed bomb was determined, whether or not the storm qualified as a bomb in the forecast atmosphere. These last data were obtained for comparison with results presented in previous studies of earlier operational models.

3. Results

a. Hits, misses and false alarms

The chronology and longitude of each bomb appears in Figs. 2–4 for the analyses and for model forecasts at the shortest and longest ranges examined. It is im-

TABLE 1. 24-h deepening corresponding to 1 Bergeron.

Latitude	Deepening (mb)	Latitude	Deepening (mb)	Latitude	Deepening (mb)
90°–83°	28	54°–51°	22	36°–34°	16
82°–74°	27	50°–48°	21	33°–32°	15
73°–67°	26	47°–45°	20	31°–30°	14
66°–63°	25	44°–43°	19	29°–27°	13
62°–58°	24	42°–40°	18	26°–25°	12
57°–55°	23	39°–37°	17	24°–23°	11

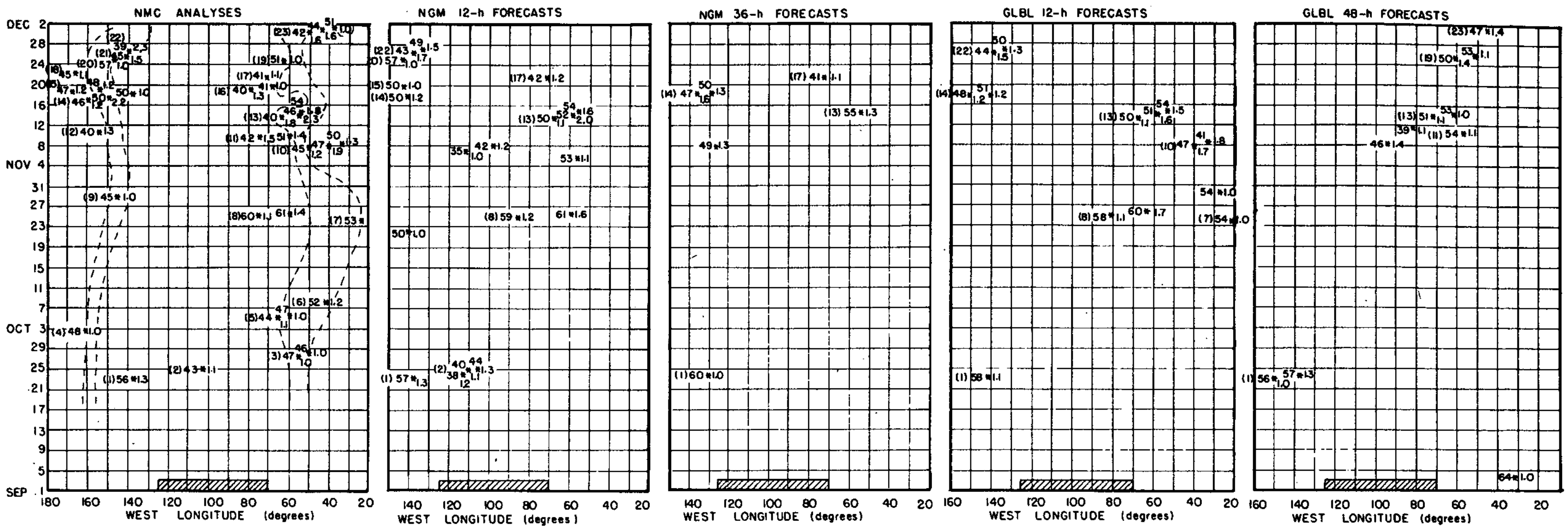


FIG. 2. Time and longitude of explosive cyclones for 1 September–2 December 1986 for (a) NMC analyses, (b) NGM 12-h forecasts, (c) NGM 36-h forecasts, (d) GLBL 12-h forecasts, and (e) GLBL 48-h forecasts. An asterisk is entered for any 24-h period in which the cyclone's deepening rate was at least 1 Bergeron. Number to the right of the asterisk is the Bergeron value. Number immediately to the left is the latitude. Serial number of the storm is in parentheses. The inner pair of dashed lines denotes the edge of the NGM C-grid. The outer pair denotes the edge of the display area. The hatched area represents the approximate longitude range of North America. Missing data are denoted by "m." See text for definition of forecast periods.

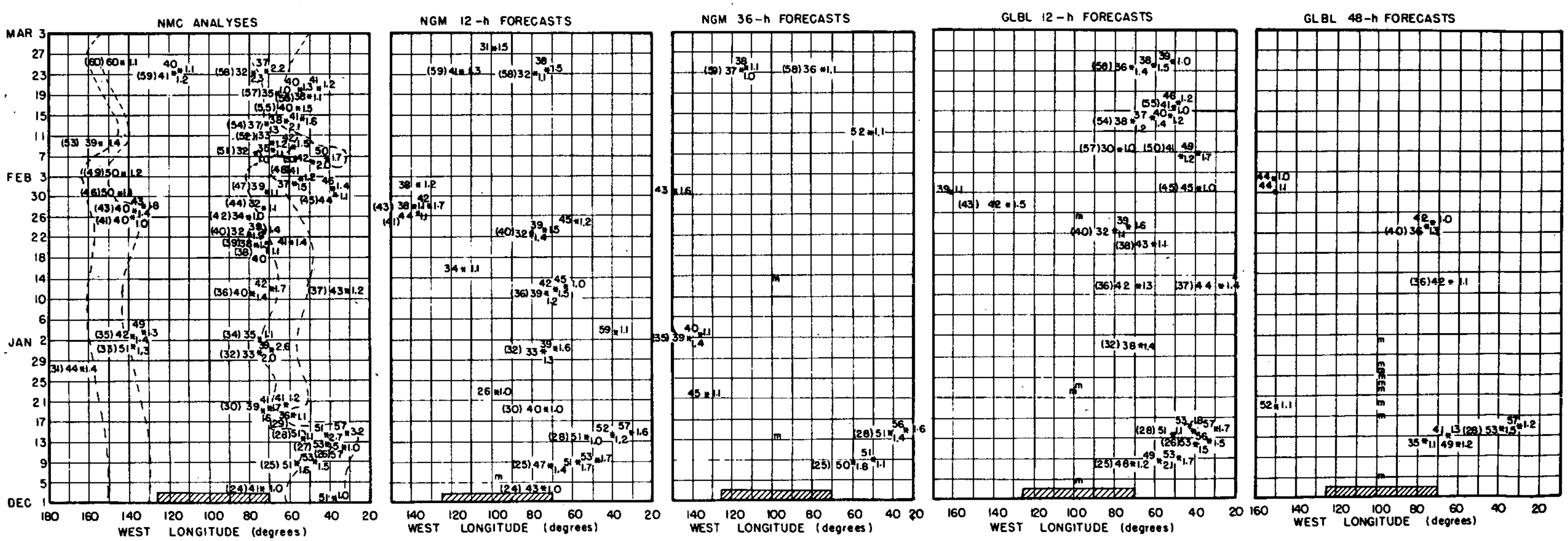


FIG. 3. As in Fig. 2 except for 1 December 1986–3 March 1987.

