Skill in Prediction of Explosive Cyclogenesis over the Western North Atlantic Ocean, 1987/88: A Forecast Checklist and NMC Dynamical Models

FREDERICK SANDERS

9 Flint Street, Marblehead, Massachusetts

EUGENE P. AUCIELLO

NOAA, NWS Forecast Office, Boston, Massachusetts

(Manuscript received 19 October 1988, in final form 21 February 1989)

ABSTRACT

Analyses and predictions of explosive cyclogenesis over the western North Atlantic Ocean during the 1987/88 cold season were compared. The analyses were the manual and automated series produced at the National Meteorological Center (NMC). The forecasts were those produced by the nested grid model (NGM) and the "aviation run" of the global spectral model (AVN) at NMC, and also by a simple checklist employed by the National Weather Service (NWS) Forecast Office, Boston.

Skill of the forecasts has evidently improved since the preceding year. Probability of detection of an event in a specified 24-h period, with the manual analyses used as verification, approached 72% for the NGM in the range of 0–24 h with a false alarm rate of 17%. In the range of 36–60 h, the values for the AVN forecasts were 42% and 30%. When the automated analyses were used for verification, forecast performance was somewhat better.

The accuracy of the checklist forecasts was comparable to that of the AVN forecasts but not as good as that of the NGM predictions, in the small sample available for comparison.

Deepening in the NGM forecasts over the range of 12–24 h was 2 mb less than in the manual analyses, with a correlation of 0.55. The accuracy was limited mainly by errors in timing, with the model failing to well represent the initial analyzed deepening but catching up later. The automated analyses displayed a similar failure, with a correlation of 0.49 between analyses. Uncertainty in initial analysis is a major factor limiting present accuracy, especially at short range.

1. Introduction

A considerable recent increase in skill of operational dynamical models at the National Meteorological Center (NMC) in the prediction of explosive cyclogenesis was found by Sanders (1987). Because this skill appeared to be most pronounced in the western North Atlantic region, and because the field-project phase of the Experiment on Rapidly Intensifying Cyclones over the Atlantic (Hadlock and Kreitzberg 1988) will have occurred during the 1988/89 season, it is important to have a continuous record of this skill. We therefore evaluated the performance of the nested-grid model (NGM) and the global spectral model during the 1987/88 cold season.

We also verified the forecasts made by employment of a checklist developed by Auciello and Sanders (1986) from the results of a 1985 workshop on oceanic storms reported by Lange et al. (1986). This was designed as an objective operational technique for the forecasting of explosive cyclogenesis up to 36 hours prior to the event over an area between 38° and 45°N, and 55° and 75°W. It has been used by marine forecasters at the National Weather Service Forecast Office (WSFO) in Boston during the past three cold seasons.

2. Data and procedures

The Massachusetts Institute of Technology (MIT) files of surface maps from September 1987 through April 1988 were examined for the occurrence of explosive cyclogenesis in the area between 30° and 55°N and between the east coast of North America and longitude 50°W. Explosive deepening was considered to have occurred when the observed 24-h deepening was at least 24 mb after the deepening (mb) was multiplied by the ratio of the sine of the observed latitude divided

Corresponding author address: Frederick Sanders, 9 Flint Street, Marblehead, MA 01945.

TABLE 1. Cyclones that reached ≥1.0 bergerons in either the NH or FH analysis series, or in the NGM or AVN model output, as received on DIFAX.

					NGM			AVN		
No.	Date	DDGG†	NH§	FH@	12	24	36	24	36	48
1	3–4 Oct 1987	0400	Y#	N	N	N	N	N	N	N
2	11–12 Oct 1987	1212	N	N	N	N	N	N	N	Y
3	22–23 Oct 1987	2300	Y	Y	Y	Y	X‡	Y	Y	Y
4	5–7 Nov 1987	0600	N	N	Y	Y	N	N	Y	Y
		0612	N	N	Y	Y	Y	Y	Y	N
5	11–13 Nov 1987	1112	Y	N	N	N	N	N	N	N
		1200	Y	Y	Y	Y	N	Y	N	X**
		1212	Y	Y	Y	Y	N	N	N	N
6	20–21 Nov 1987	2012	Y	Y	Y	Y	Y	Y	Y	Y
		2100	N	N	N	N	Y	Y	N	Y
7	21–22 Nov 1987	2200	N	N	Y	N	N	N	N	N
8	29 Nov–1 Dec 1987	3000	Y	N	N	N	N	N	N	N
		3012	N	N	Y	N	N	N	N	X**
9	4–6 Dec 1987	0500	N	N	N	N	N	N	N	Y
10	11–12 Dec 1987	1100	Y	N	Y	N	N	N	N	N
		1112	Y	Y	N	N	N	N	N	N
11	16–17 Dec 1987	1600	Y	Y	N	N	N	Y	Y	Y
		1612	Y	Y	N	N	Y	N	N	N
12	22–25 Dec 1987	2300	Y	Y	Y	N	N	N	N	Y
		2312	Y	Y	Y**	Y**	Y**	Y	N	X
		2400	Y	Y	X**	X**	X**	Y	Y	X**
		2412	Y	Y	X**	X**	X**	Y	Y	X**
13	25–27 Dec 1987	2600	N	N	N	N	N	N	Y	Y
		2612	Y	Y	X**	Y	Y	Y	Y	Y
14	29–31 Dec 1987	2912	Y	Y	Y	Y	Y	Y	Y	Y
		3000	Y	Y	Y	Y	Y	Y	Y	Y
		3012	N	Y	N	N	Y	N	N	N
15	2–3 Jan 1988	0212	N	N	N	N	N	N	N	Y
16	4–6 Jan 1988	0412	Y	Y	Y	Y	N	Y	Y	Y
		0500	Y	Y	Y	Y	N	Y	X	Y
		0512	Y	Y	Y	Y	Y	Y	Y	X
		0600	Y	Y	Y	Y	N	Y	Y	Y
17	16–17 Jan 1988	1612	Y	Y	X**	Y**	Y**	Y	Y	Y**
18	25–27 Jan 1988	2600	Y	Y	Y	Y	N	N	X	X
		2612	Y	Y	X	N	N	Y	N	X
19	2–3 Feb 1988	0300	N	N	Y	Y	Y	N	N	N
20	4–6 Feb 1988	0412	Y	N	Y	X	N	N	N	N
		0500	Y	Y	Y	Y	N	Y	N	Y
		0512	Y	Y	Y	Y	Y	Y	N	X**
		0600	Y	Y	Y	Y	N	N	N	N
21[i]	6–8 Feb 1988	0700	Y	N	X**	X**	N**	N	N	X**
		0712	N	Y	Y	X**	X**	N	Y	X**
		0800	Y	Y	Y	Y	X**	Y	Y	X**
22[k]	12–14 Feb 1988	1212	Y	Y	Y	Y	Y	Y	Y	N
		1300	Y	Y	Y	Y	N	N	N	Y
		1312	N	N	X	N	N	N	N	N
23	4–6 Mar 1988	0500	Y	N	N	N	N	N	N	N
		0512	Y	Y	Y	Y	Y	X	N	N
		0600	Y	Y	Y	Y	N	N	N	N
24	7–9 Mar 1988	0800	Y	Y	Y	Y	Y	Y	Y	X**
		0812	Y	Y	Y	Y	Y	Y	Y	X**
25	10–11 Mar 1988	1100	Y	Y	N	N	N	N	N	N
26	14–16 Mar 1988	1500	Y	Y	N	N	X**	N	N	N
		1512	Y	Y	Y	N	N	N	N	X**
27	19–21 Mar 1988	1912	Y	Y	Y	N	Y	Y	Y	Y
		2000	Y	Y	Y	Y	Y	Y	Y	Y
		2012	Y	Y	Y	N	Y	Y	N	N
28	29–30 Mar 1988	3000	Y	N	X**	X**	X**	N	N	X**
29	19–21 Apr 1988	2000	Y	X	N	N	N	N	N	N**
		2012	Y	X	N	N	N	N	N	N**

† DDGG is date and UTC of center of 24-h period.
§ NH is manual Northern Hemisphere analysis.
@ FH is automated "front-half" analysis.
# Y is deepening of at least 1.0 bergeron; N means otherwise.
‡ X is missing data.
* is FS analysis.
** is storm off east edge of map. Estimate of center could sometimes be made.
[i] Storm #21 was east of 50°W from 0718 onward.
[k] Storm #22 data were interpolated at 1300 and 1312.

