AD-A007 846

AUTOMATED LONG RANGE (ANALOG) FORECAST MODEL

Donald R. McConathy, Jr., et al

Fleet Numerical Weather Central
Monterey, California

December 1974

UNCLASSIFIED
Security Classification

AD-A007846

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
FLEET NUMERICAL WEATHER CENTRAL

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE
AUTOMATED LONG RANGE (ANALOG) FORECAST MODEL

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
TECHNICAL NOTE 74-1, DECEMBER 1974

5. AUTHOR(S) (First name, middle initial, last name)
DONALD R. MCCONATHY JR., LT, U.S. NAVY and
CARL D. THORMEYER, LT, U.S. NAVY

6. REPORT DATE
DECEMBER 1974

7a. TOTAL NO. OF PAGES
75

7b. NO. OF REFS
18

8a. CONTRACT OR GRANT NO.

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

Naval forces have two principle requirements for long-range weather forecasts: (1) to protect ships in transit, and (2) to aid in operational planning. Since the state-of-the-art precludes adequate dynamic model forecasts beyond a few days, a synoptic climatology approach has been developed for longer ranges. Additional products include sea condition probabilities, a list of Best Analog Dates, and Terminal Probability Forecasts for selected Naval Installations. Four years of statistical verification have consistently demonstrated skill over persistence beyond two days, with skill over climatology out to six to seven days prior to 1973 and out to 8 days or more since then.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED
Security Classification

FLEET NUMERICAL WEATHFR CENTRAL's AUTOMATED LONG RANGE (ANALOG) FORECAST MODEL

By

Donald R. McConathy Jr., LT, U. S. Navy

and

Carl D. Thormeyer, LT, U. S. Navy

Technical Note No. 74-1

December 1974

Fleet Numerical Weather Central
Monterey, California

ia

This Technical Note is intended to replace the FNWC Technical Note No. 49 of September 1969. It includes a summary of all development work done to date at FNWC on the Long Range Forecast Model. The paper has been written in such a way as to be useful in both understanding the mechanics of the model and in the interpretation of the products. In this respect the publication is also a users guide to the analog products.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table

ABSTRACT

Naval forces have two principle requirements for long-range weather forecasts: (1) to protect ships in transit, and (2) to aid in operational planning. Since the state-of-the-art precludes adequate dynamic model forecasts beyond a few days, a synoptic climatology approach has been developed for longer ranges. To produce a hemispheric long-range forecast, a CDC 6500 computer searches Fleet Numerical Weather Central's 28-year data base at two levels (500 and 1000mb) to find synoptic situations most closely resembling current weather patterns. The 40 best matches (or analogs) are then subjected to a 10-day history check to eliminate criss-crossing or unseasonal historical weather sequences. At least 20 analogs remain, and the ten best of these are averaged point by point over the Northern Hemisphere to produce a composite initial analog match. This average composite is then projected out ten days yielding hemispheric forecasts of large-scale features and trends (e.g., blocks, long-waves, stormtracks). Additional products include sea condition probabilities, a list of Best Analog Dates, and Terminal Probability Forecasts for selected Naval Installations. Four years of statistical verification have consistently demonstrated skill over persistence beyond two days, with skill over climatology out to six to seven days prior to 1973 and out to 8 days or more since then.

1.0 INTRODUCTION

Modern military operations require long-range planning. Environmental conditions are a major consideration in such planning. Historically, these conditions were described through the use of climatological atlases. Developments in recent years have required more detailed long-range forecasts than climatology can provide.

Most efforts in long-range forecasting have concentrated on dynamical or climatological/statistical approaches. Dynamic forecast models, including the Fleet Numerical Weather Central baroclinic multi-layer primitive equation models, do not show skill beyond a few days. The large data voids over the ocean areas often preclude a sufficiently accurate specification of the initial analysis. A successful dynamic solution of the initial value problem which currently limits short range forecasts appears to be several years into the future. In the interpretation of large-scale flow patterns, statistical approaches to prediction of weather parameters (e.g., conditional probabilities) have been useful, especially in the very long-range forecast periods of 15 days or more. Neither the climatological/statistical nor dynamic approaches provide for the critical 3 to 15 day forecast period.

Elliot (1) described a weather typing method used in forecasting for military operations during World War II. Several years ago Fleet Numerical Weather Central (FNWC) turned to a synoptic climatology (i.e., historical analog) approach to forecasts beyond a few days. Recent efforts by FNWC include developments by Wolff and Woodworth (2), Wolff and Thormeyer (3) and Thormeyer (4). Considerable experimentation has been carried out by private and government meteorologists in the United States and abroad [Hartranft, Restivo and Sabin (5), Elliot (6), White and Palson (7), Hofmann (8), Lund (9), Lorenz (10) and Augulis (11)].

2.0 PHILOSOPHY BEHIND SYNOPTIC CLIMATOLOGY FORECAST TECHNIQUES

Use of synoptic climatology methods in long-range forecasting is based upon the following assumption: if weather sequences in past history can be found which closely match the current sequence, the same dynamic interactions (even though these interactions are not fully understood) which took place then are apparently also taking place at the present time. Further, it is assumed that since the end results were essentially the same, the series will persist long enough to serve as a medium and long-range forecasting tool. The validity of this assumption depends on many factors - quality of initial match, specific synoptic pattern in question (common or unique), amount of meridional flow, etc. If analog forecasts can be shown to be more skillful than random chance, climatology or persistence, then there must exist a dynamic relationship within the analog selection technique to whatever is predicted. This philosophy is applied subjectively by experienced forecasters through the use of "recall" (using their mental catalog of analogous patterns) to make current forecasts.

FNWC has developed and tested several analog models. The most successful model, developed by Hanna (15), has been utilized operationally since July, 1970. The present effort differs from previous methods in the following ways:

(1) The selection of analogs is totally automated; no human judgement enters into the process.

(2) The availability of advanced high-speed digital computers allows a rapid and comprehensive search of a large data base.

(3) Use of two levels considers the thermal structure, thus sharpening the selection process.

(4) The synoptic data base is more extensive than any previously used and encompasses possible cyclic effects through a double sunspot (22-year) cycle.

(5) Use of a pattern separation technique [Holl (12), Thormeyer (13)] improves the selection procedure by separating large and small-scale disturbances.

3.0 THE HISTORICAL DATA BANK

The analog technique requires finding a nearly perfect match. Since an ideal homogeneous data bank of several hundred years' duration is unattainable, FNWC endeavored to acquire a maximum data base. The result is a set of Northern Hemisphere surface and 500 mb analyses from 1946 to the present. The earlier years of this data base were acquired from several different sources (14) while FNWC analyses have been used since 1962. All of the non-FNWC data has been re-analyzed in a manner approximating the current objective numerical analysis scheme. Although this does not achieve absolute homogeneity, it insures continuity from map to map and improves the vertical consistency of analyses from earlier years. This data base is being continually updated by FNWC's Climatology Department.