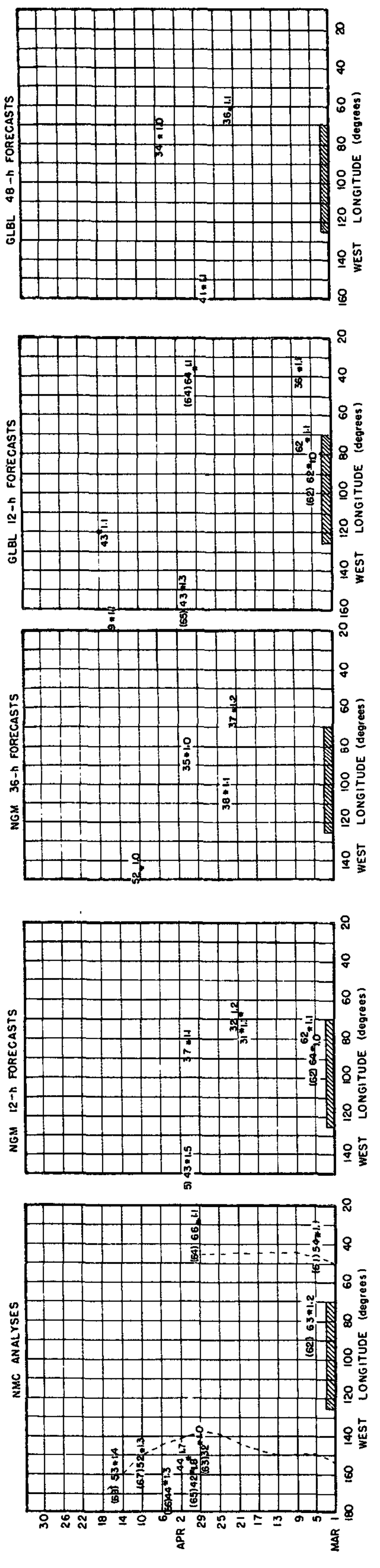


FIG. 4. As in Fig. 2 except for 1 March–22 April 1987.

mediately apparent that bombs occur in the appropriate regions with some frequency in all datasets, although somewhat less often as the forecast range increases; this decrease was monotonic with range. A substantial number of storms are not detected by the forecast model, and the models produce occasional false alarms.

The number of analyzed bombs and the number that were detected by each model at each forecast range, as well as the number of false alarms produced, is given in Table 2. Whether comparison between analysis and prediction for a given storm represented a success or not was judged subjectively. Often the timing was correct to the nearest 12 h, and the position error was less than a few degrees of longitude and latitude. At other times, when the explosive deepening occurred in an adjacent 12-h period or the position error was greater, generosity prevailed. Close examination of Figs. 2–4 will enable the reader to form an independent judgment.

The NGM success in detection was much greater when the storm occurred within the C-grid than when it was in the region where the calculations on the C- and B-grids interact, as anticipated. The false alarm rate was larger in the C-grid as well. The success in detection decreased with range in both models, except in the 24-h GLBL predictions (12-h forecast GLBL atmosphere). The number of successful forecasts in this particular model atmosphere may be understated, because of the way the forecast deepening in the first 12 h was determined. The GLBL initial analysis was not available on DIFAX, and substitution from the NGM analysis or from the preceding 12-h GLBL forecast when the NGM analysis was not available were the only feasible alternatives. The GLBL analysis probably showed a less-deep central pressure than the more highly resolved NGM, so that the estimated 24-h deepening was probably less than what the model actually produced.

Note in Table 2 that the GLBL model performed much less well in the Pacific than in the Atlantic region. The initial analysis was less well known in the former area. The small number of successes occurred with storms relatively close to the Pacific coast or in the Gulf of Alaska, where the soundings from stations in the Aleutian chain may have provided a relatively good initial analysis. (To call storm No. 14 a success is to be extremely generous. The model atmospheres showed the cyclogenesis, if at all, too weakly, too late and too far east.) The NGM failed to show much distinction in success between the Atlantic and Pacific, perhaps because it excluded a large area in the central Pacific where the initial analysis was particularly uncertain.

False alarms were notably fewer than detection successes in the Atlantic, and about equal in the Pacific. Even here, the probability of a bomb in the analyses, given the prediction of one, was much greater than the probability, given no such prediction—since up to 484

TABLE 2. Events (E), hits (H), and false alarms (FA); probability of detection (POD), false-alarm rate (FAR), and Critical Success Index (CSI).* Modified CSI values in parentheses (see text).

12-h			24-h			36-h			48-h		
E POD	H FAR	FA CSI	E POD	H FAR	FA CSI	E POD	H FAR	FA CSI	E POD	H FAR	FA CSI
					NGM, all						
61	27	12	61	24	7	61	13	6			
.44	.31	.37	.39	.23	.35	.21	.32	.19			
		(.34)			(.40)			(.22)			
					NGM, C-grid						
35	18	11	35	17	7	35	9	5			
.51	.38	.39	.49	.29	.40	.26	.36	.22			
		(.32)			(.36)			(.24)			
					NGM, C–B overlap						
26	9	1	26	7	0	26	4	1			
.35	.10	.33	.27	.00	.33	.15	.20	.15			
		(.39)			(.47)			(.25)			
					GLBL, all						
68	27	3	68	27	12	67	18	9	52	9	10
.40	.10	.38	.40	.31	.32	.27	.33	.24	.17	.53	.14
		(.47)			(.42)			(.35)			(.19)
					GLBL, Atlantic and North America						
44	20	3	44	22	7	43	15	7	36	7	7
.45	.13	.43	.50	.24	.43	.35	.32	.30	.19	.50	.16
		(.48)			(.53)			(.45)			(.24)
					GLBL, Eastern Pacific						
24	7	0	24	5	5	24	3	2	16	2	3
.29	.00	.29	.21	.50	.22	.12	.40	.12	.12	.60	.10
		(.43)			(.21)			(.13)			(.09)

* An event is the occurrence of deepening at a rate of at least 1 Bergeron at some time during the passage of the analyzed storm across the verification area. A hit is the successful forecast of an event. A false alarm is an unsuccessful forecast of an event. Probability of detection (POD) is H/E. False-alarm rate (FAR) is FA/(H + FA). Critical Success Index (CSI) is given by CSI = $[(POD)^{-1} + (1 - FAR)^{-1} - 1]^{-1}$.

forecasts were examined, containing a larger but uncounted number of candidate cyclones. Over North America, forecast bombs and analyzed ones were rare. In the NGM there was a slight tendency of the model to overpredict continental bomb occurrence, thus suggesting a weak continuation of a systematic error first noted by Leary (1971).

b. Critical Success Index and skill score

Critical Success Indices (Donaldson et al., 1975), equivalent to "threat scores," were calculated for the results categorized in Table 2. In the NGM forecasts, the superiority of forecasts in the C-grid over those in the overlap area is seen together with the deterioration with range. The increased probability of detection (POD) in the former is sufficient to offset the adverse increase in the false-alarm rate (FAR). Qualitatively similar contrasts are seen in two categories of GLBL forecasts, but the differences are larger because FAR is larger in the Pacific than in the Atlantic. Overall, there seems to be little to choose between the two models, but the GLBL in the Atlantic region appears to be superior to the NGM in the C-grid, because the latter is hampered by a relatively large FAR.