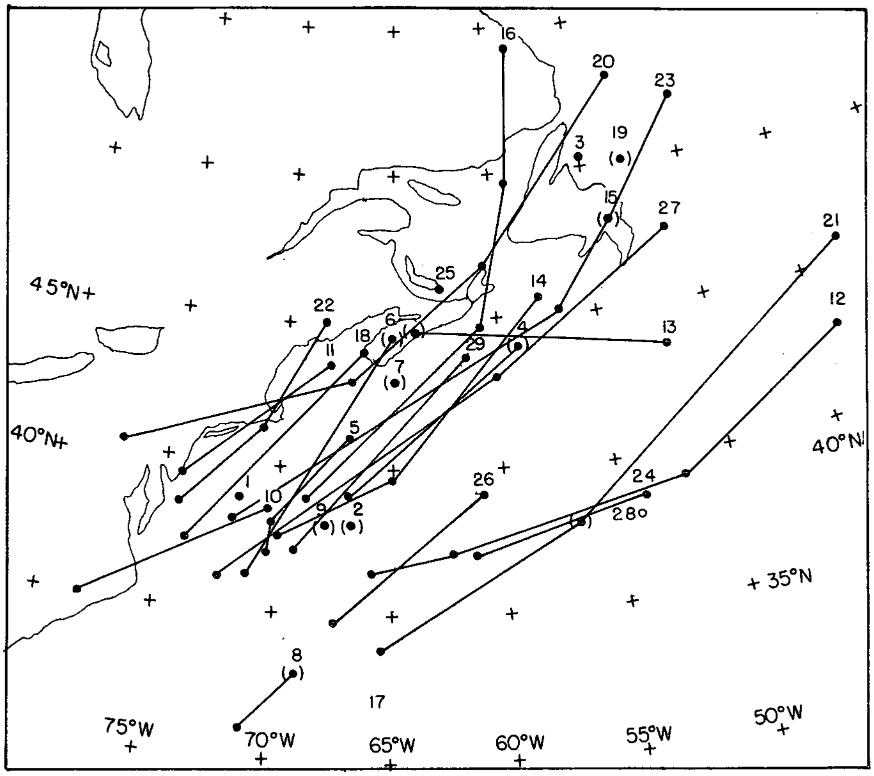

FIG. 1. Positions of cyclones at the center of 24-h periods when explosive deepening was analyzed in the NH series or was predicted by one or both of the NMC models. Positions for overlapping periods for the same cyclone are connected by solid line. Parentheses indicate that explosive deepening occurred in a forecast but not in the analysis. The number of the storm is shown as listed in Table 1.

by the sine of 60° (at least one bergeron as defined by Sanders and Gyakum 1980).

Data from forecast maps produced each 12 hours by two models, the NGM and the "aviation run" (AVN) of the global spectral model, were compared with data from two sets of analyses from NMC, the manual Northern Hemisphere series available every 6 hours (NH) and the automated series available every 12 hours, for the limited "front half" (FH) of the hemisphere extending from the western Pacific region to the eastern Atlantic region.

The checklist, which was initially designed to use output from the limited-area fine-mesh model, is now applied to NGM analyses and forecasts, as described by Auciello and Sanders (1987). It comprises six categorical questions:

1) Does a 500-mb absolute vorticity maximum of 17×10^{-5} or more exist in the NGM initial analysis in an area from 30° to 50°N and 85° to 110°W?

2) Does this maximum maintain initial intensity or strengthen on successive 12-h NGM forecast charts out to 48 h?

3) Is this maximum forecast to move at an average of 30 kt or more through 48 hours?

4) Does this maximum cross the coast between 32° and 45°N during the forecast interval?

5) Does a jet streak of 110 kt or greater exist at 250 or 300 mb within a 300-nm (550-km) radius in the semicircle south of the initial vorticity maximum?

6) Does the NGM develop a surface low of 990 mb or deeper during the next 48 hours over an area from 38° to 45°N and 55° to 75°W?

Explosive cyclogenesis is predicted when four or more of the above questions are answered affirmatively. This criterion was selected subjectively through use of the checklist during preceding seasons. No study has been made to determine if the criterion is optimal or to what extent the questions may be redundant. Checklist forecasts made each 12 hours from 1 October 1987 to 31 March 1988 were evaluated by reference to the NH analyses.

In two instances, slight reanalysis of the verifying maps was done, reflecting the influence of later observations. Otherwise the latitude, longitude, and central

TABLE 2. For the 1987/88 season, events (E), hits (H), false alarms (FA) probability of detection (POD), false-alarm rate (FAR), and critical success index (CSI).

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 Western Atlantic vs NH											
40 .72	29 .17	6 .63	42 .62	26 .10	3 .58	41 .37	15 .21	4 .33			
AVN Western Atlantic vs NH											
			44 .57	25 .07	2 .54	45 .40	18 .18	4 .37	33 .42	14 .30	6 .36
NGM Western Atlantic vs FH											
34 .82	28 .22	8 .67	36 .69	25 .11	3 .64	33 .52	17 .15	3 .47			
AVN Western Atlantic vs FH											
			38 .66	25 .07	2 .62	37 .51	19 .14	3 .48	26 .58	15 .29	6 .47

pressures of cyclones were taken as given on the maps, subject to some difficulty in reading values at times when clarity was a problem or when the cyclone was very near the edge of the map area. The NGM display area did not extend eastward to 50°W in the lower latitudes, and the 60-h AVN display was slightly more severely shortened north of 40°N. Some data were lost because of these limitations, and some because the maps were missing from the MIT archive.

A chronological listing of all 24-h periods when explosive deepening was occurring in one or more of the map series, or in one of the relevant forecasts, is shown in Table 1. (In the NH analyses, 24-h periods beginning at 0600 or 1800 UTC were not examined.) The time range of the forecasts is the center of the 24-h forecast period. Thus, 12 refers to forecasts from initialization time to 24 hours later, and 48 refers to forecasts from 36 to 60 hours after initialization. (There was no evaluation of 12-h forecasts for the AVN because maps were not available for the initialization time.) Some storms produced only one such 24-h period, but others showed explosive deepening in the analyses over as many as four overlapping 24-h periods spanning 60 hours. Each period was considered separately in the detailed analysis.

The locations of the cyclones in the NH analyses at the midpoint of each 24-h period when explosive deepening was occurring are shown in Fig. 1. The implied tracks are similar to those shown for earlier years (e.g., Sanders 1986), except that the center of the first period

TABLE 3. Same as in Table 2, but for the 1986/87 season. Comparison was with only the NH analyses.

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 C-grid vs NH											
36 .51	18 .38	11 .39	35 .49	17 .29	7 .40	35 .26	9 .36	5 .22			
AVN Atlantic and North America vs NH											
			44 .50	22 .24	7 .43	43 .35	15 .32	7 .30	36 .19	7 .50	7 .16

TABLE 4. As in Table 2, but for only the first 24-h period in which either NH-analyzed or predicted explosive cyclogenesis occurred, for the 1987/88 season. Comparison is with the NH analyses.

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						
19	11	3	20	9	2	20	7	1			
.58	.21	.50	.45	.18	.41	.35	.12	.33			
					AVN						
			23	10	0	22	10	2	19	9	5
			.44	0	.44	.46	.17	.42	.47	.36	.38

was located relatively far to the southwest. Note the concentration of storms that began explosive deepening in the area between 35° and 40°N and between 70° and 75°W.

3. Results

a. The 1-bergeron threshold

1) THE DYNAMICAL MODELS

The analysis of performance in individual 24-h periods is shown in Table 2, in which forecasts from each model are verified against each of the analyses series. Considering first the NH analyses as "ground truth", we find that the probability of detection (POD) in the NGM 12-h forecasts approached three out of four but lowered rapidly between 24 and 36 hours to only slightly more than one in three. A modest number of false alarms occurred at all ranges, most frequently at 12 hours, but the false alarm ratio (FAR) exceeded one in five only in the 36-h forecasts. The critical success index (CSI) (Donaldson et al. 1975) was 0.63 for the 12-h forecasts and was slightly lower at 24 hours, but then dropped substantially at 36 hours to only 0.33.