From this start, 1000mb contour and 500 - 1000 mb thickness fields were generated. FNWC's pattern separation technique (12) was then applied to acquire large and small scale disturbance fields at 500 and 1000 mb. This technique, applicable to any continuous field, objectively divides the total flow into large (SL) and small(SD) scale disturbance fields and a thermally-driven planetary vortex (SV). The SV plus the SL yield the residual field (SR) which, at 500 mb, depicts the long wave pattern. Since the SV is a measure of the zonal wind strength, the large and small scale features, when isolated, are essentially undiluted by the zonal component of the total flow.

4.0 DEVELOPMENT OF THE MODEL

4.1 The Single Analog Approach (Past)

From its inception in March 1969 until July 1970, the operational forecast model was based on the concept of isolating one best historical analog although a near-perfect match could not always be found. Throughout the hemisphere, the best correlations of large-scale features with those of the current chart averaged from 0.70 to 0.75 and rarely exceeded 0.85, except at 1000 mb in the summer when correlations were routinely in the 0.90 to 0.95 range. With such poor initial specification, it was not surprising that the single analog model showed only occasional skill over climatology and persistence, regardless of the selection method used.

4.2 The Composite Analog Approach (Present)

It was determined that when ten hemispheric fields are averaged together gridpoint for gridpoint, thus forming a composite analog, the correlation of this new analog is always higher than that of any of the ten individual components. Further, the composite looked more like the current chart than did any of the individual charts. Composite correlations have routinely been 0.10 to 0.20 greater than the highest individual correlation included in the composite. Initial large scale composite correlations in winter and summer reach 0.95 or greater in one or more sectors (see section 5.2.1) and often exceed 0.90 across the hemisphere. However, in the transition seasons they drop to as low as 0.75.

In developing this technique, Hanna (15) points out that if one considers the analog approach as a crude application of Time-Series Analysis and Prediction, powerful statistical techniques can be applied. If the entire past of a dynamical system is known, then nothing is gained by knowing the dynamics as well. Strictly speaking, a direct application of the time-series approach would be to relate today's weather pattern with the entire available past at every point in the hemisphere, thus producing prediction coefficients. Although the multiple time-series problem has essentially been solved theoretically, a tremendous effort is required to apply it directly to meteorological prediction. The evolution of the composite analog technique attempts to approximate a strict application of the multiple time-series prediction.

There are three major steps involved in the computation of the composite long-range forecast: (1) selection of the initial analog candidates making up the composite, (2) a history check to eliminate unseasonal and non-parallel patterns, and (3) construction of the composite itself. Figure 1 is a flow diagram of the current composite model.

4.2.1 The Initial Date Selection. All years in the data base are searched ± 30 days from the current calendar date (ignoring out-of-season months) for the following four parameters:

SL 500	Large-scale disturbance field at 500 mb
SL 1000	Large-scale disturbance field at 1000 mb
D 500	500 mb height field
D 1000	1000 mb height field

Hemispheric correlations between the historical charts and current charts are computed using all grid points north of 20N (approximately 1500) according to the standard equation

$$r = \frac{\Sigma(x - \bar{x}) \quad (y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 \; \Sigma(y - \bar{y})^2}} \quad , \qquad (1)$$

where r is the correlation coefficient, and x and y are corresponding gridpoints on the current and historical data fields being correlated. Note that since the 500 and 1000 mb height fields include both large and small scale disturbances, the large-scale features are weighted about three to one over the small-scale in the selection procedure. Some 1680 individual correlations for each of the four parameters are available at this point; these are normalized (Equation 2) and averaged over the four parameters for each historical date.

$$N = \frac{r - \bar{r}}{\sigma} \quad (2)$$

In this equation, r is the correlation value of a specific chart with the current data, $\bar{r}$ is the mean of all 1680 correlations for that parameter, and σ is the standard deviation of the 1680 correlations for that parameter. The principal function of normalizing is to place the correlations of all four parameters on an equal basis. The normalized values for all four parameters from the same historical date are averaged together, yielding one correlation (N) for each date. These 1680 normalized and averaged correlations are then ranked, with the best 40 selected as candidates. A typical range in the value of N is from about 2.30 for the highest candidate to about 1.50 for the 40th. A perfect match is near 4.00. Other selection techniques that have been tried in the past are discussed briefly in Appendix A.

4.2.2 The History Check. In an attempt to eliminate unseasonal and non-parallel weather patterns (which have evolved differently yet correlate well with current charts), a gross 10-day history check is made. The history of each of the 40 candidates is correlated with the history of the base date (D_o-1 to D_o-5, and D_o-6 to D_o-10). The history check insists that the D_o-1 to D_o-5 average be no less than the D_o value minus 1.25, and that the D_o-6 to D_o-10 average be no less than the D_o value minus 2.50. If these criteria are met, the candidate remains eligible for inclusion in the composite. Normally, 25 to 30 of the original 40 dates survive the history check. To ensure enough dates for the composite formation, at least 20 dates must remain after the history check. If necessary, the 1.25/2.50 limits are expended (at increments of 0.125/0.250) until the 20 dates are obtained. Recent comparison of the history check against a no-history check indicates the history check provides little or no additional skill.

4.2.3 Final Selection and Composite Formation. From the dates which pass the history check the ten with the highest D_O normalized averages are selected for the composite. These 10 dates are then averaged together for each desired parameter thus generating a standard FNWC Northern Hemispheric 63 x 63 field (16). This field becomes the initial match. Daily forecasts are produced by adding 24 hours to each of the historical dates making up the composite, and averaging these new dates for each additional day's forecast desired. Forecasts are generated from zero to 18 days at one-day intervals; however, routine dissemination is limited to 10-12 days due to a lack of skill beyond that time frame.

The relative success of the composite approach can be explained physically by the fact that most individual analogs show phase errors in one section of the globe or another. Further, on a given day, the bulk of the sector phase errors (among the 40 analog candidates) tend to be in the same portion of the globe. By averaging the ten strongest candidates, the well matched areas tend to be enhanced while the gross phase errors tend to be toned down through smoothing. This appears to be due to the fact that the analog candidates making up the composite are not random, but rather share some loose relationship of weather sequences.

4.3 The Current Operational Model. The operational analog forecast model, based on the composite approach described above, has been in routine operation since 1 July 1970. It was run weekly until October 1971 and twice weekly since. The initial match was based on the 00Z synoptic time data, but to obtain a better initial field the program was switched to the updated 12Z synoptic data in November 1972. This change allows the maximum data input to the initial analysis and eliminates any effect of comparing 12Z climatology fields with a 00Z synoptic field.

5.0 PRODUCTS AND USAGE

Due to the known characteristics of analogs (Appendix B), forecast products are designed to yield primarily large scale information. The composite concept currently in operation sacrifices detail in order to minimize undesirable phase-errors. Although extremes cannot be portrayed, large-scale information is not seriously affected by the smoothing procedure.