Inspection of Figs. 2–4 shows that many of the bombs in the analyses missed by the forecast models were of marginal strength, with deepening no more than 2 mb above the critical threshold value. To see what influence this effect might have on CSI, values were calculated considering an event in the analyses to require a maximum Bergeron value of 1.3 or more (moderate or strong bombs as stratified by Sanders, 1986a) but to allow a hit to be recorded if the corresponding forecast maximum value was 1.0 or larger. More false alarms would then occur, of course. In Table 2, these modified CSI values were an improvement in the NGM-mesh and GLBL-Atlantic samples where the stronger analyzed storms tended to be relatively well predicted and the FAR was relatively low. Some deterioration was noted, however, in the NGM C-grid and in GLBL Eastern Pacific.

The CSI takes no account of successful negative forecasts: cases in which a cyclone fails to deepen explosively in either analysis or forecast. A calculation of skill based on the Brier Score for probability fore-

TABLE 3. Maximum Bergeron values ($b_{max} > 1.2$) (mb) for analyses and forecasts.

		Forecast						
		NGM			GLBL			
b_{max}	Analysis	12-h	24-h	36-h	12-h	24-h	36-h	48-h
>2.3	28, 32							
2.3	13, 22							
	58							
2.2	14							
2.1	54				25			
2.0	50	13						
1.9	10, 40							
1.8	43, 65		25	25, 28	10, 8		11, 25	
1.7	30, 36	22, 25	28		8, 10	10, 28	28	
	45	43						
1.6	23, 25	8, 28	13, 32	14, 46	13, 40	8	13	
		32						
1.5	11, 21	36, 40	8, 14		22, 26	26, 32	8	28
	27, 39	*, 58			43, 58			
	48, 51	65						
	55							
1.4	7, 8			35	32, 37	*, 40	46, 54	*, 19
	31, 35				54	46, 50		23
	53, 68							
1.3	1, 12	1, 2	20, 40	8, 13	36, 65	11, *	10, 40	1, *
	16, 33	59				36, 58	55	40
	57, 66					*, *		
	67					65		

* False alarm.

casts, as used by Sanders (1973), seemed desirable since it can be related to reduction of variance and does take successful negative forecasts into account. For reasonable estimates of the number of negative successes (see Appendix) there was a close correlation between the Brier Score and CSI, the former being about 0.05 lower.

c. Maximum Bergeron values and extreme deepening

The maximum Bergeron value for a given storm was somewhat smaller in the forecasts than in the analyses on the average (but not in all instances). The identifying number of each storm in the analyses is listed by maximum Bergeron value in Table 3 for all cases in which it was larger than 1.2. The maximum Bergeron values for the corresponding storms in the seven forecast periods, provided they were at least 1.3, are also shown. The relatively small number of cases in the NGM forecast atmosphere is attributable to poor performance in the C–B overlap area and to the occurrence of a number of the analyzed cases altogether outside the NGM display area.

Track and central pressure for a spectacular example, storm No. 28, are illustrated in Fig. 5. The minimum value of 900 mb in the NMC analysis at 0000 UTC 15 December is based on reports from a buoy and from an adjacent ship. It may be an exaggeration, but not a gross one. An extrapolation from the observation points to the center, based on gradient-wind balance, was reported by Bosart (personal communication) to yield a value of 908 mb. According to Burt (1987), minimum pressures in operational analyses by the British and West German services were 916 and 912 mb, respectively. Any of these estimates make this storm evidently the deepest extratropical cyclone in the historical record.

The NGM forecasts (Fig. 5b) initialized at 0000 UTC 13 December, and thereafter, show strong bomb development despite a location in the overlap area throughout the period of maximum intensification. (The track initialized at 1200 UTC on the thirteenth could not be determined because only the 36-h forecast was received on DIFAX.) The final central pressures were roughly 30 mb higher than the consensus of analyses. The track forecasts were quite good, forecast positions not differing from the analyzed ones by much more than the difference between the two analyses except at 0000 UTC on the 15th, when the northwestward turn of the analyzed track was not caught by the 48-h forecast and was overdone by the 12-h prediction. The GLBL forecasts (Fig. 5c) were quite similar, showing slightly deeper central pressures, on average, but missing the final turn to the northwest except at the shortest range. Despite the failure of both models to reach the extraordinary central pressures estimated in the analyses, they were the lowest seen in the model forecasts during the season.

In general, a particularly explosive deepening might

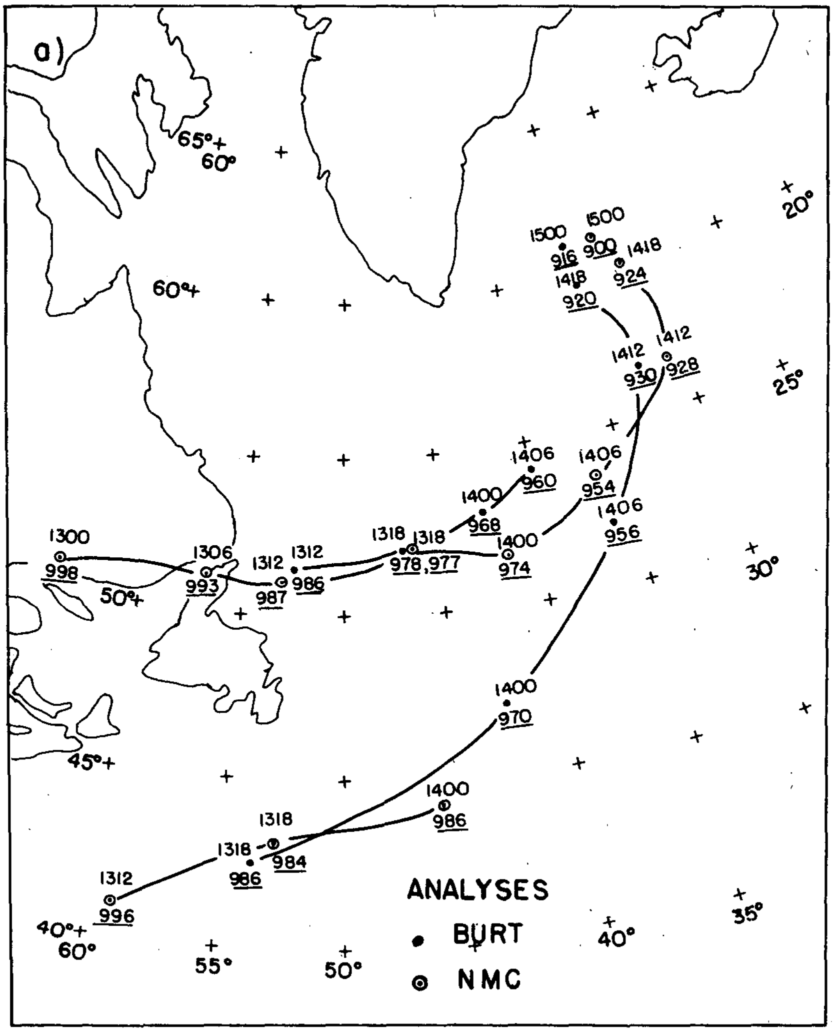


FIG. 5. Tracks of the cyclone of 13–15 December 1986, with positions at 12-h intervals. The date and UTC are indicated by the four digits above the station dot. The central pressure, in mb, is shown below the dot. (a) Analyses, by NMC and by Burt (1987); (b) forecasts by the NGM model; (c) forecasts by the GLBL model. In panels b and c, the date and UTC of the initialization are shown in parentheses above the dot.