The AVN 24- and 36-h forecasts showed results comparable to those for the NGM—slightly worse at 24 hours and somewhat better at 36 hours. There was little further deterioration in the 48-h forecasts, a slight increase in POD being offset by an increase in FAR.

Skill in the 1987/88 forecasts is difficult to compare with the skill reported by Sanders (1987) for the preceding 1986/87 cold season, shown in Table 3. In the earlier study, the inner fine mesh of the NGM, over which the skill was seen to be greater (than outside), extended only a short distance off the East Coast during much of the season. The fine-mesh results in the coastal zone were combined with those from the eastern Pacific area, where the forecast skill is lower. The evaluation for the AVN in the earlier study was carried out over the present verification area as well as over the North American continent and the central Atlantic Ocean. On the other hand the verification method was more generous; the forecast of explosive cyclogenesis was considered successful when the model, at the specified range, predicted sufficient deepening at least once for a particular cyclone, whether or not the timing and duration was correct. In the present study, the analyzed and predicted deepening are compared separately for each 24-h period for each cyclone.

With these caveats in mind, we find in Table 2 substantial improvement in the skill of both models. In the more recent season, the PODs were higher and the FARs were lower at all ranges. Since the end of the 1986/87 cold season, the horizontal resolution of the AVN model was substantially increased. In the NGM model the expanded C-grid in the western Atlantic region was used for the entire season. While we have not attempted to identify what changes were responsible for the forecast improvement, these presumably helped.

When the 1987/88 forecasts are verified against the FH analyses, Table 2 shows that the number of hits and false alarms changes only slightly. The number of events, however, is sharply reduced, so that POD is improved and the values of CSI rise, substantially in a number of instances. In those cases when explosive

TABLE 5. Intercomparison of NGM and AVN predictions of explosive cyclogenesis with checklist predictions, October 1987–March 1988.

NGM			AVN			Checklist		
E POD	H FAR	FA CSI	E POD	H FAR	FA CSI	E POD	H FAR	FA CSI
14	13	2	14	10	3	14	11	5
.93	.13	.81	.71	.23	.59	.79	.31	.58

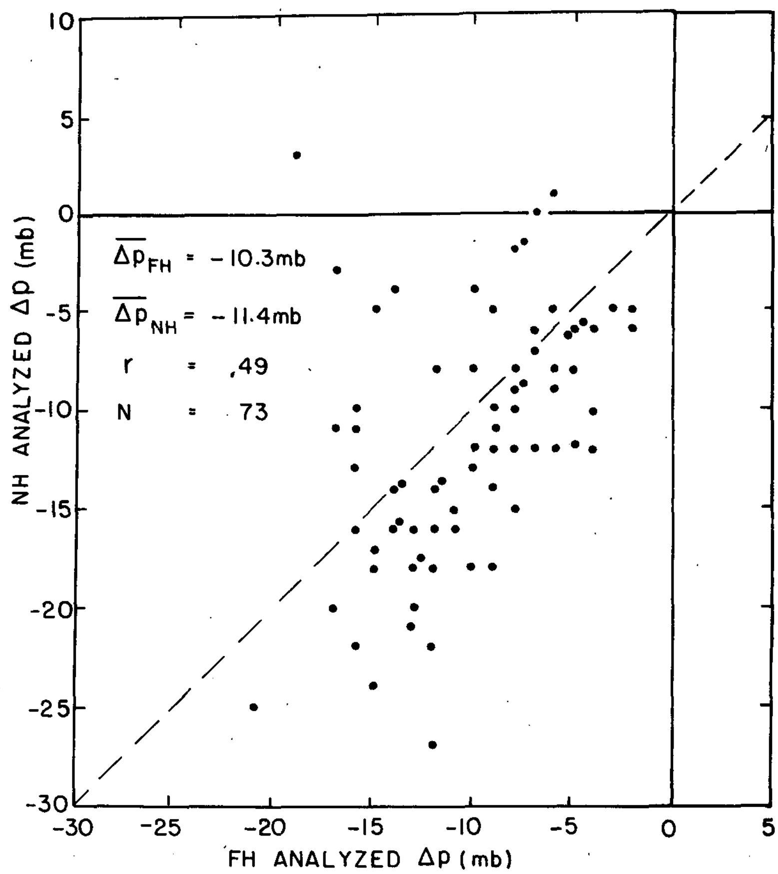

FIG. 2. (a) Scatter diagram of 12-h deepening in NH analyses vs in FH analyses. (b) Mean central pressure vs time, relative to the central pressure at the start of explosive cyclogenesis, for the FH and NH analyses. The number of pairs for each 12-h increment is shown in panel b. The number decreases with time owing to the smaller number of cyclones that maintain explosive deepening rates over longer times.

cyclogenesis occurred in the NH but not in the FH analyses, the plotted observations tended, with some exception, to support the stronger deepening in the former set. It appears that the FH analyses rely excessively on a first-guess, derived from the short-term forecast and not enough on the current observations. This increment of skill (when the forecasts are verified against the FH analyses instead of the NH analyses) is likely spurious. The FH analyses resemble the model forecasts more and the observations less, so that while the forecasts are really no better, the verifying analyses are worse. It is difficult to establish what actually happened at sea far from a dense and reliable network of observations.

In the process of collecting data, initiation of explosive deepening was perhaps not as well predicted as its continuance. To check this possibility, we evaluated the performance of the models for only the first 24-h period for each cyclone in which either NH-analyzed or predicted explosive deepening occurred. The results in Table 4 show that the probability of detection in the first 36 hours (i.e., in the 12-h and 24-h forecasts) was not as high for initial deepening as for later deepening, and that CSIs are commensurately lower. The effect is not seen in the forecasts for longer ranges. It is possible that the initial analysis was to blame; it tends to be particularly difficult at this time, and the forecast model may be unusually sensitive. At longer ranges, the model forecast may depend less on the detail of the initial analysis. This does not mean, however, that the model is better than the observations at predicting when explosive cyclogenesis is about to start; note that false alarms were relatively numerous in the 48-h AVN forecasts.

2) THE CHECKLIST

The checklist forecasts were not made or verified for specific 12-h periods. For a particular storm, an affirmative forecast was considered to have been made if

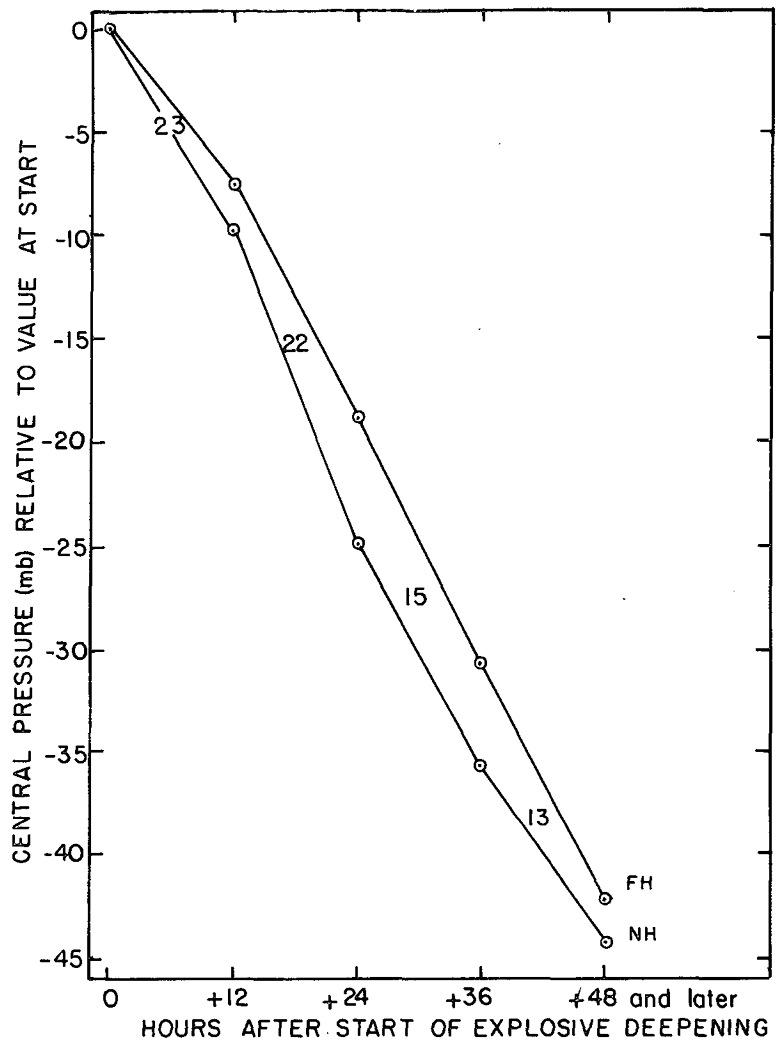

FIG. 2. (*Continued*)