5.1 Hemispheric Products

5.1.1 Forecast Fields. Several charts are routinely produced at daily increments out to 12 days.

The large-scale 500 mb analog forecast, SL 5 AL, depicts long-wave and block positions with associated relative intensities. Examination of the entire sequence of forecast charts provides information on the anticipated movement, persistence or intensity changes of these larger-scale systems. Since the detail in the SL 5 AL's is not as sharp as in a regular SL 500 chart, stormtracks can be implied from gradient, curvature, and direction of flow. SL 5 AL charts are plotted with a contour interval of 20 meters (vice 30 meters on the synoptic SL 500) in order to accomodate the loss of detail caused by the averaging process. Analog forecasts of the residual field (SR 500) and D-value field (D-500) are not normally used due to the inherent over-smoothing.

The surface pressure analog forecast, PS AL, show the most likely positions for surface highs and lows on a given day in the forecast period. They can be thought of as probability charts - the stronger a high or low is on a given forecast chart, the more probable it is that a system of the same type will be located near that position. The predicted strength of, and changes in, semi-permanent systems (such as the Siberian High and Aleutian Low in winter, and the Eastern Pacific/Atlantic Highs and Asian Thermal Low in summer) can be particularly useful. Examination of the entire sequence of forecast charts yields information on expected principle stormtracks including Western Pacific tropical cyclones. PS AL charts are plotted with a three-millibar contour interval to compensate for the loss of detail.

5.1.2 Product Limitations. The lack of exactness in analog forecasting, when considered with the characteristics of analogous weather patterns (Appendix B), limits the overall skill of the forecasts. Unlike dynamic models, the latest forecast series may not be as good in a given geographical area as was the previous one, in which case the older forecast should be used. Further, the smoothing inherent in the model limits the forecast usefulness to large-scale features, and precludes any possible accurate depiction of intense small-scale features or rapidly moving storms. Even with the large-scale features, the nature of the composite model enables it to predict only about half the intensity of the overall change that actually occurs making it more useful qualitatively than quantitatively. Phase errors also occur, which means ± two days from the forecasted date should be allowed in verifying predicted large-scale changes.

5.1.3 Examples of Analog Products. Two sets of forecast charts, summer and winter, are included as Figures 7 - 25. Table 1 gives comparative statistical information on the initial match and verification scores for each set. One shortcoming of the analog products is that the relative magnitude of most systems is from 10 to 30 percent weaker than the verification chart. This is caused by the smoothing technique used in creating the composite and is especially prevalent on the surface (PS) charts where small scale systems are incorporated into a "mean" composite field. When interpreting these charts, the geographical location and relative strength of the major analog features is more important than trying to make a perfect correlation of individual systems.

Initial Match - Composite Analog compared to Base Date Chart(D_o)

Parameter	Hemi	U.S.	Pearl	Guam	Rota
00Z 14 JUN 71					
D 1000	.8164	.7237	.7323	.7196	.8869
SL 500	.8179	.4781	.9007	.9008	.8614
00Z 8 NOV 71					
D 1000	.6427	.4629	.6942	.7816	.5479
SL 500	.8029	.7968	.7599	.8657	.8235

Analog Verification - Composite Analog Compared to Verifying Chart during Forecast Period

Date	Initial Match	Day 1 - 5	Day 6 - 10	Day 11 - 15	Day 1 - 10	Day 1 - 15
14 JUN 71	.734	508	.376	.481	.442	.455
08 NOV 71	.663	.583	.567	.377	.575	.509

Table 1 - Comparative Initial Match and Verification Statistics for Composite Examples

The following is a summary of the major features (good and bad) for the composite forecast charts.

A. SL 500 - Summer Series

00Z 14 JUN 71 - Initial Match (Figures 6-7). This chart is very good in all areas except over the U.S. where the Low off British Columbia is not shown. Note the very close match of the Low on the Kamchatka Peninsula.

00Z 18 JUN 71 - 4 DAY FCST (Figures 8-9). Major systems are slow indicating a possible phase error of 1-2 days. Relative location and intensity of most systems is good, with the over development of the Low in the Eastern Atlantic and the orientation of the blocking ridge over Canada being the major discrepancies.

00Z 21 JUN 71 - 7 DAY FCST (Figures 10-11). The forecast did not show the reintensification of the East Pacific Low. It weakened the block over North America too much and maintained the Low in the Eastern Atlantic which actually moved northward (Note low verification score for the 6-10 day forecast period in Table 1).

00Z 25 JUN 71 - 11 DAY FCST (Figures 12-13) An extremely good match is seen over the Eastern U.S. and Atlantic. Intensification of the High over Alaska is not predicted.

B. PS - Summer Series

00Z 14 JUN 71 - Initial Match (Figures 14-15). The combination of different contour interval plus smoothing make this chart closer to an SR 1000 chart. With this in mind, all the major features are noted, especially the deep Low in the NW Pacific. Again relative strengths of systems are the important consideration.

Throughout this entire series, the SL 500 Low over Eastern Canada has remained as a nearly stationary system, acting as an anchor point for the rest of the hemisphere. Studying Figures 6-15 along with Table 1 will give some insight into how best to utilize the analog charts.

C. SL 500 - Winter Series

00Z 8 NOV 71 - Initial Match (Figures 16-17). As Table 1 indicates, this is a good match over the entire hemisphere. Major discrepancies are the mid-ocean Highs which are over-smoothed and too far south. Also the Low over North Africa is too weak and displaced southward, and this remains a problem throughout the forecast period.

00Z 11 NOV 71 - 3 DAY FCST (Figures 18-19). A good match is seen in the Western Pacific and the Polar region. However, the forecast does very poorly over the Eastern Pacific and the U.S. region where a major High has formed over New Mexico with Lows along either coast. Note also the development of the Low over the Western Mediterranean.

00Z 14 NOV 71 - 6 DAY FCST (Figures 20-21). Considerable improvement is noted over the +3 DAY FCST. All major features are depicted with the exception of the Eastern Atlantic High/ Low combination which is shown as being much too weak.

00Z 18 NOV 71 - 10 DAY FCST (Figures 22-23). The Pacific Ocean region is a near perfect match. However, the Atlantic and Europe sectors are a rather poor match with the Lows in the Western Atlantic and Northern Europe missed completely and the East Atlantic High much too weak.

D. PS - Winter Series

00Z 8 NOV 71 - Initial Match (Figures 24-25). All major features are shown but the small scale features are smoothed out, i.e., four Lows over Northern Europe on analysis are smoothed to one large Low on composite.

This series illustrates how an analog forecast which starts out good can become rather poor after two or three days and then be quite good again by day five or six. Use of the primitive equation forecast charts should aid in determining the relative fit of the analog composite forecast on a day-to-day basis.

ERRATA SHEET

1. Abstract, line 8: Should be "1000 mb" vice "100 mb".

2. P. 11, section 5.1.3 line 6: Should be "30" percent, vice "300".

3. P. 23, section 6.1.2.2, third line from end of section: Should be "25" year, vice "5" year.

4. P. 27, section 8.0 - Acknowledgements: Third line from end: Should be "Cecilia".

5.2 Regional Products

5.2.1 Analog Dates Message. This message lists the 40 initial historical dates including those making up the composite along with the average large-scale correlations [SL 500 + SR 1000)/2] for the hemisphere and the four sectors for each historical date. A sample message is shown in Figure 2; sector boundaries are depicted in Figure 3.* For comparative purposes, correlation values for the initial match generated by the composite on the same date are also given in Figure 2. It should be noted that the hemispheric composite correlation value is greater than any one of the 10 dates making it up. This information is of value to those users having some form of historical weather information - historical charts, single-station hourly or daily climatological data, etc. Since the use of sectors in analog selection has been observed to improve the skill in regional forecasts under certain conditions, the user of regionally chosen dates to extract corresponding climatological information for forecast purposes can be a valuable forecast tool.