be regarded as 20 mb or more in 12 h (c.f. Abbey et al., 1987). A listing of all such instances during the bomb cases shown in Figs. 2–4 is given in Table 4. Except for the GLBL at 60 h, both models predicted this extreme deepening on a number of occasions, although less frequently as the forecast range increased, and substantially less frequently than in the analyses. The decrease with lengthening forecast range suggests that precision in the details of the initial analysis, rather than internal instability in the models, may be responsible for extreme deepening, but a sensitivity study would be required to make the explanation more than speculative.

d. Twelve-hour deepening rate

The format in which these results have been presented is a novel one, prompted by the abundant bomb production by this season's version of the NMC operational models. To provide a direct comparison with earlier results, analyzed values of 12-h deepening for the analyzed bomb cases are plotted in Fig. 6 against corresponding predicted values in the period from 12–24 h, for each model. Comparison with results given in Fig. 15 of Sanders and Gyakum (1980) for the (6-L) and (7-L) Primitive-Equation (PE) models and in Fig. 12 of Sanders (1986b) for the Limited-area Fine-Mesh Model (LFM) shows the distinct improvement in the NGM and GLBL performances during the 1986/87 cold season.

Qualifications, however, are in order. In the NGM forecasts the improvement is seen only in the C-grid (Fig. 6a) and there only clearly in the Atlantic and North American region. In the Pacific (Fig. 6b), there was virtually no correlation between predicted and an-

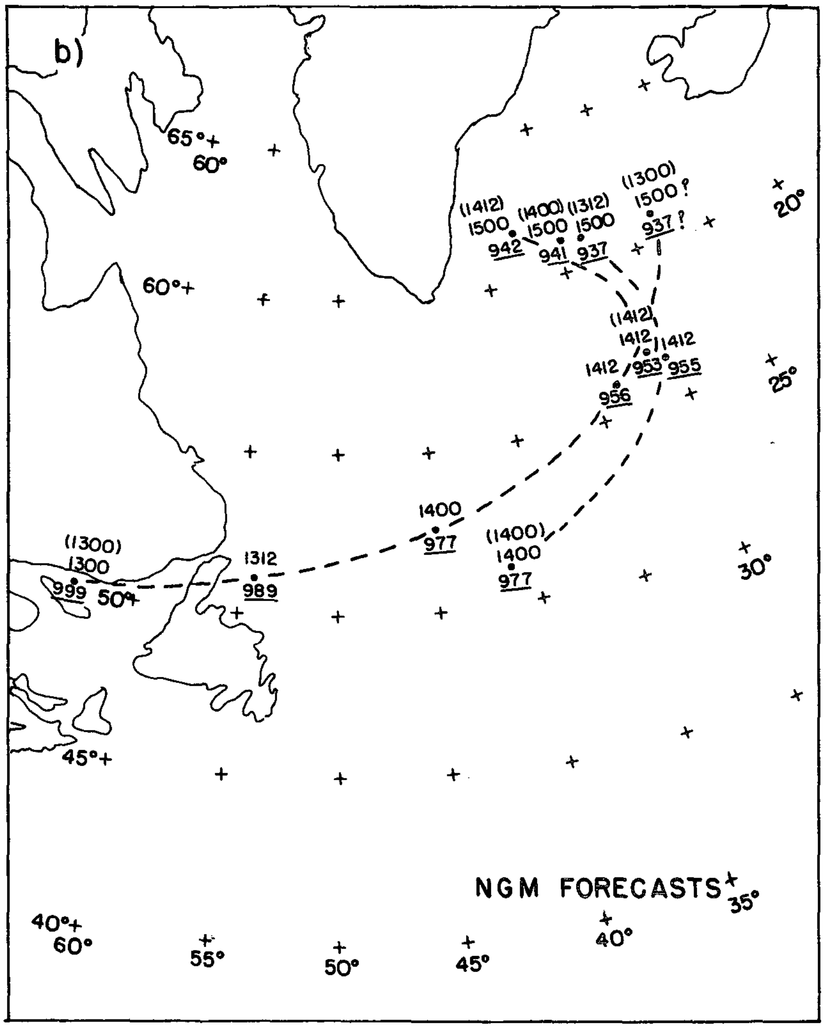


FIG. 5. (*Continued*)

alyzed central pressure falls, because of numerous 12-h timing errors in the forecasts of explosive deepening (especially with storms entering the C-grid from the overlap area). In the overlap area itself, the performance appears to be roughly comparable to the results for the LFM in 1981–84, although the older studies excluded instances in which the analyzed 12-h deepening was smaller than 10 mb. We may hope that NGM performance in the Pacific will improve in this respect, with the expansion of the C-grid (Hoke, 1987).

There is a dramatic Pacific deficit in the forecasts produced by the GLBL model as well (Figs. 6c and d). Here (Fig. 6d) the results do not look obviously better than those in the overall area for the 7-L PE model in 1978–79 (Sanders and Gyakum, 1980). On the other hand, in the Atlantic and North American area, seen in Fig. 6d, the current results represent a distinct improvement over the LFM performance for 1981–84. When the current GLBL performance is evaluated only in the area of the earlier LFM study (Fig. 6e), the improvement is even greater.

Mean shortfalls in NGM-predicted deepening ranged from 1.1 mb for C-grid forecasts outside the Pacific region to 6.8 mb for the overlap area. In the GLBL forecasts, the average shortfalls were 6.0 mb in the Pacific region and 3.1 mb elsewhere. Some shortfall is to be expected even if the forecasts were not biased overall, because the case selection (based on bomb events in the analyses) is biased toward large deepening in the analyses.

4. Discussion

The question arises why this improvement has occurred. In the case of the NGM, examination of a limited number of predictions of oceanic cyclogenesis in the period February–April 1986, immediately following its operational implementation, suggested a perfor-