one or more checklists referring to a particular storm contained affirmative answers for four of the six questions, as described above. The forecast was regarded as a hit if a cyclone qualified as explosive in any 24-h period centered within 36 hours after the initial time of one or more of the checklists referring to it. Otherwise it was a false alarm.

During the 1987/88 season, there were 15 explosive cyclogenetic events in the checklist verification area, of which 12 were hits. In Table 1, these were storms No. 1, 4, 5, 6, 10, 11, 12, 14, 16, 19, 26 and 27. Storms No. 18, 22 and 23 were misses. The total of six false alarms included storms No. 9 and 17 as well as additional cases on 20 December, 21 January, 16 February and 24–25 February. The remaining storms in Table 1 comprised eight that occurred outside the WSFO verification area, three that were false alarms in the model forecasts but not in the checklist forecasts, and one that occurred after 31 March. The POD for the checklist forecasts was thus 0.80, the FAR was 0.33, and the resulting CSI was 0.57.

This performance represents a considerable improvement over the preceding season, when there were 17 explosive events and POD, FAR, and CSI were 0.76, 0.48, and 0.45, respectively. Factors contributing to the improvement included greater familiarity of the forecasters with the procedure, the shift to the NGM forecasts as a basis for the checklist, and increase in the skill of the NGM in the verification area.

A direct comparison was made between model forecasts and checklist forecasts for storms in the sample common to both. A number of the model-forecast cases (Table 1, No. 3, 8, 13, 17, 21, 24, 25, and 28) were discarded because the explosive cyclogenesis occurred

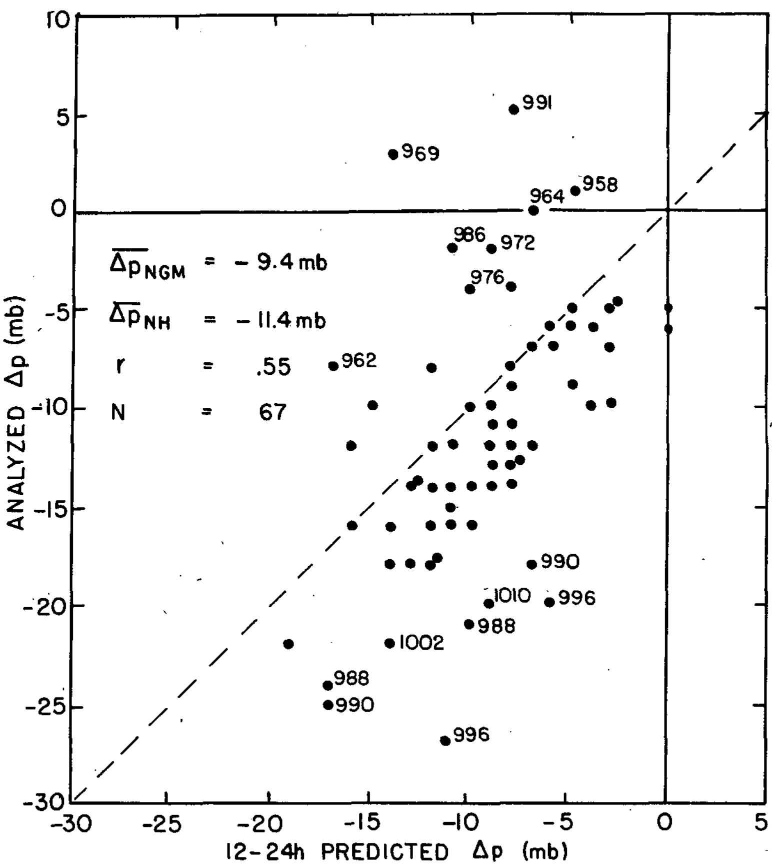

FIG. 3. As in Fig. 2., but for NH analyses vs 12–24 h segments of NGM forecasts. The values plotted by a number of points in (a) are analyzed central pressures at the start of the 12-h period. In (b), note that the forecast segments are taken from different forecast runs.

beyond the checklist verification area, while No. 4 (counted as a checklist case) was eliminated because the center was interpreted as being just over land and No. 29 because the WSFO experiment had ended. The remaining cases yielded the results shown in Table 5, based on the generous verification criterion used in the earlier study (Sanders 1987) and in the checklist evaluation above; i.e., the timing and duration of the forecast deepening was not required to conform to the analysis evolution, provided forecast and analysis clearly referred to the same storm.

The checklist and the AVN forecasts produced nearly the same CSI, but the NGM scored distinctly higher. The sample size, however, is small, as attested by the difference in performance before and after 6 January 1988. At this point, the checklist was ahead, with a CSI of 0.80, followed by the NGM with 0.78 and the AVN at 0.55. After that time, the NGM scored 0.86, followed by the AVN at 0.67 with the checklist a poor third at 0.33. Additionally, during the past cold season, October 1988–March 1989, there were 16 events. The checklist scored 13 hits and produced 7 false alarms, yielding POD, FAR, and CSI of 0.81, 0.35, and 0.57, respectively, close to the values for the 1987/88 season.[1] No verification of the model forecasts has yet been undertaken. A longer period of record would be required for drawing reliable conclusions concerning the relative skill of models and checklist, but it appears that improvement of the checklist performance would most profitably focus on elimination of the relatively large number of false alarms. We feel that use of the checklist should be encouraged, along with careful examination of the model surface prognoses.

b. 12-h deepening

To get more information than is afforded by using the threshold of 1 bergeron, we examined 12-h predicted deepening for both models during the period of explosive deepening as shown by either of the analyses. These predictions were compared to the values from each of the analysis series.

---

[1] Note added in proof.

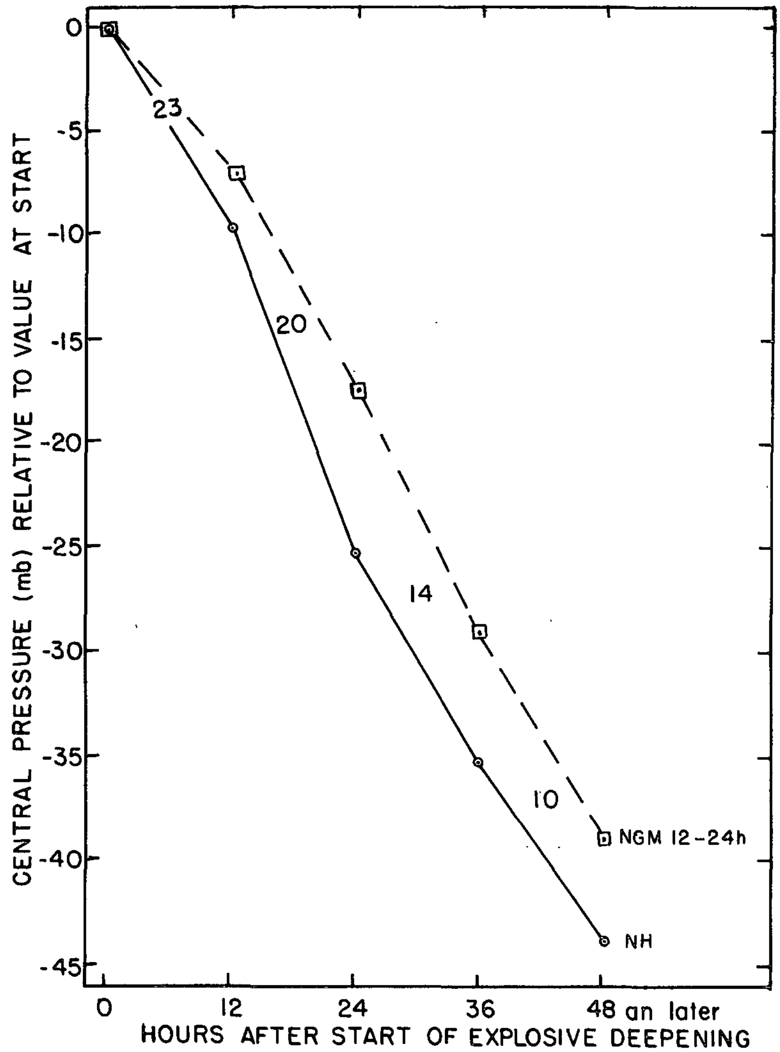

FIG. 3. (*Continued*)