5.2.2 Fifteen-Day Forecast of Daily Average Weather. These forecasts utilize the sector correlations described above. Ten dates from the list of those remaining after application of the history check are chosen for each geographic region based on sector strength. Specifically, U.S. West Coast forecasts use sectors I and II, while East Coast forecasts use only sector I (Figure 3). A sample forecast is shown in Figure 4. These forecasts predict daily average wind direction (by quadrant) and speed, probable maximum and minimum temperature ranges (the wider the range, the less certain the forecast), probable mean temperature, and probability of occurence of rain(R), snow(S), thunder(T), freezing rain/sleet(ZR), drizzle(L) and fog(F).

*Routine dissemination is via the Air Force Automated Weather Network (AWN), MANOP heading ABXN KNWC NRØ1.

Mean temperatures are most useful for showing trends. Trends are also shown by the increases/decreases in probabilities of the different weather parameters from one day to the next throughout the forecast period. If one chooses to use the probability information as a specific daily forecast, the probably of occurence should be compared with that expected by climatology before the forecast is made. For instance, if the climatological probability of rain on any given day is 10 per cent, a forecast of 40 per cent chance of rain strongly suggests that rain should occur on that day.

Additional information can be implied from forecasts for other stations in the same geographic region. For example, if a forecast for Quonset Point, Rhode Island, during the spring predicted cool and damp weather, while the forecast for Washington, D.C., for the same time period was warm and dry, considerable "back-door" cold front activity between the two stations would be indicated. Similarly, rainfall probabilities for U.S. West Coast stations could indicate expected southward penetration of precipitation associated with Eastern Pacific storms.

Routine dissemination of this product to Naval commands was terminated in early 1974 due to lack of user interest. It is anticipated this product will remain dormant until a better forecast technique becomes available.

5.2.3 Sea Height and Direction Probabilities. Using the same regional principles described earlier, sea height and direction probabilities can be generated over ocean areas for use in Optimum Track Ship Routing (OTSR) and for planning subsequence diversions in track. This product is described elsewhere by Selfridge (17).

5.2.4 Limitations. Regional products are limited by the characteristics of currently prevailing weather patterns in that region and also by the time-phase problem. Further, the probabilistic nature of these forecasts precludes accurate prediction of extremes. As with hemispheric forecasts, a previously issued forecast may verify better than the most recent at any given station.

6.0 VERIFICATION

From a verification standpoint, the analog forecast model described in this report is different from dynamic models. Some of the ramifications of this are described in Appendix B. Specific verification of hemispheric and regional analog products is described below.

6.1 General Skill of Hemispheric Products

6.1.1 Qualitative. Substantial differences occur in skill from one geographic region to another on virtually every composite analog forecast issued. Across the hemisphere, however, it has been subjectively observed that general statement type forecasts based on the analog (e.g., "Long wave currently in Eastern Pacific expected to progress to Western U.S. toward end of first week") can be expected to verify about 65 per cent of the time during the first week of the forecast period, and 52 per cent of the time during the second week.

6.1.2 Quantitative. A rigorous quantitative verification program has been developed by Hanna (15) for the hemispheric analog model. No numerical verification has been attempted for the specific geographic regions. Verification is based on the average of normalized scores for all four selection parameters out to 15 days. Statistics are generated as averages for days 1-5, 6-10, 11-15, 1-10 and 1-15. The initial match scores (D_o) are included for comparative purposes. The verification is done by comparing the composite analog scores with corresponding scores for persistence (initial chart used as a forecast chart for each day of the forecast period) and climatology (monthly mean charts used as a forecast chart for each day of the forecast period).

6.1.2.1 Per Cent of Perfect Prediction. Since computation of these scores involves normalized values, a certain strictness occurs which results in relatively low-appearing scores. Table 2 shows comparative annual and seasonal verification percentages for the 47 month period July 1970 to May 1974. The following conclusions are drawn from these statistics:

Seasonal and Year Averages

Period	D_o	+1 to +5	+6 to +10	+11 to +15	+1 to +10	+1 to +15
			Summer (Jun, Jul, Aug)			
Persistence	100.0	41.5	7.8	1.7	24.7	17.6
Climatology	43.0	42.6	41.6	39.8	42.1	41.4
Composite(1)	66.6	48.2	33.7	28.3	41.0	36.9
Composite(2)	70.2	53.6	43.3	36.4	48.0	44.6
			Fall (Sep, Oct, Nov)			
Persistence	100.0	36.6	9.2	1.0	22.9	16.2
Climatology	37.7	37.6	36.3	36.7	36.9	36.8
Composite(1)	68.7	44.9	31.0	26.2	38.0	34.1
Composite(2)	70.6	48.8	32.1	30.6	40.5	37.3
			Winter (Dec, Jan, Feb)			
Persistence	100.0	47.8	19.4	9.7	33.6	26.3
Climatology	38.3	37.0	35.2	34.4	36.1	35.6
Composite(1)	71.8	51.3	35.5	32.6	43.4	39.8
Composite(2)	75.3	55.8	37.3	29.3	46.6	40.8
			Spring (Mar, Apr, May)			
Persistence	100.0	40.2	10.1	5.0	25.2	19.0
Climatology	36.0	37.1	38.2	39.4	37.7	38.4
Composite(1)	66.8	48.4	31.0	26.9	39.7	35.4
Composite(2)	71.0	52.6	39.5	36.9	46.0	43.0
			Overall (Jul 1970 - May 1974)			
Persistence	100.0	40.7	11.0	3.0	25.9	18.9
Climatology	38.3	38.3	37.5	37.1	37.9	37.7
Composite(1)	68.2	47.5	32.6	27.6	40.1	36.0
Composite(2)	72.3	53.1	37.9	33.1	45.5	41.4

(1) Period Jul 70 to Nov 72 and (2) Period Nov 72 to May 74

TABLE 2 - Verification Scores: Per Cent of Perfect Score, July 1970 - May 1974

(1) The initial analog match is substantially (80 - 90 per cent except 60% in summer) better than climatology but consistently worse than persistence (which is always 100 per cent). While the initial match averages only 70 per cent of perfect, the correlation value for large-scale features averages about 0.85.

(2) Persistence is numerically superior to the analog only out to two days. Beyond that, the analog is superior to 15 days and probably more.

(3) Prior to the initial data selection change in 1972, the analog was superior to climatology in the first five-day period for all seasons. It's not quite as reliable as climatology during the six-to-ten day period, except during winter when they were about the same. And the analog is significantly worse than climatology during the 11-to-15 day period. Since the initial analog selection was modified to the updated 12Z fields, a significant improvement in all verification scores is noted. With the exception of fall, these statistics indicate the analog is now superior to climatology out to at least 8 days and, in many cases, may be superior beyond.