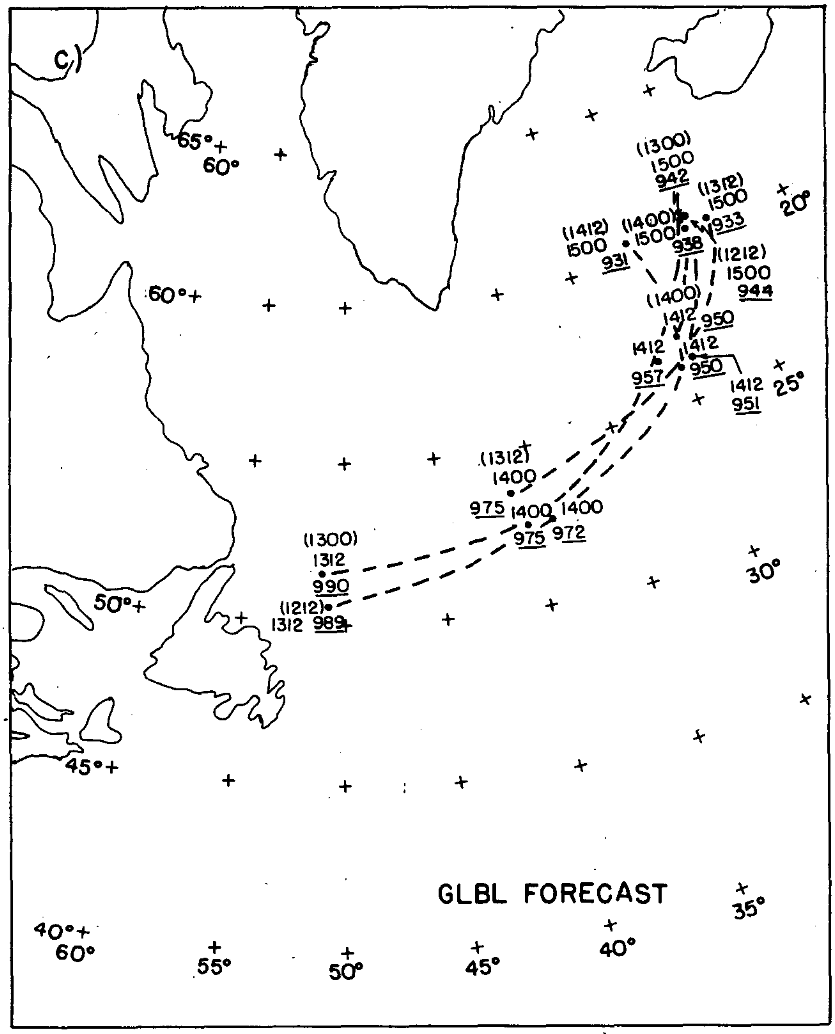


FIG. 5. (*Continued*)

mance no better than that demonstrated for the LFM and perhaps worse. The sample size, however, was small.

A number of changes in the initial analysis for the NGM (DiMego, 1987) were made early in 1987 (Petersen, personal communication), producing a more highly resolved first guess and more effective use of satellite soundings. Difficulties were encountered with the satellite data, however, in November and again near the end of the period of study. A comprehensive change in the model since the preceding season was the installation of a more complete (and presumably improved) formulation of radiative and mixing processes and of surface fluxes of heat, moisture and momentum, as well as a modification of the representation of cumulus convection. These were reported by Tucillo and Phillips (1986). Although the motivation for these changes was mainly improved prediction of precipitation over land and an effective simulation of the diurnal variation of the boundary layer over land, collateral improvements in oceanic forecasting seem to have occurred as well.

In the case of the GLBL, a cursory examination of model performance during the 1985/86 cold season indicates that numerous bombs were predicted then as well, although no attempt was made to determine how successfully. Substantial improvement has occurred (except in the Pacific) in the 8 years since the time of the 7-L PE forecasts evaluated by Sanders and Gyakum (1980). This must be attributed to the considerable improvement of spatial resolution and representation of physical processes which has occurred in this model, but further study would be required to determine when and why the major increases in skill occurred.

Differences in present skill in different domains also call for explanation. Evidently, the simultaneous calculation in the NGM fine mesh and adjacent coarser mesh vitiates any direct advantage of the fine mesh in

TABLE 4. Instances of extreme deepening, in excess of 20 mb in 12 h. MM denotes month. The forecast range refers to the end of the 12-h period denoted by DD date and UU time (UTC).

(a) Analyses

Storm no.	MM	DDUU	Deepening (mb)
8	10	2512	20
10	11	0812	20
11	11	1000	22
13	11	1400	20
	11	1412	26
14	11	1712	22
21	11	1800	24
22	11	2600	29
25	12	0900	20
28	12	1412	46
	12	1500	28
32	12	3100	22
39	1	2012	21
40	1	2300	23
43	1	2712	20
45	1	3100	20
49	2	0400	22
50	2	0612	22
58	2	2312	24
65	3	3100	20
66	4	0500	20
68	4	1500	24

(b) NGM 12-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
8	10	2512	27
14	11	1412	23
22	11	2700	21
25	12	0900	27
28	12	1412	22

(c) NGM 24-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
1	9	2312	20
8	10	2512	24
13	11	1412	28
22	11	2700	20
25	12	1812	21
28	12	1400	22

(d) NGM 36-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
13	11	1412	22
14	11	1812	20
28	12	1412	21
	12	1500	24

(e) NGM 48-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
14	11	1812	20
25	12	0900	20

(f) GLBL 12-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
8	10	2512	25
14	11	1712	21
21	11	2700	20
25	12	0812	22
	12	0900	28
26	12	1200	22
28	12	1412	27
50	2	0612	22
65	3	3112	24

(g) GLBL 24-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
8	10	2512	24
10	11	0812	24
	11	0900	20
13	11	1412	24
25	12	0900	22
28	12	1412	25

(h) GLBL 36-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
8	10	2512	23
13	11	1412	20
25	12	0900	21

(i) GLBL 48-h forecasts

Storm no.	MM	DDUU	Deepening (mb)
11	11	1012	21
28	12	1412	21

the overlap area. The simple (if expensive) solution is to enlarge the area of the fine mesh (Hoke, 1987). This seems to be a matter of the technique and logistics of computation, to which the present type of study can make little direct contribution.

In one respect or another, however, both models show much worse performance in the Pacific than in the Atlantic. North America contributed few bombs during this season, in either the analyses or the forecasts, so the distinction was essentially between the two oceanic regions. Although no quantitative information is at hand, we have the impression that the coverage of observations by ships and aircraft, and perhaps satellite as well, is substantially poorer in the Pacific than in the Atlantic. This, if so, confounded the achievement of high skill in two ways: the initial conditions for the forecast were not as well known as they were in the Atlantic, nor were the final verification conditions. The important upwind conditions aloft, of course, were well documented for the Atlantic bombs by the North American rawinsonde network. Bombs in the eastern Pacific enjoyed no such benefit, except perhaps in the Gulf of Alaska.

There is a possibility that some physical ingredient, perhaps convection, is not well handled in the model and is more important in perhaps the southern half of the cases. For examination of this question, forecast results for the GLBL model were summed over the forecast periods shown in Table 2 and then separated in each of the two areas by the modal latitude (40.5°N