First, however, the two analysis series were compared. While the 12-h deepening in the NH series was only 1.1 mb greater than that in the FH, there was a great scatter about the line of perfect agreement (Fig. 2a), and the correlation coefficient was not large. The two most outlier points occurred with storm No. 24, in which spectacular deepening occurred first in the NH analysis from 0000 to 1200 UTC 8 March and then 12 h later, more weakly, in the FH analysis, while the NH showed filling. In this case, no ship observations were available close to the center at the time of analyzed deepening, and the NH version must have relied on satellite imagery, in which a pronounced vortex developed. A later discussion (NOAA 1988) based on additional observations received by mail failed to mention strong winds prior to 0000 UTC 9 March. Therefore, what happened in the real atmosphere is in considerable doubt. Even if these two points are removed from the sample, however, the correlation coefficient rises only to 0.62, reflecting much disagreement between the two verification series.

This instance is an extreme example of a discrepancy between analyses that happened in a number of instances; most others were reasonably well documented by observations in support of the NH analyses. Note in Fig. 2b that, on the average, the NH center was about 6 mb deeper than the FH center 24 hours after the onset of explosive deepening. Beyond this time, the NH deepening tended to ease while the FH closed the gap, so that at 48 hours the center was only about 2 mb deeper in the NH than in the FH.

For comparison of the forecasts with results derived from earlier models (Sanders and Gyakum 1980; Sanders 1986, 1987) we first looked at the forecast seg-

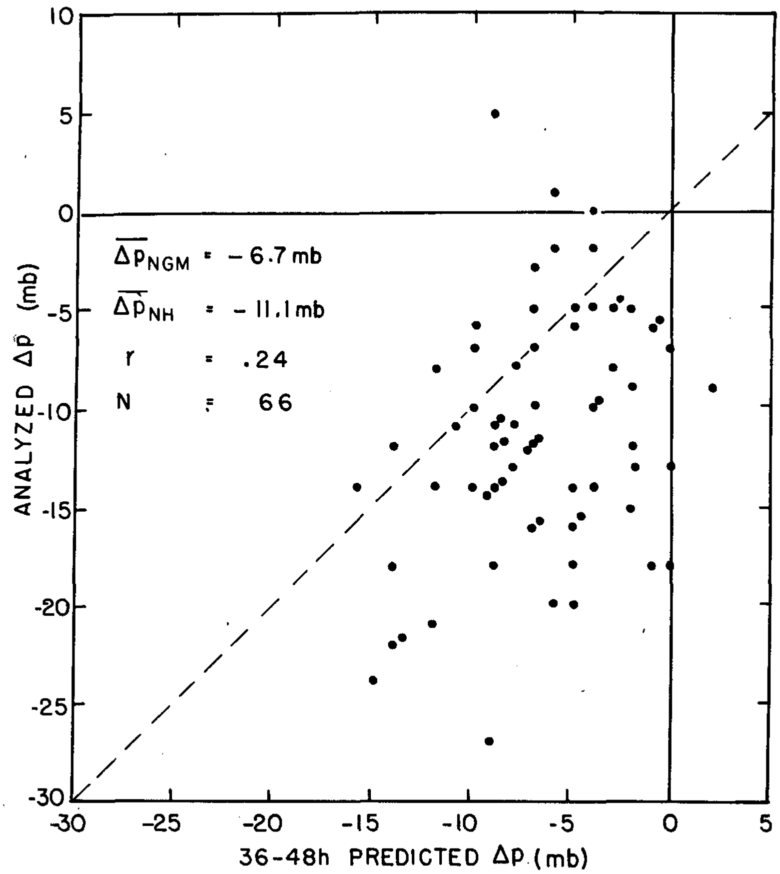

FIG. 4. As in Fig. 3, but for 36–48 h segments of NGM forecasts.

ments from 12 to 24 hours after initialization. Comparison of the NGM forecasts with the NH analyses is shown in Fig. 3. A substantial clustering of points shows close adherence to the "perfect forecast" line, but with a shortfall in predicted deepening of about 2 mb (Fig. 3a). There are also outlying points in both directions, so that the correlation coefficient is a modest 0.55. These results are comparable to those shown by Sanders (1987) for the NGM during the preceding cold season, but not as good as the results for the Atlantic and North American regions in that season. The average shortfall in deepening was about 1 mb more during this season than last; to judge from subjective inspection of the NH charts, this difference may be attributable to a tendency for analysts to draw slightly more intense centers this season than last. It should be noted that the correlation coefficient here is slightly higher than the coefficient between the two sets of verifying analyses.

The outliers are of two types. In a number of cases, the model simply did not catch the extremely rapid deepening shown in the analyses. Note that the strongest predicted deepening was 19 mb in 12 hours, while the analyses showed eight instances of deepening of 20 mb or more. These instances tended to occur near the beginning of intensification, as implied by the pressure values at the start of the extremely rapid falls. Examination of the charts for these instances shows some cases well documented by observations, but others, like the one discussed above, with a doubtfull analyzed deepening. (If the two data points corresponding to this case are removed from the sample, the correlation coefficient rises to 0.63, still slightly higher than the value for the comparison between the two verifying analyses.) Note from Fig. 3b that when the forecasts for 12 to 24 hours are stratified according to time after the onset of analyzed deepening, the model failed to replicate the rapid NH deepening in the first 24 hours. Later, however, like the FH analyses, the model tended to catch up. The outliers in Fig. 3a representing excessive forecast deepening were instances in which the NH center had already reached considerable depth, while the NGM center was still on the way down.

When these same NGM forecasts are compared with the deepening in the FH analyses (not shown), the average shortfall in the forecasts was only 0.6 mb and the correlation coefficient between forecast and analyzed