The model scores best in winter and worst in summer when weather conditions are changing least rapidly with respect to the climatology. Interpolation of these comparative scores shows that, on the average since Nov 1972, the analog is superior to climatology to about eight days in all seasons except fall where the skill is slightly less.

6.1.2.2 Comparative Verification Statistics. From the verification scores of each forecast period, it can be ascertained whether the analog model performed better or worse than climatology and persistence and in which season. This information has been summarized in Tables 3a and 3b using each forecast period between 1 July 1970 and 31 May 1974. These tables indicate the percentage of cases for

ANALOG BETTER THAN OR EQUAL PERSISTENCE (PERCENT OF CASES)

SEASON (CASES)	INITIAL MATCH	DAYS 1-5	DAYS 6-10	DAYS 11-15	OVERALL 1-10	OVERALL 1-15
SUMMER(54)	0%	74%	94%	89%	94%	94%
AUTUMN(60)	0%	80%	92%	88%	93%	95%
WINTER(38)	0%	68%	89%	89%	87%	89%
SPRING(40)	0%	75%	90%	95%	90%	90%
ANNUAL(192)	0%	75%	91%	90%	93%	93%

ANALOG BETTER THAN OR EQUAL CLIMATOLOGY (PERCENT OF CASES)

SEASON (CASES)	INITIAL MATCH	DAYS 1-5	DAYS 6-10	DAYS 11-15	OVERALL 1-10	OVERALL 1-15
SUMMER(54)	100%	74%	22%	11%	48%	28%
AUTUMN(60)	100%	82%	13%	13%	48%	28%
WINTER(38)	100%	94%	51%	36%	79%	62%
SPRING(40)	100%	80%	35%	18%	58%	40%
ANNUAL(192)	100%	81%	29%	19%	57%	37%

TABLE 3a - Verification Comparisons by Time Periods: Percent of Cases better than stated parameter, July 1970 - November 1972.

ANALOG BETTER THAN OR EQUAL PERSISTENCE (PERCENT OF CASES)

SEASON (CASES)	INITIAL MATCH	DAYS 1-5	DAYS 6-10	DAYS 11-15	OVERALL 1-10	OVERALL 1-15
SUMMER(26)	0%	85%	96%	92%	96%	96%
AUTUMN(32)	0%	92%	92%	96%	92%	96%
WINTER(52)	0%	79%	90%	89%	85%	90%
SPRING(54)	0%	85%	96%	98%	98%	100%
ANNUAL(164)	0%	84%	93%	93%	92%	95%

ANALOG BETTER THAN OR EQUAL CLIMATOLOGY (PERCENT OF CASES)

SEASON (CASES)	INITIAL MATCH	DAYS 1-5	DAYS 6-10	DAYS 11-15	OVERALL 1-10	OVERALL 1-15
SUMMER(26)	100%	100%	58%	35%	84%	69%
AUTUMN(32)	100%	96%	42%	27%	73%	62%
WINTER(52)	100%	98%	56%	29%	77%	64%
SPRING(54)	100%	94%	52%	32%	78%	67%
ANNUAL(164)	100%	97%	51%	29%	77%	65%

TABLE 3b - Verification Comparisons by Time Periods: Percent of Cases better than stated parameter, December 1972-May 1974.

each forecast period and season that the analog was better or worse than climatology and persistence. Generally, the same conclusions as in 6.1.2.1 above can be drawn from Table 5. In particular note the improvement since Nov 1972.

In addition to specific verifications, some general relationships between the analog, persistence, and climatology can be made. Figures 5a and 5b depicts the monthly averages of the 1-5 day verification statistics. The pattern shows a more stable appearance since October 1971 when the analog was switched to twice/week runs. Also included is the initial analog score which aids in evaluating the deterioration of the model during the initial verification period. Figure 5b shows the relationships since the initial data selection change in 1972. The higher initial analog score (D_o) and 1-5 day score should be noted indicating the marked improvement in model performance since that change. In many cases there is a direct correlation between the verification score of persistence and/or climatology with the analog score. The variation in the persistence score is a measure of how rapidly the day-to-day pattern is changing across the hemisphere with lower score indicating the greatest change. The climatology score is a measure of how seasonable a given pattern is compared to a 15 year mean. In this case low scores indicate an unseasonable pattern which would generally indicate a below normal analog verification.

6.2 General Skill of Regional Products

6.2.1 Qualitative. It has been observed that regionalized composite analog forecasts designed for specific locale and purpose appear to be somewhat superior to the more general hemispheric forecasts.

6.2.2 Quantitative. Available resources have precluded comprehensive verification of the regional approach. However, an exercise was undertaken from December 1971 through March 1972 utilizing the regional approach for predicting the occurence or non-occurence of daily rainfall in Monterey, California. A comparison was made of forecasts based on the following:

(1) Climatology - daily probabilities of rainfall.

(2) Persistence - Use the pattern from the past 14 days to forecast for the next 14.

(3) Composite - ten best hemispheric dates.

(4) Regionalized - dates used in the operational forecast for Monterey.

(5) Hand-selected - dates which considered past history at the station as well as degree of zonal flow.

Time phase errors were not considered. The results, shown in Table 4, demonstrate that using regionalized analogs may improve the forecast; however, these results are not sufficient to make any general statement.

PREDICTION OF DAILY RAINFALL OVER 14-DAY PERIODS

AT MONTEREY, CALIFORNIA, 1 DECEMBER 1971-31 MARCH 1972

(34 TWO-WEEK FORECASTS)

METHOD	PER CENT OF DAYS CORRECTLY FORECASTED	
	Week 1	Week 2
Daily Climatology	60.9%	60.9%
Persistence (previous two weeks)	59.6%	60.7%
Hemispheric Composite Dates	61.7%	54.2%
Automated Regionalized Dates	67.7%	56.2%
Hand-Picked Regionalized Dates	68.4%	60.6%

TABLE 4

7.0 SUMMARY

There is little doubt that long-range forecasting using synoptic models is an imperfect art. The skill demonstrated by the average composite analog forecast model over other conventional means of ascertaining hemispheric weather patterns beyond a few days gives purpose to the technique. Despite this success, our inability to select the ten analog dates that ultimately turn out best (Appendix A) suggests that much more work needs to be done and improved selection models must be developed. The planned shift to a regression type composite in the future should improve both the skill and the detail of the forecasts.

This type of model is not designed to be permanent. With the advent of improved dynamic models using fourth and fifth generation computers, and with improved initialization through the use of automatic weather stations and satellites, dynamic prediction should become superior to synoptic and statistical techniques. At this point, analogs will best be used to relate dynamic forecast patterns to expected actual weather. Until that time, however, the development of the analog approach described in this report appears to be the best method of providing operationally usable and totally automated long-range forecasts for Naval operations.