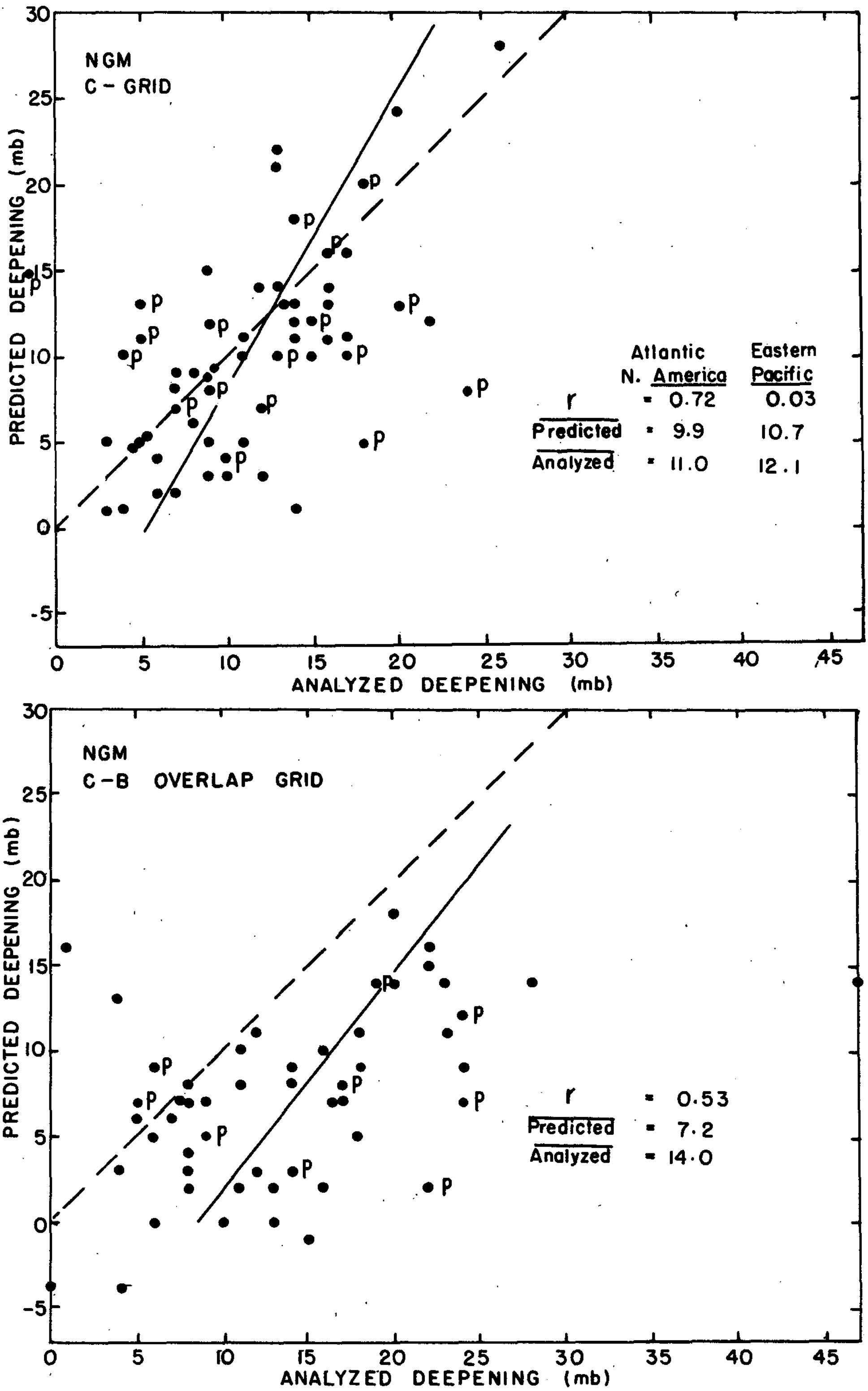
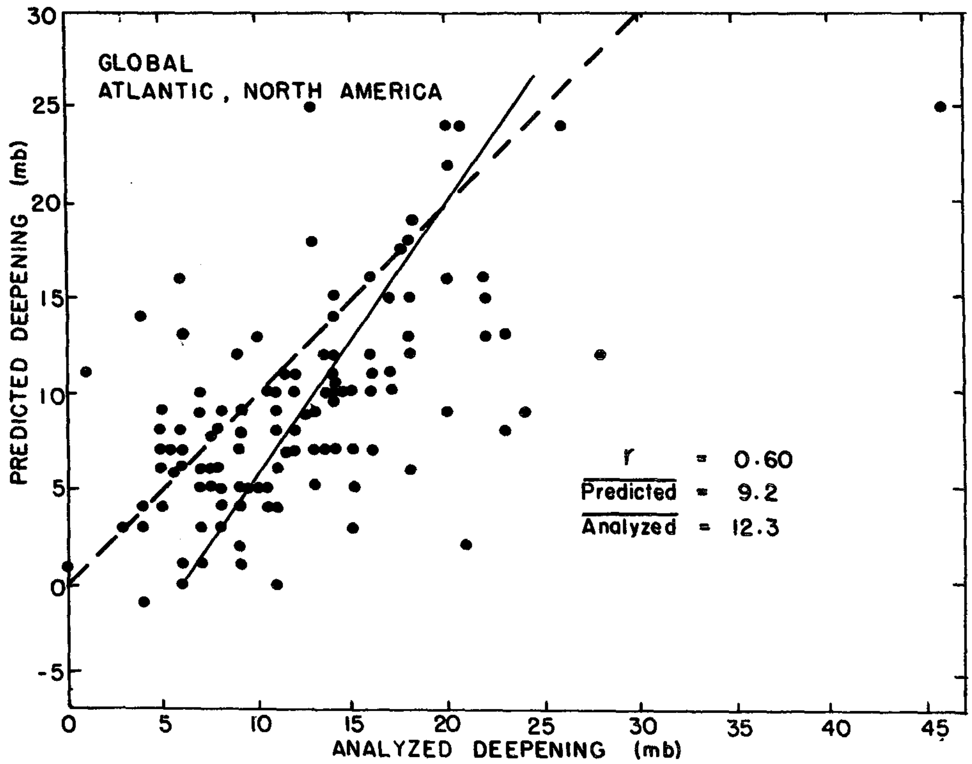


FIG. 6. Scatter diagram of predicted and analyzed deepening for cases of explosive cyclones in the analysis atmosphere. The forecasts are for the period 12–24 h after initialization. The dashed line represents a perfect forecast. The solid line is the regression line for analyzed deepening, given the forecast value. All points are used to determine this line, except in panel a, where the points for Pacific cases, labeled P, are excluded. Mean values of the analyzed and predicted deepening are given, together with the correlation coefficient *r*. (a) NGM forecasts in the C-grid; (b) NGM forecasts in the mesh area; (c) GLBL forecasts in the Atlantic and North American region; (d) GLBL forecasts in the eastern Pacific region and in the Atlantic and North American portion of the C-grid.



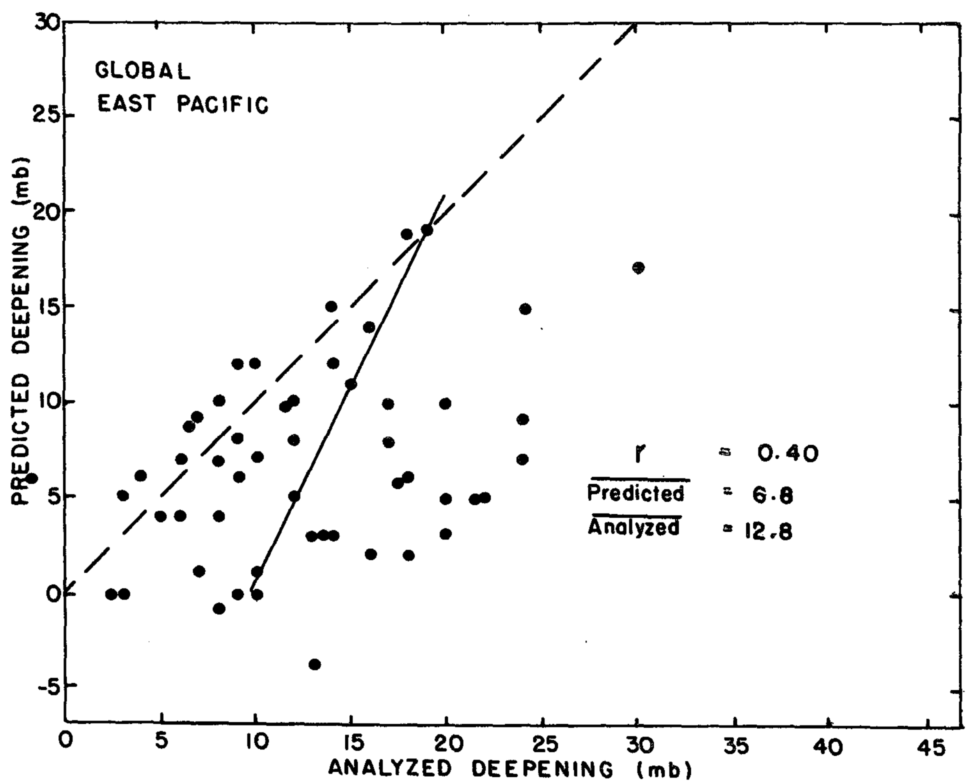


FIG. 6. (*Continued*)