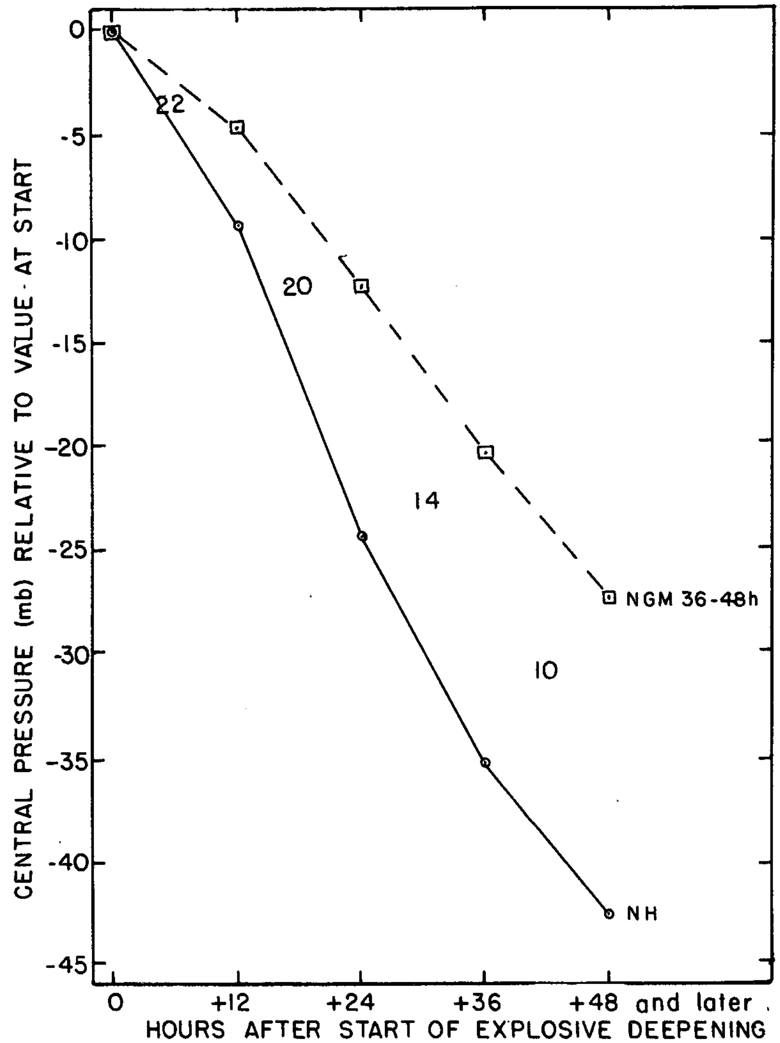

FIG. 4. (*Continued*)

deepening was 0.61. The difference between the forecast and the analyzed deepening varied little throughout the cyclogenetic episode, on the average. As in the discussion of the CSI scores, we feel that this improvement is spurious, reflecting largely the failure of the FH analyses to capture the early rapid deepening seen in the NH analyses, however doubtful the latter might be on occasion.

Examination of the segments of the NGM forecasts for the range from 36 to 48 hours shows a marked degradation when verified by the NH analyses. The scatter of points in Fig. 4a is very broad and the correlation coefficient is low. The average predicted deepening underestimated the analyzed deepening at all stages in the development (Fig. 4b), the average discrepancy being 4.4 mb while the largest was 7.3 mb between 12 and 24 hours.

When the 12–24 h segments of the AVN forecast deepening are compared to the NH analyzed values (Fig. 5a), we see a performance slightly worse than that of the NGM; i.e., the shortfall in predicted deepening, 2.6 mb, is greater in this case, and the correlation coefficient is somewhat lower. Once again, the model failed to catch the full strength of the early deepening (Fig. 5b), but this time there was little tendency of the forecast to catch up, since the average shortfall remained nearly constant after the first 24 hours.

Examination of the 36–48 h segments of the AVN forecasts, however, shows in Fig. 6 a somewhat better performance than the corresponding NGM predictions. The underprediction by 3.6 mb is notably less than that of the NGM and little different from the AVN result at shorter range, while the correlation coefficient is not quite so low as for the NGM. This improvement

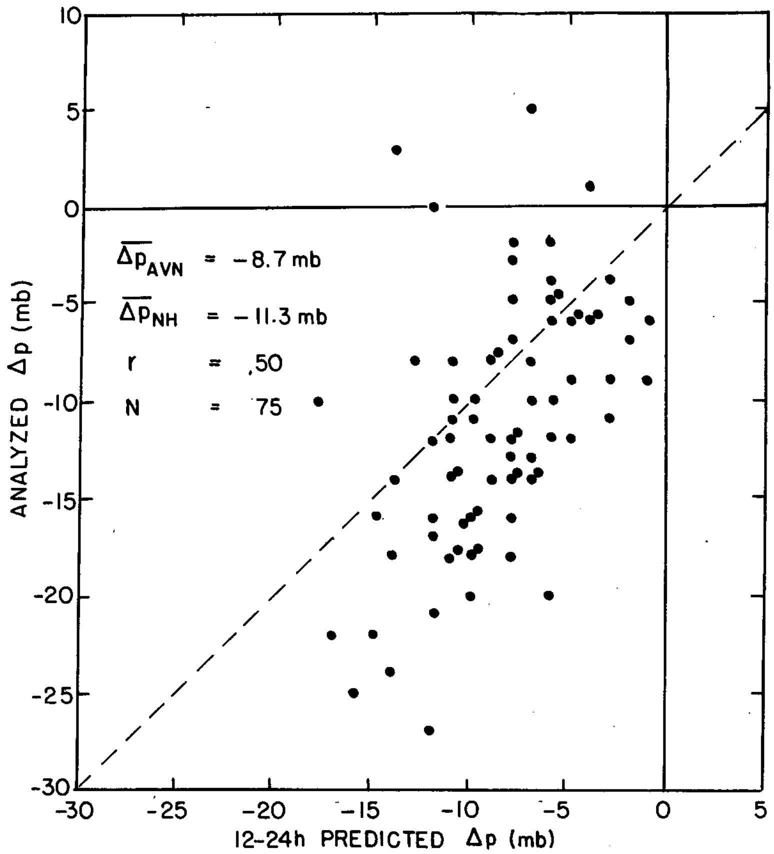

FIG. 5. As in Fig. 3, but for 12–24 h segments of AVN forecasts.

(or reduction of deterioration), is consistent with the slightly better CSI performance of the AVN in the 36-h forecasts shown in Table 1. In both models there is plenty of room for improvement at this range.

Verification of the 12–24 h AVN forecasts with reference to the FH analyses rather than the NH analyses (not shown) indicates a substantially better performance, with an average underestimate of the analyzed deepening of 1.4 mb and a correlation coefficient of 0.60. As with the NGM comparison, however, this improvement is felt to be spurious.

4. Conclusions and further discussion

Skill in prediction of explosive cyclogenesis over the western North Atlantic Ocean by the NMC dynamical models appears to have increased during the 1987/88 cold season, relative to the preceding year. For both the NGM and the AVN model runs, the probability of detection of the event was evidently greater and the false alarm rate was evidently reduced. As examples of the most recent performance, for the NGM forecasts in the range from zero to 24 hours, nearly three out of each four events was detected while one in six forecasts was a false alarm; for the AVN forecasts in the range from 36 to 48 hours, somewhat fewer than half the events were detected and nearly one in three of the forecasts was a false alarm. These results were obtained by reference to a series of manually analyzed surface maps for verification. When an automated series was used for verification, forecast performance was somewhat better.

A simple checklist used by forecasters at WSFO Boston, when verified by the same criterion and for the same set of storms, performed approximately as well as the AVN forecasts but not as well as the NGM forecasts. The sample size was extremely small, however.

Model predictions of 12-h deepening were compared to deepening in the analyses. For the NGM model in the range from 12 to 24 hours, the predicted deepening was, on the average, 2 mb less than the manually analyzed value with a correlation of 0.55. For the AVN run in the same range, the shortfall in predicted deepening was 2.6 mb and the correlation was 0.50. At the longest range for which comparison was possible, 36 to 48 hours, NGM performance was substantially