8.0 ACKNOWLEDGEMENTS

Many people have at one time or another been involved with the development of FNWC's Automated Long-Range Forecast Model. The contributions of several deserve special recognition. First we recognize the contributions of Mr. Peter Hanna, whose statistical techniques were the key to the development and testing of the current operational model. Recognition must also go to Lieutenant Commander R. C. Corkrum and Mr. James A. Zuver who, while attached to FNWC's Climatology Department, almost singlehandedly assembed the historical data base over a three year period. Special thanks also goes to Mr. Richard Elms who programmed and streamlined the current operational model and many of the developmental models along the way.

Additional thanks go to Lieutenant Commanders J.B. Tupaz and R.D. Gilmore, Captain C. Barteau, Captain C.R. Ward, Mr. Peter Hanna and Mr. W.E. Hubert for reviewing the manuscript. Finally a special thanks goes to the authors wives, Evelyn and Cecilia, for their help in reviewing the manuscript and their patience and understanding during the many extra hours required to complete this project.

REFERENCES

1. Elliot, R. D., Extended Weather Forecasting by Weather Type Methods, U.S. Navy Department, Washington, D.C., February 1944.

2. Wolff, P. M. and W. C. Woodworth, Hourly Point Forecasting at FNWC, Technical Memo #16, Fleet Numerical Weather Facility, Monterey, California, September 1966.

3. Wolff, P. M. and C. D. Thormeyer, Extended Range Analog Forecast Program at Fleet Numerical Weather Central, Technical Note #49, Fleet Numerical Weather Central, Monterey, California, September 1969.

4. Thormeyer, C. D., "The Development of Automated Short-Range Terminal Forecasting using Analog Techniques," Automated Weather Support, Proceedings of the 6th AWS Technical Exchange Conference, U.S. Naval Academy, 21-24 September 1970 (published as Air Weather Service Technical Report #242, April 1971).

5. Hartranft, F. R., J. S. Restivo and R. C. Sabin, Computerized Map Typing Procedures and their Application in the Development of Forecast Aids, Technical Paper 70-2, Aerospace Sciences Division, Headquarters 4th Weather Wing, Ent Air Force Base, Colorado, August 1970.

6. Elliot, R. D., "Extended-Range Forecasting by Weather Types," Compendium of Meteorology, American Meteorological Society, Boston, Massachusetts, 1951.

7. White, R. M. and W. C. Palson, Jr., On the Forecasting Possibilities of Empirical Influence Functions, Air Force Cambridge Research Center, Journal of Meteorology, Vol. 12, No. 5, October 1955.

8. Hofmann, A., Monthly Forecasts for Germany - 6 Year Verifications, Scientific Report No. 3, Forschungsabteilung des Deutchen Wetterdienstes, Frankfurt/Main, June 1956.

9. Lund, I. A., Map Pattern Classification by Statistical Methods, Journal of Applied Meteorology, Vol. 2, No. 1, February 1963.

10. Lorenz, E. N., Atmospheric Predictability as Revealed by Naturally Occurring Analogues, Journal of the Atmospheric Sciences, Vol. 26, No. 4, July 1969.

11. Augulis, R. P., Precipitation Probabilities in the Western Region Associated with Winter 500-mb Map Types , ESSA Technical Memorandum WBTM WR 45-1, Salt Lake City, Utah, December 1969.

12. Holl, H. M., Scale and Pattern Spectra and Decompositions, Technical Memorandum #3, Meteorology International, Inc., Monterey, California, 1 November 1963.

13. Thormeyer, C. D., Scale and Pattern Decomposition at Fleet Numerical Weather Central - Explanation and Usage Guide, Fleet Numerical Weather Central Working Paper, Monterey, California, 24 February 1970.

14. FNWC Data Bank Inventory, Fleet Numerical Weather Central, Monterey, California.

15. Hanna, P., Composite Analog Statistics 1970-1971, Internal unpublished manuscript, Fleet Numerical Weather Central, 21 July 1971.

16. United States Naval Weather Service Numerical Environmental Products Manual (in publication).

17. Selfridge, S. W., An Extended-Range Probablistic Sea Condition Forecast Data Base as Applied to the Automated Optimum Ship-Routing Problem, Technical Note #3, Project M174, Meteorology International, Inc., Monterey, California, 1972.

18. Kesel, P. G. and F. J. Winninghoff, Fleet Numerical Weather Central's Four Processor Primitive Equation Model, Fleet Numerical Weather Central, Monterey, California, January 1971.

APPENDIX A

EXPERIMENTATION WITH THE SELECTION PROCESS

A major problem in developing a composite analog forecasting method involves optimum selection of the historical dates making up the composite. In-house studies have demonstrated that (a) the maps with the highest initial or past correlations do not necessarily produce the best analogs and (b) the model rarely comes close to picking the 10 analog candidates from the original 40 that ultimately turn out best. However, the current selection method has produced results comparable to or better than all other selection methods so far investigated and it also has been the most stable. The principal techniques that have been tested are discussed and summarized below.

1. Selection Variations for the Initial Match (D_o)

1.1 High-Low Pairs. This early scheme was based on matching the count and location of highs and lows on current SL 500 charts with those on historical SL 500 charts. This method failed partly because of such analysis quirks as large highs shown split into two centers one day, merged into one on the next, and split again on the third. Thus, good candidates were screened out prematurely.

1.2 500 mb Large-Scale Patterns. Another early model selected candidates solely on the basis of the SL 500 chart. This failed partly because there was no correlation with surface features and partly because of the incomplete data base at the time.

1.3 Hemispheric Zonal Wind. This third early screen was a function of the square of the difference between current zonal winds and zonal winds of the potential analog, and of the square of the cosine of the latitude. Preliminary results indicated that this initial screening mechanism was somewhat unstable.

1.4 Screening Order of Large-Scale Parameters. From March 1969 until July 1970, the operational analog selection model used a screening procedure in which a certain percentage of the analog candidates were eliminated with each pass through the screen. Most screens consisted of from two to four passes utilizing various parameters. The most effective screens eliminated all but 10 - 20 per cent of the available dates on the first pass, 50 per cent of the remainder on each succeeding pass, leaving 40-60 candidates for further processing. Screens which were considerably looser or tighter provided worse results.

Parameters screened and order of screening varied as developmental work continued. At first, only large-scale parameters were used; later, large-scale parameters were used for the initial screens only. The three major large-scale parameters investigated were:

SL 500 (Large-scale disturbance field at 500 mb)

SL 1000 (Large-scale disturbance field plus zonal component at 1000 mb)

SR 5-10 (Large-scale disturbance field plus zonal component of 500-1000 mb thickness)

Surprisingly, use of the thickness parameter in the screen produced the least satisfactory results. Although this parameter is a measure of the hemispheric large-scale thermal pattern, the very small spread in correlation values (relative to those large-scale values at 500 and 1000 mb) negated its value. In any event, use of both SL 500 and SR 1000 implies the large-scale thickness. The order of these latter two parameters in the screen was found to make little difference in the ultimate outcome.