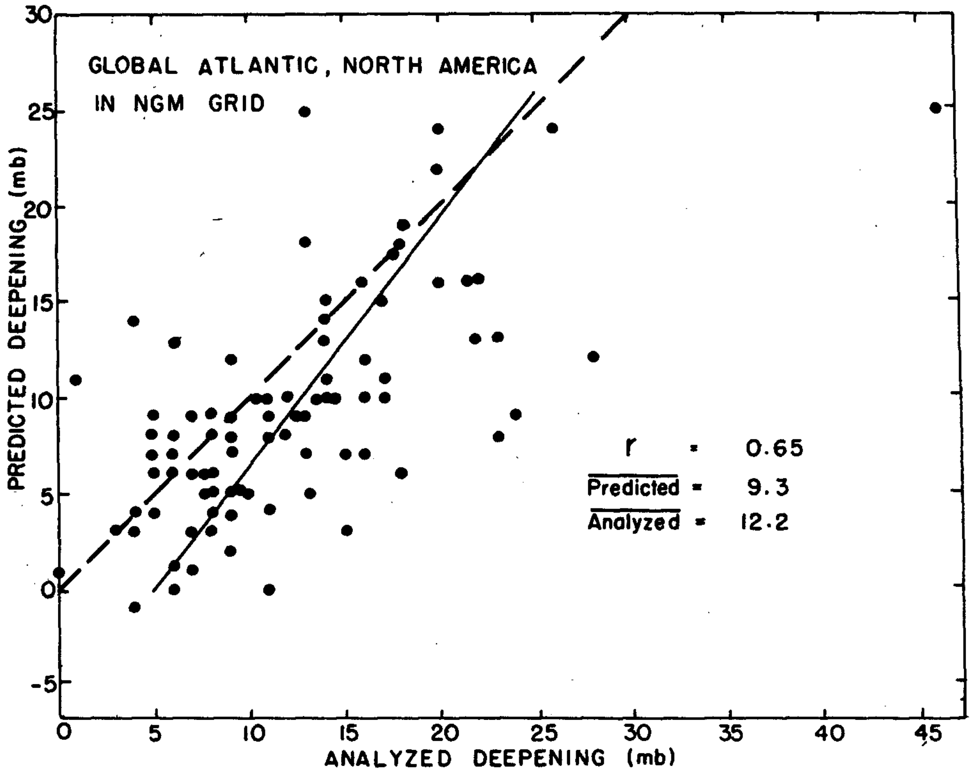


FIG. 6. (*Continued*)

for the Atlantic and North America and 45.5°N for the Pacific). The values of CSI were lower for the southern half of the cases in both areas: 0.304 vs 0.353 for the Atlantic and North America and 0.092 vs 0.229 for the Pacific. Thus, although we cannot rule out the possibility of a physical difference with latitude, an additional factor (presumably data coverage) is likely responsible for the Pacific deficiency, since the northern Pacific fares worse than the northern Atlantic. This factor may play a role in the latitudinal decrease of skill as well.

Over North America, performance in both models was relatively poor, but in a small number of cases. After summation over the forecast ranges, POD for the NGM was 4/9 and for the FAR was 10/14, yielding a CSI value of 0.21. For the GLBL model, POD was 1/12 and FAR was 4/5, for a CSI value of 0.06. If these cases were excluded from the sample, the GLBL Atlantic CSI values rose about 0.02. The excessive fraction of false alarms suggests that the tendency for overprediction of cyclone depth over land, first noted in the NMC models by Leary (1971), has not been entirely eliminated.

Despite the limitations just discussed, the present skill in both models is substantial. While a more complete study of the predicted vs analyzed deepening or filling of all cyclones would be required to establish optimum ways of inferring explosive cyclogenesis from the dynamical forecasts, the present results are certainly sufficient to indicate that no physical ingredient critical for oceanic cyclogenesis is missing in the models as they now stand. (Orographic development may well be another matter.) Explosive cyclogenesis over the sea is therefore not a mystery in the sense that the models are helpless in dealing with the phenomenon. The research problem before us now is to improve the skill and to extend the range of predictability of this event. To do this may require a better understanding than we now have of why the models do as well as they do, an irony in view of the traditional scientific view of the relationship between understanding and prediction.

An attractive approach to the research problem is to study the small static and symmetric stability of the environment of the bomb, a factor indicated by Reed and Albright (1986) to be critical. In particular, it would be interesting to investigate whether the responsible fluxes of heat and moisture from the sea surface act well in advance to condition the atmosphere for subsequent explosive cyclogenesis, as suggested recently by Namias (personal communication), or whether they act concurrently, as argued by Emanuel (1986) for tropical cyclones.

Even a casual look at the chronology of explosive cyclogenesis given in Figs. 2–4 suggests that these events occur in regimes rather than randomly in time. Note, for example, the rarity prior to the second week in

November, followed by frequent cyclogenesis in both oceans until the beginning of December. Thereafter, in the Pacific, discrete episodes occurred around the beginnings of January, February, and April, with little activity at other times. In the Atlantic there was frequent cyclogenesis during the middle of December, then only sporadic activity until the last week of January. Afterwards, activity was frequent until the last week of February, marking the effective end of the season in this ocean. Traditional ideas concerning the planetary-scale aspects of cyclogenesis in general appear to be applicable to bombs in particular.

A study of either observations or model output could shed light on these questions. Sensitivity to data type and coverage is a question of the greatest practical interest, and needs to be studied in the framework of a large analysis and prediction system of the type in place at NMC.

5. Concluding summary

Prognostic charts for the Atlantic, North America and eastern Pacific, derived from the NGM and global-spectral models initialized at NMC and received on DIFAX, were examined for the occurrence of explosively deepening cyclones for September 1986–April 1987. The criterion for an explosive event was a 24-h deepening of 24 mb at latitude 60°N, normalized geostrophically for other latitudes (one Bergeron), beginning at 0000 or 1200 UTC. Events were noted in forecast periods ranging from 0–24 to 24–48 h for the NGM and from 0–24 to 36–60 h for the global model. Manual analyses from NMC, also received on DIFAX, were used for verification.

Skill in prediction of the one-Bergeron event, in terms either of the Critical Success Index (CSI) or the Brier score for probability forecasts (S), with an estimated climatological control, was found at all ranges in both models. There was a high correlation between CSI and S, the latter being about 0.05 lower. The CSI values for the NGM forecasts exceeded 0.39 for the 0–24 and 12–36-h forecasts in the innermost fine grid and as low as 0.15 in the 24–48-h forecasts in the area where the finest grid and the surrounding coarser one overlap. In the global forecasts, CSI ranged from a maximum of 0.43 in the 12–24-h forecasts in the Atlantic and North American region to a minimum of 0.10 in the 36–60-h forecasts for the eastern Pacific. Explosive cyclogenesis over North America was rare in the models and even rarer by far in the analyses, leading to very low CSI values in both models.