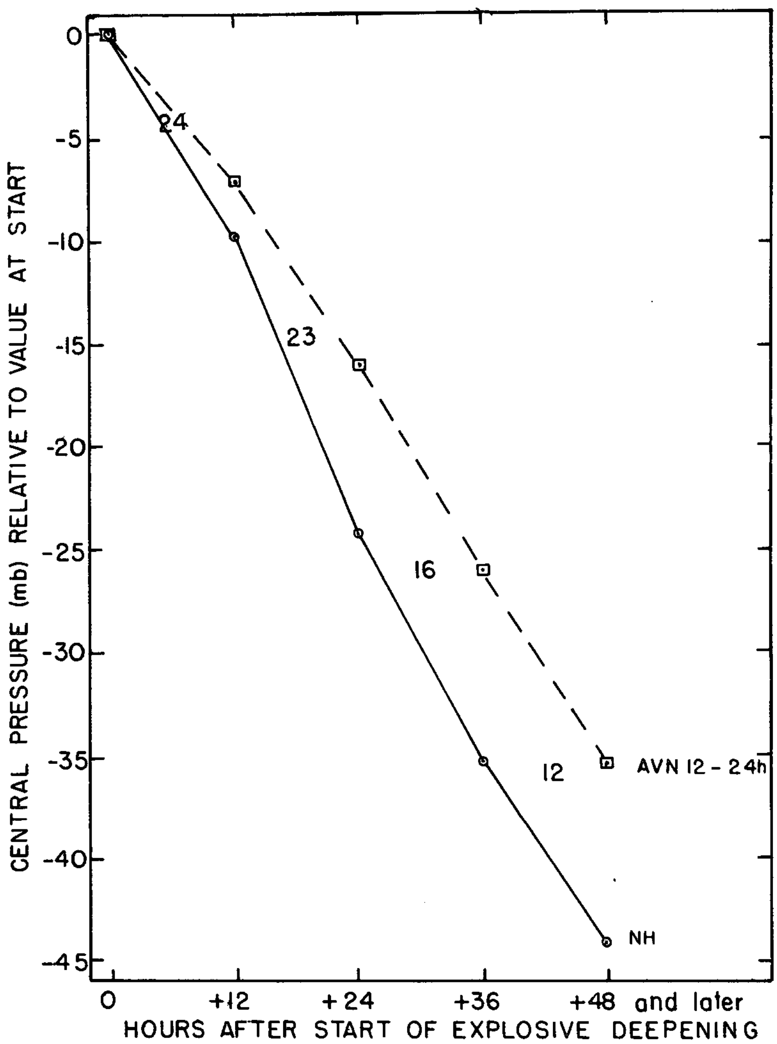

FIG. 5. (*Continued*)

worse: a shortfall of 4.4 mb and a correlation of 0.24. Deterioration in the AVN forecasts was not quite as great, with corresponding values of 3.6 mb and 0.31.

The correlations for the shorter range were limited by the failure of the models to capture accurately the initial onset of explosive deepening in the manual analyses. The automated analyses also failed to show the full intensity of this initial onset, although they tended to catch up when the cyclone had completed deepening in the manual analyses. Comparison of the 12-h deepening in the two sets of analyses showed a mean shortfall of 1.1 mb in the automated product but a correlation of only 0.49.

Thus, there is some doubt about what actually happened in the real atmosphere, but the evidence tended to support the manual analyses. In any case, since the correlation between analysis and the forecast in the range from 12 to 24 hours is as good as the correlation between two sets of analyses, it appears that the analysis will have to be more certain before it will be possible to demonstrate substantial further increase in forecast skill.

A consistent tendency is for relatively poor short-term forecasts from both models in predicting the early part of the explosive intensification. This is seen in both the CSI results for the 1-bergeron threshold in Tables 2 and 3, and in the verification of 12-h predicted deepening seen in Figs. 3b and 5b. It is evident from Fig. 1 that most of the starting points for explosive intensification in the NH analyses were over relatively warm water, along the edge of the Gulf Stream or in the Sargasso Sea. The models may still not represent adequately the way in which heat and moisture from the sea surface is transferred into the atmosphere, dis-

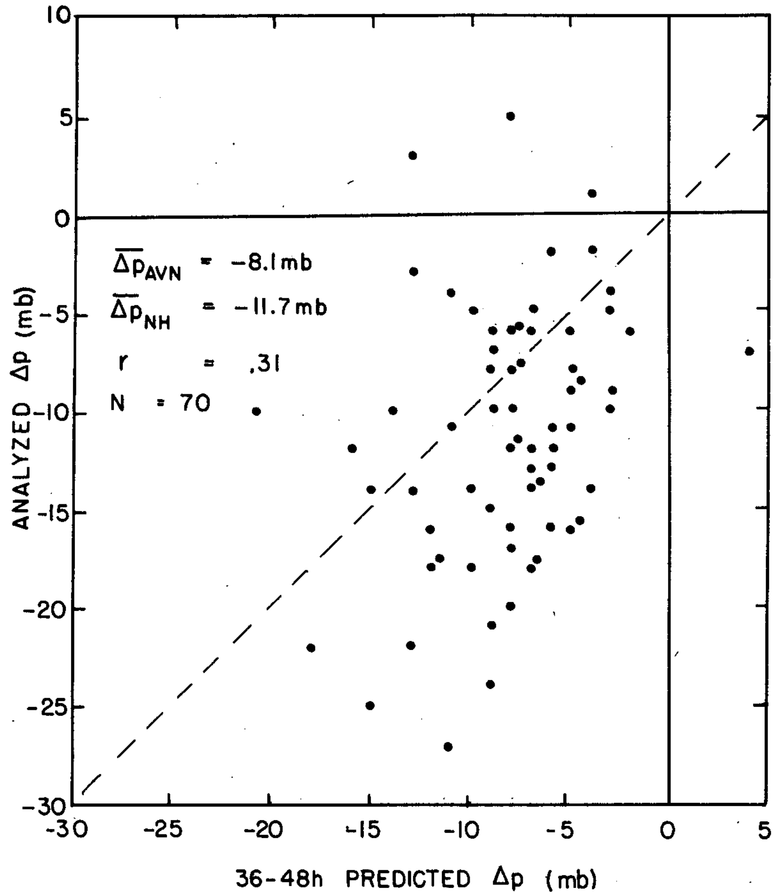

FIG. 6. As in Fig. 3, but for 36–48 h segments of AVN forecasts.

tributed aloft and condensed. These processes may be particularly important at this early time, if the interaction with the oncoming upper-level disturbance has not yet become strong.

We did not undertake detailed diagnosis of individual cases. From the NGM initial analyses, however, we identified the track of the 500-mb vorticity maximum ultimately associated with the deep surface low through the period of cyclogenesis for each case. At the start this maximum was at a distance of slightly more than 1000 km from the surface center. Given the typical scale of mobile upper-level systems, this suggests that the surface system was beneath just the outer fringe of the associated upper-level cyclonic vorticity–advection field at this time. In fact, in about half the cases, the early development of the storm showed an association with an additional weak upper-level vorticity maximum in the southeastern states, south of the major band of baroclinic westerlies. This weak maximum tended to lose its identity as the more powerful northern one approached and surface cyclogenesis proceeded. In contrast, 24 hours later an unambiguous vorticity maximum lay at an average distance of about 600 km from the surface center. Interaction at this time was likely strong, and the models performed better at short range.

From recent research on maritime cyclogenesis, the complex role played by surface transfers of sensible and latent heat is beginning to emerge. In sensitivity tests of a channel model (Nuss and Anthes 1987), the direct effect of the sensible fluxes is to weaken the baroclinicity in the environment of the developing cyclone, and to weaken the thermal wave that supports the storm. Hence, in this sense, the fluxes oppose cyclogenesis. Chen et al. (1983), however, find surface fluxes important at short range through their influence on latent heating. Moreover, circumstantial evidence for a direct cyclogenetic effect of fluxes has been presented recently by Davis and Emanuel (1988), so the issue is still unresolved.

On the other hand, there is general agreement between observational studies (Rogers and Bosart 1986; Reed and Albright 1986; Reed and Blier 1986a, 1986b; Wash et al. 1988) that small or weakening stability (or