1.5 Small-Scale Parameters. For a time, independent small-scale (SD) features at 500 and 1000 mb were included near the end of the basic screen as final selectors. There turned out to be little relation between correlations of large and small scale features from the same chart, making this effort a failure. There appear to be four basic reasons for this: (1) There are an infinite number of small-scale synoptic patterns that can be superimposed onto the same basic hemispheric large-scale pattern; (2) Due to the larger mesh size of much of the older historical data, many small-scale features that were probably there had been smoothed out during reanalysis; (3) Satellites have vastly increased our knowledge of oceanic synoptic features, thereby making questionable many oceanic analysis prior to the early 1960's; and (4) Even with satellites, wavelengths of the SD scale over sparse data areas are not always accurately depicted, especially at 500 mb.

1.6 Total Disturbance Fields. In order to utilize small-scale information, yet keep it dependent upon good large-scale correlations, the total disturbance field (TDF) was devised. This field was created by subtracting the same zonal component of the flow out of both current and historical charts. The TDF technique put the 500 and 1000 mb fields on a relatively equal basis when used as a final eliminating parameter in the screening process. More importantly, it increased the spread of the correlation values, making elimination and/or selection easier. This was especially vital at 500 mb where height values correlate very strongly with latitude. Once the large-scale had been accounted for, this technique proved to be the most effective screen in the operational model prior to July 1970. However, the 40-60 remaining analog dates prior to any history check in that model were almost always very similar to the 40 dates remaining using concurrent tests of the present operational model. Since the current model uses the normalizing technique described earlier, the need for a TDF was eliminated.

2. Checks for Historical Continuity. For an analog sequence to be good into the future, it is reasonable that the analog sequence should be accurate in the immediate past.

Some form of history check is useful to narrow the number of candidates from the initial 40 to the 10 in the composite. Although a number of variations of the history check have been investigated, none have resulted in significant improvement in skill over the current operational version. Since the 10 dates chosen operationally rarely approach the skill of the best 10 possible (Table 5), investigations into improved history checks continue. A summary of the testing in this area to date is presented below:

2.1 Manual History Checks. Before the automated version of the analog selection model went into operation, historical continuity of the 40-60 remaining analog candidates was examined manually. Special emphasis was placed on the two to three days immediately preceding the current date. Over an eight-month period, both manual and automated models were run concurrently, and neither was consistently shown better than the other (Table 5). The automated version, however, has the advantage of saving considerable time and manpower.

2.2 Three and Six Day Weather Patterns. This was an early attempt to automate the history check by incorporating a historical continuity of three or six days and involved two techniques to match historical weather sequences with those currently occurring.

(1) An average of the absolute values of the SD patterns for the three or six days prior to the current date was computed. This was designed to be a measure of the strength of the activity of the transient small-scale features.

(2) An average of the absolute values of the 24-hour changes in SD patterns for the three or six days prior to the base date was computed. This was designed to be a measure of the movement of the transient small-scale features.

Based on six widely varying test cases with this model, going back only three days tended to be superior to going

SKILL OF COMPOSITE ANALOG TESTS USING VARIOUS HISTORY SCREENS

(APPENDIX A)

1 JULY 1970 - 14 JUNE 1971 (50 CASES)

TEST INVOLVED	AVERAGE RAW SCORE	COMPARED TO OPS MODEL		
		BETTER	WORSE	SAME
Best 10 Possible (Hindsight)	.974*	50	0	0
Operational Model	.465	0	0	50
Manual History Check (35 Cases)	---	17	18	0
Ten-Day Average History	.448	21	28	1
Five-Day Average History	.448	18	28	4
Tightest History Screen	.400	19	31	0
Tighter than Ops Model	.446	13	23	14
Slightly looser than Ops Model	.455	20	17	13
Loosest (No Past Requirement)	.466	24	18	8

*Note: Perfect prediction would be normally between 3.00 and 4.00. A Composite of the best 10 possible would score considerably higher than the average of the 10 best as above.

TABLE 5

back six; even so, this was neither significantly better nor worse than ignoring this history check altogether. Although theoretically sound, results of this scheme were inconclusive since it suffered from the same drawbacks of historical SD patterns as discussed earlier. Since there is apparently an infinite number of hemispheric SD variations that can occur in a given large-scale pattern, averaging these over a three or six day period showed very little spread in the correlation values among the top 200-500 candidates. Plots of these patterns showed very ill-defined fields rather than obvious stormtracks.

2.3 Past History Check - Five and Ten Day Averages. In the first experiment, the 20 dates (of the original 40) with the highest average normalized values over the previous 10-day period were saved. The 10 best D_o matches of these were chosen for the composite. The experiment was repeated for a shorter history period of only five days. In both experiments, the current operational version was substantially better (Table 5). There was very little difference between ten days of history and five.

2.4 Variations in the History Screen Limits. The 1.25/2.50 limits for periods D_o-1 to D_o-5 and D_o-6 to D_o-10, respectively, were varied in order to produce tighter and looser screens as follows:

(1) Set the history screen to the tightest value possible, increasing it at increments of 0.125/0.250 until at least 10 of the original 40 dates pass it. Use the best 10 of these remaining dates at D_o for the Composite (tightest screen).

(2) Same as (1) except insist that at least 20 dates pass the history screen. This is a slightly looser screen but tighter than the operational version in most cases.

(3) Same as operational model except limits are slightly looser at 1.50/3.00 vice 1.25/2.50.

(4) Insist on no past requirement whatsoever. (No history check; thus loosest possible screen)

The one general conclusion from this set of tests was that, over the 50-week period from 1 July 1970 through 14 June 1971, the looser the history screen when compared to the operational screen, the better the composite analog tended to verify. For screens looser than the operational version, however, the 50-week average of 15-day verifying scores were generally the same as or slightly lower than those of the operational model. This further implies less stability than with the operational version. See Table 5.

One possible explanation of the apparent superiority of the looser screens is as follows: Although a history check is applied to eliminate criss-crossing and retain parallel weather sequences, weather sequences are often inconsistent with each other. Thus, an analog with a good past parallel sequence may not produce a good forecast. Similarly, an analog with poor past history may produce an excellent future match. The degree to which this is true seems to be dependent upon the uniqueness and/or stability of the past and current weather patterns.

2.5 Variations in the History Screen Ratio. The history screen normally has a 1:2 ratio for the two five-day periods; i.e., 1.25/2.50. This ratio was adjusted several different ways in an attempt to improve the date selection for the Composite. No consistent results were found.

3. Possible Future Selection Techniques. There are several approaches which might improve the analog selection procedure although none have been thoroughly tested. It must be remembered that even small changes (good or bad) may have a marked influence on model performance.

a. Use of the Primitive Equation (PE) Model Forecasts. The FNWC PE model described by Kesel (18) has shown more skill in predicting large-scale features than small-scale ones.

Large-scale skill is often good at 48 hours and may still be quite good at 72 hours. Several methods of combining PE and Analog forecasts have been considered.

(1) Use of the PE as Part of the History Screen. This method is valid if we can assume that PE large-scale forecasts verify reasonably well out to 48 hours. Once the initial 40 historical analog dates are available, they could be narrowed down to 10 by considering the large-scale correlations of each date at both SL 500 and SR 1000 for the period D_o-2 to D_o+2. This five-day period includes the current chart vs. the historical analog, the one and two-day-old histories, and the 24 and 48-hour large-scale progs vs. each analog date advanced 24 and 48 hours. The highest 10 two-level five-day averages could then be selected for a composite analog forecast. Very limited tests of this technique, utilizing a "perfect prog" concept, have indicated potential value of this technique.