Extreme deepening, at rates of at least 20 mb in 12 h, occurred in both models, but with a frequency that decreased with range and was substantially less than the frequency in the analyses even at the shortest range. For the period 12–24 h after initialization, predicted deepening was correlated with the corresponding deepening in the analyses, for instances in which explosive cyclogenesis occurred in the latter. Correlations for the NGM were as high as 0.72 in the Atlantic portion of the fine mesh but only 0.03 in the eastern Pacific portion, owing to frequent 12-h errors in timing. Average forecast deepening fell short of analyzed 12-h deepening by only 1.1 and 1.4 mb in the two regions. Correlations in the global-model forecasts were 0.65 and 0.40 for the Atlantic and Pacific regions, respectively, with respective shortfall in predicted deepening of 3.1 and 6.0 mb.

Although no systematic study of the causes for improvement in forecasts over the past 8 years was undertaken, it appears that improved analysis, increased horizontal and vertical model resolution and improved model representation of surface and boundary-layer fluxes, and of the effects of cumulus convection, all contributed. In particular, the introduction of a more sophisticated formulation of the roughness length over the sea in the past year (Phillips, 1986) appears to have been especially beneficial in the NGM.

Acknowledgments. The author is grateful to the Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology; to Thomas McGuire at the NWS Forecast Office, Boston; and to Peter Leavitt, Weather Services Corporation, for providing facilities and access to maps. He also wishes to thank Isabelle Kole for preparation of figures and Ralph Petersen, John Brown, and James Tucillo, NMC, for discussion of recent changes in analysis procedures and in the forecast models. This research was supported by the Office of Naval Research under Contract N00014-C-0785.

APPENDIX

Relationship between CSI and Skill Measured by Brier Score

The forecast skill, as measured by the Brier Score, is the fractional improvement of a set of probability forecasts which vary from day to day over a single repeated forecast of the climatological probability. In the present application this climatological forecast was obtained from an estimate of the relative frequency of bombs from the appropriate sample. The Brier Score for this set of forecasts was compared to the score for a set in which the probability took on one of two values, depending on whether or not the model predicted a bomb. These probabilities were likewise taken from the results of the present sample.

In this study, successful negative cases were not counted, so the total number of cyclones had to be estimated. It was assumed that over the 484 times examined, three individual cyclones were initiated in or entered the NGM C-grid each two times. In the mesh area the number of cyclones was assumed to be 75% of this value. In the GLBL Atlantic and North American area two individual cyclones were assumed to enter or develop each synoptic time. The number was presumed to be half this value in the Pacific area.

There proved to be a close correlation ($r = 0.97$) over the 14 pairs (three in each NGM category and four in each GLBL category) between CSI and the skill S. With the quantities expressed as decimal fractions, the linear regression equation is

$$S = -0.056 + 0.97\,\mathrm{CSI}.$$

Evidently in this instance it makes little difference whether the negative successes are counted or not. To test the sensitivity of S to the assumed number of cyclones, it was recalculated for the 24-h GLBL Atlantic and North American area on the assumption of only half as many cyclones. The original skill of 0.36 dropped to 0.34. Evidently the sensitivity is not great. On the basis of either score, modest skill is apparent, extending to the maximum range of the available prognostic charts.

REFERENCES

Abbey, R. F., R. Hadlock and C. W. Kreitzberg, 1987: ERICA Overview Document, 41 pp. [Available from Office of Naval Research, Code 1122, 800 N. Quincy St., Arlington, VA 22217-5000.]

Anthes, R. A., Y.-H. Kuo and J. R. Gyakum, 1983: Numerical simulation of a case of explosive marine cyclogenesis. *Mon. Wea. Rev.*, **111**, 1174–1188.

Burt, S. D., 1987: A new North Atlantic low pressure record. *Weather*, **42**, 53–56.

Chen, T.-C., C.-B. Chang and D. J. Perkey, 1983: Numerical study of an AMTEX '75 cyclone. *Mon. Wea. Rev.*, **111**, 1818–1829.

DiMego, G. J., 1987: The National Meteorological Center Regional Analysis System. *Mon. Wea. Rev.*, **115**, (in press).

Donaldson, R. J., Jr., R. M. Dyer and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. *Proc. Ninth Conf. on Severe Local Storms*, Norman, Amer. Meteor. Soc., 321–325.

Douglas, S. G., and T. T. Warner, 1987: Utilization of VAS satellite data in the initialization of an oceanic-cyclogenesis situation. *Mon. Wea. Rev.*, **115**, (in press).

Emanuel, K. A., 1986: An air–sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. *J. Atmos. Sci.*, **43**, 585–604.

Hoke, J. E., 1987: Improving the horizontal resolution of the Nested-Grid Model. Technical Procedures Bulletin No. 368, National Weather Service, Silver Spring, MD 20910.

Leary, C., 1971: Systematic errors in operational National Meteorological Center primitive-equation surface prognoses. *Mon. Wea. Rev.*, **99**, 409–413.

Liou, C.-S., and R. L. Elsberry, 1987: Heat budgets of analyses and forecasts of an explosively deepening maritime cyclone. *Mon. Wea. Rev.*, **115**, 1809–1824.

Nuss, W. A., and R. A. Anthes, 1987: A numerical investigation of low-level processes in rapid cyclogenesis. *Mon. Wea. Rev.*, **115**, (in press).

Orlanski, I., and J. J. Katzfey, 1987: Sensitivity of model simulations for a coastal cyclone. *Mon. Wea. Rev.*, **115**, (in press).

Phillips, N. A., 1986: Surface mixing and oceanic drag coefficient in the nested-grid model. Technical Procedures Bulletin No. 360, National Weather Service, Silver Spring, MD 20910.

Reed, R. J., and M. D. Albright, 1986: A case study of explosive cyclogenesis in the Eastern Pacific. *Mon. Wea. Rev.*, **114**, 2297–2319.

Sanders, F., 1973: Skill in forecasting daily temperature and precipitation. *Bull. Amer. Meteor. Soc.*, **54**, 1171–1179.

——, 1986a: Explosive cyclogenesis over the west central North Atlantic Ocean, 1981–84. Part I: Composite structure and mean behavior. *Mon. Wea. Rev.*, **114**, 1781–1794.

——, 1986b: Explosive cyclogenesis over the west central North Atlantic Ocean, 1981–84. Part II: Evaluation of LFM model performance. *Mon. Wea. Rev.*, **114**, 2207–2218.

——, and J. R. Gyakum, 1980: Synoptic–dynamic climatology of the "bomb." *Mon. Wea. Rev.*, **108**, 1590–1606.

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.*, **110**, 254–271.

Tucillo, J. J., and N. A. Phillips, 1986: Modeling of physical processes in nested grid model. Technical Procedures Bulletin No. 363, National Weather Service, Silver Spring, MD 20910.

Uccellini, L. W., R. A. Petersen, K. F. Brill, P. J. Kocin and J. J. Tucillo, 1987: Synergistic interactions between and upper-level jet streak and diabatic processes that influence the development of a low-level jet and a secondary coastal cyclone. *Mon. Wea. Rev.*, **115**, (in press).