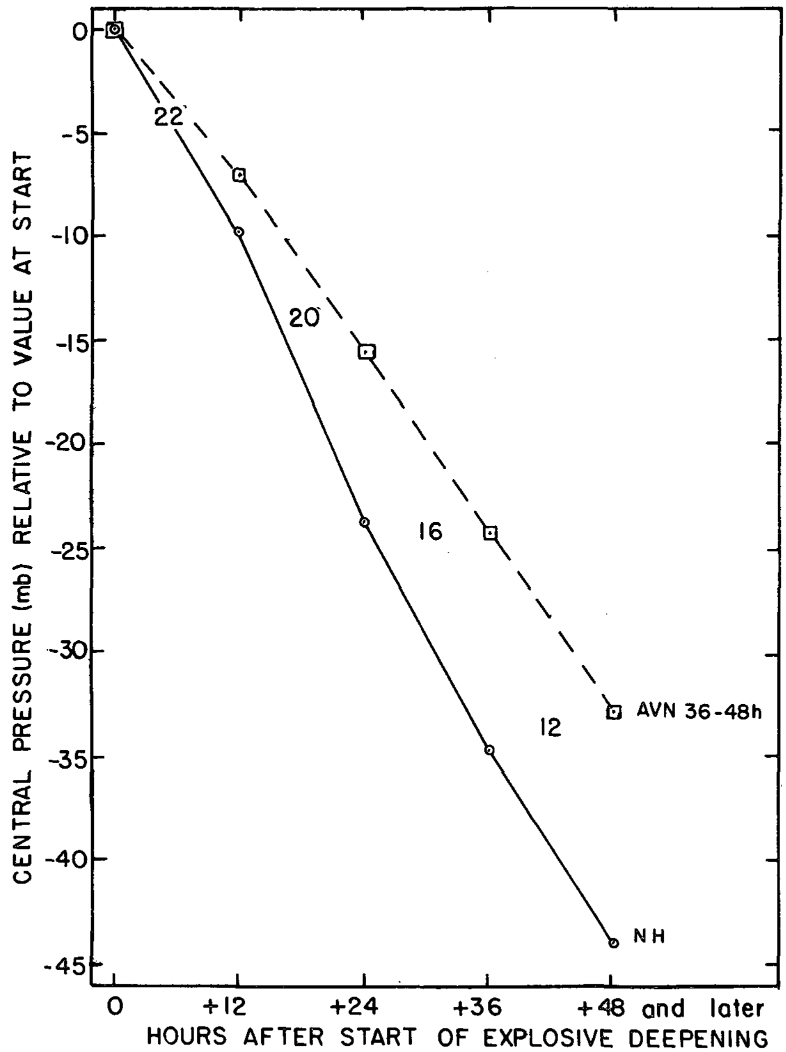

FIG. 6. (*Continued*)

conditional instability) characterizes at least the lower troposphere during rapid cyclogenesis, and Nuss and Anthes (1987) find it a necessary ingredient of strong cyclogenesis. Presumably, this small stability over the ocean is the result of conditioning of the lower troposphere by the sea surface prior to the onset of development.

Moreover, the importance of latent heating in driving the updraft, and thus the cyclogenesis, is seen in diagnosis of observations (Chang et al. 1987; Sinclair and Elsberry 1986) and in model simulations (Kuo and Reed 1988; Mullen and Baumhefner 1988; Orlanski and Katzfey 1987). The effect of large-scale latent heating, of course, can be alternatively expressed as a reduction of the effective static stability.

The surface fluxes of heat, both sensible and latent, appear to act at least indirectly, by producing a moist air-mass structure with small effective stability which then becomes involved in cyclogenesis. Since the operational models contain this structure in their initial analysis, they may be able to predict explosive deepening quite well without representing the in situ surface fluxes very well.

Along the east coasts of continents, however, surface fluxes apparently act in an additional direct way by enhancing low-level baroclinicity, with indirect effects through condensation encouraging cyclogenesis (Uccellini et al. 1987; Chen and Dell'Osso 1987). The difficulty of operational models in predicting the early stages of cyclogenesis may be due to inadequacy in representing the consequences of this process.

Alternatively, it may be that latent heat release in the early stages of cyclogenesis (the importance of which is generally acclaimed) occurs mainly on the cumulus convective scale without grid-scale saturation. Perhaps the convective parameterization is not correctly representing the situation. Such a scenario was suggested by Tracton (1973) for a number of poorly predicted cases of cyclogenesis over land.

Whatever the problem may be, at ranges beyond the first 24 to 36 hours, additional sources of error evidently affected the model performance sufficiently that predictions were no worse for the early stage of cyclogenesis than for later stages (Tables 2 and 3). These suggestions are only preliminary because there was large case-to-case variability, the analyses for some individual cases violated each of our speculative assertions, and there was uncertainty in determining what actually happened in the real atmosphere.

Acknowledgments. The authors are grateful to the Center for Meteorology and Physical Oceanography, Massachusetts Institute of Technology, for the provision of data and facilities, and to Isabelle Kole for the drafting of figures. We thank an anonymous reviewer for suggesting reasons for recent improvement in the forecasts. The work was supported in part by the Office of Naval Research under Contract N00014-85-C-0785.

REFERENCES

Auciello, E. P., and F. Sanders, 1986: An operational checklist for the forecasting of explosive cyclogenesis in the North Atlantic Ocean. *Eastern Region Technical Attachment,* No. 86-7. National Weather Service.

——, and ——, 1987: An operational checklist for the forecasting of explosive cyclogenesis in the North Atlantic Ocean. *Proc. Second Workshop on Operational Meteorology,* Canadian Meteorological and Oceanographic Society, Halifax, Nova Scotia.

Chang, C.-B., D. J. Perkey and W.-D. Chen, 1987: Observed dynamic structure of an intense oceanic cyclone. *Mon. Wea. Rev.,* **115,** 1127–1139.

Chen, S.-J., and L. Dell'Osso, 1987: A numerical case study of East Asian coastal cyclogenesis. *Mon. Wea. Rev.,* **115,** 477–487.

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

Davis, C. A., and K. A. Emanuel, 1988: Observational evidence for the effect of surface heat fluxes on maritime cyclogenesis. *Mon. Wea. Rev.,* **116,** 2649–2659.

Donaldson, R. J., Jr., 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, Amer. Meteor. Soc., 321–325.

Hadlock, R., and C. W. Kreitzberg, 1988: The Experiment on Rapidly Intensifying Cyclones over the Atlantic (ERICA) field study: Objectives and plans. *Bull. Amer. Meteor. Soc.,* **69,** 1309–1320.

Kuo, Y.-H., and R. J. Reed, 1988: Numerical simulation of an explosively deepening cyclone in the Eastern Pacific. *Mon. Wea. Rev.,* **116,** 2081–2105.

Lange, O., R. Fett, R. Landis, G. Wells, G. Rasch, E. Auciello, M. Horita and J. Jannuzza, 1986: Summary of a workshop/seminar on oceanic storms 9–13 September 1985, Seattle, Washington. *Bull. Amer. Meteor. Soc.,* **67,** 1272–1277.

Mullen, S. L., and D. B. Baumhefner, 1988: The sensitivity of numerical simulations of explosive oceanic cyclogenesis to changes in physical parameterizations. *Mon. Wea. Rev.,* **116,** 2289–2329.

NOAA, 1988: Marine weather review - North Atlantic weather log. *Mariners Wea. Log,* **32,** 45.

Nuss, W. A., and R. A. Anthes, 1987: A numerical investigation of low-level processes in rapid cyclogenesis. *Mon. Wea. Rev.,* **115,** 2728–2743.

Orlanski, I., and J. J. Katzfey, 1987: Sensitivity of model simulations for a coastal cyclone. *Mon. Wea. Rev.,* **115,** 2792–2821.

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

——, and W. Blier, 1986a: A case study of comma cloud development in the Eastern Pacific. *Mon. Wea. Rev.,* **114,** 1681–1695.

——, and ——, 1986b: A further study of comma cloud development in the Eastern Pacific. *Mon. Wea. Rev.,* **114,** 1696–1708.

Rogers, E., and L. F. Bosart, 1986: An investigation of explosively deepening oceanic cyclones. *Mon. Wea. Rev.,* **114,** 702–718.

Sanders, F., 1986: Explosive cyclogenesis over the west central North Atlantic Ocean, 1981–1984. Part I: Composite structure and mean behavior. *Mon. Wea. Rev.,* **114,** 1781–1794.

——, 1987: Skill of NMC operational models in prediction of explosive cyclogenesis. *Wea. and Forecasting,* **2,** 322–336.

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

Sinclair, M. R., and R. L. Elsberry, 1986: A diagnostic study of baroclinic disturbances in polar air streams. *Mon. Wea. Rev.,* **114,** 1957–1983.

Tracton, M. S., 1973: The role of cumulus convection in the development of extratropical systems. *Mon. Wea. Rev.,* **101,** 573–593.

Uccellini, L. W., R. A. Petersen, K. F. Brill, P. J. Kocin and J. J. Tucillo, 1987: Synergistic interactions between an upper-level jet streak and diabatic processes that influence the development of a low-level jet and a secondary cyclone. *Mon. Wea. Rev.,* **115,** 2227–2261.

Wash, C. H., J. E. Peak, W. E. Calland and W. A. Cook, 1988: Diagnostic study of explosive cyclogenesis during FGGE. *Mon. Wea. Rev.,* **116,** 431–451.