(2) Use of the PE as a "Jumping Off" Point. The easiest way to combine the PE and analog forecast models would be to use the latter as an extension of the former. Since the Analog tends to show more skill than the PE in terms of comparison with persistence after 48-96 hours, the 48 or 72 hour time frame would be a logical "jumping off" point. The analog model could be revised to search for the 40 analog dates based on the P.E. large-scale progs instead of the current analyses. The history check with adjusted weighting factors could still go back 10 days (D_o+2 to D_o-2 and D_o-3 to D_o-7), thus adding stability, especially in instances where the progs were partially in error. Despite the likelihood of some error in the large-scale progs, the unexplained variance resulting from the impossibility of perfect matches might easily be greater than the variance caused by a large-scale prog being in error. A large-scale prog might correlate higher with its verifying chart than with a composite analog based on it.

This could either harm or help such an analog selection scheme. No testing or modeling has been undertaken involving this technique.

b. The Use of Zonal Indices. An in-house study in 1969 seemed to indicate a periodicity of about 12 days during the winter months. The Composite analogs which do best tend to be the ones which have initially matched a hemispheric synoptic pattern containing strong meridional flow, and in a relatively common pattern of well-defined blocks and long-waves. A well-established two-block five-wave pattern appears to be most favorable. Those analogs which start off matching a highly zonal pattern usually perform worse, despite a tendency for higher initial correlations. Thus, the use of zonal indices as possible selection criteria for analog candidates is undergoing renewed investigation.

4. Summary of the Selection Process Experimentation. Few selection methods have been found which are significantly better or worse than the others. It appears that simply designed selection models perform as well (if not better) than highly complex approaches. The 40 dates consist of essentially the same groupings of historical weather sequences, regardless of the selection method used to isolate them initially, as long as the SL 500 and SR 1000 parameters are involved. The most stable selection method found up to now is the normalizing technique currently in operation. The worst results have been obtained when using very tight or very loose parameter screens, when using thickness parameters, when using only one height level, or when using small-scale disturbance fields. With regard to history checks, it can be said that (a) the operational weighted history screen is superior to a straight "averaged" screen, (b) there is little difference in skill between a five-day and ten-day history check, and (c) the tighter the history screen, the worse the skill of a composite analog forecast. For maximum stabilization, however, a gross history check is considered prudent. For the future, the PE model, measures of zonal indices, and relative sector correlations all appear promising for incorporation into the selection procedure.

APPENDIX B

CHARACTERISTICS OF ANALOGOUS WEATHER PATTERNS

Several years of investigation into the long-range forecast problem has produced the following conclusions about the characteristics of analogous weather patterns.

1. The Initial Match

1.1 A perfect or near perfect match of an individual historical analog using the current data base is unattainable.

1.2 Hemispheric analog matches and forecasts rarely verify equally throughout the entire Northern Hemisphere at the same time, but will match better in certain sectors of the globe.

1.3 A high correlation at 500 mb does not ensure the same at the surface and vice versa, especially in individual sectors.

1.4 A more meaningful match is obtained using large-scale rather than small scale features.

1.5 More good analog matches occur with common, well-established large-scale weather patterns than with transitory, unorganized or highly zonal patterns.

1.6 Initial composite matches are better in summer and winter than in transition seasons. Overall, the best correlations are found over the Eastern Atlantic and Europe in summer with worst correlations over North America and the Western Atlantic during transition seasons. The above information is summerized in Table 6, which shows the hemispheric and geographical sector correlations of the Total Disturbance Field (TDF) of the composite analog. These figures are averaged over two periods of time: (1) Jul 70 through Nov 72, (2) Nov 72 through Aug 74. The marked improvement (28%) in the initial match after November 1972 can be attributed to the initial data selection change discussed in Para. 4.3. Since TDF's (Appendix A) are a measure of both large and small scale features with the zonal component removed, these correlation values are generally 0.10 to 0.15 lower than those of the corresponding large-scale pattern alone.

VARIATION OF INITIAL COMPOSITE ANALOG MATCH BY SEASONS
AVERAGE OF 500 MB and 1000 MB TOTAL DISTURBANCE FIELDS
(ZONAL FLOW REMOVED)

	HEMIS 360 DEG	N. AMER 40W-130W	PACIFIC 130W-160E	WESTPAC 160E-60E	ELANT/MED 60E-40W
ANNUAL					
(JUL70-NOV72)	.6949	.6466	.6623	.6470	.7039
(NOV72-AUG74)	.8750	.8344	.8528	.8625	.8740
SUMMER (JUN-AUG)					
70-72	.7398	.7110	.6411	.6482	.7642
73-74	.9035	.8518	.8510	.8684	.9136
AUTUMN (SEP-NOV)					
70-72	.6454	.5887	.6591	.6242	.6293
72-74	.8523	.8192	.8211	.8542	.8622
WINTER (DEC-FEB)					
70-72	.7352	.7006	.7069	.7069	.7305
72-74	.8771	.8451	.8550	.8715	.8917
SPRING (MAR-MAY)					
71-72	.6720	.5986	.6506	.6206	.7104
73-74	.8579	.8153	.8720	.8523	.8230

TABLE 6

2. Characteristics of Prediction Ability

2.1 Long-lasting individual analogs are rare. These occasional winter-time analogs can often be followed on a day-to-day basis for two to four weeks. A good example occured from 25 December 1969 through 3 February 1970 which matched on a hemispheric basis almost daily the period from 7 January through 16 February 1960, or 41 days. A more recent case showed the period from 10 February through 12 March 1972 matching the period 11 January through 11 February 1951 or 32 days.

2.2 The inability to forecast the location and intensity of individual storms using analog techniques limits the techniques to forecasting large scale changes.

2.3 The best initial matches do not necessarily ensure the best forecasts. A well established synoptic pattern will normally produce better forecasts than those with unique or highly zonal patterns, regardless of initial correlation values.

2.4 Unlike dynamic forecast models, analog forecasts do not in general lose accuracy continuously with time. In a given geographic area, the three-day forecast may verify better than the initial match, the seven-day forecast better than the three-day, etc. Thus, each composite analog forecast issued must be judged on its merits in each geographical area of interest. If a newer forecast is not performing as well as an older forecast after a few days, then the older forecast should continue to be used.

2.5 The progression of analogs is often somewhat irregular, making it necessary to allow for time-phase errors in all operational forecasts. Generally an allowance of $\pm$ 2 days in verifying the occurence of a predicted weather event is adequate.

2.6 Use of recurring weather patterns (or cycles) for long range forecasting is prevalent in both the literature and in practice. These cycles range from six-day weather types to 22-year double sunspot cycles. An extensive investigation of cyclic effects in the analog model was undertaken but the results were inconclusive. Occurrence of analogs was fairly evenly distributed across the entire 28-year history with no apparent cyclic influence except for a very slight hint of a 10 - 12 year pattern.

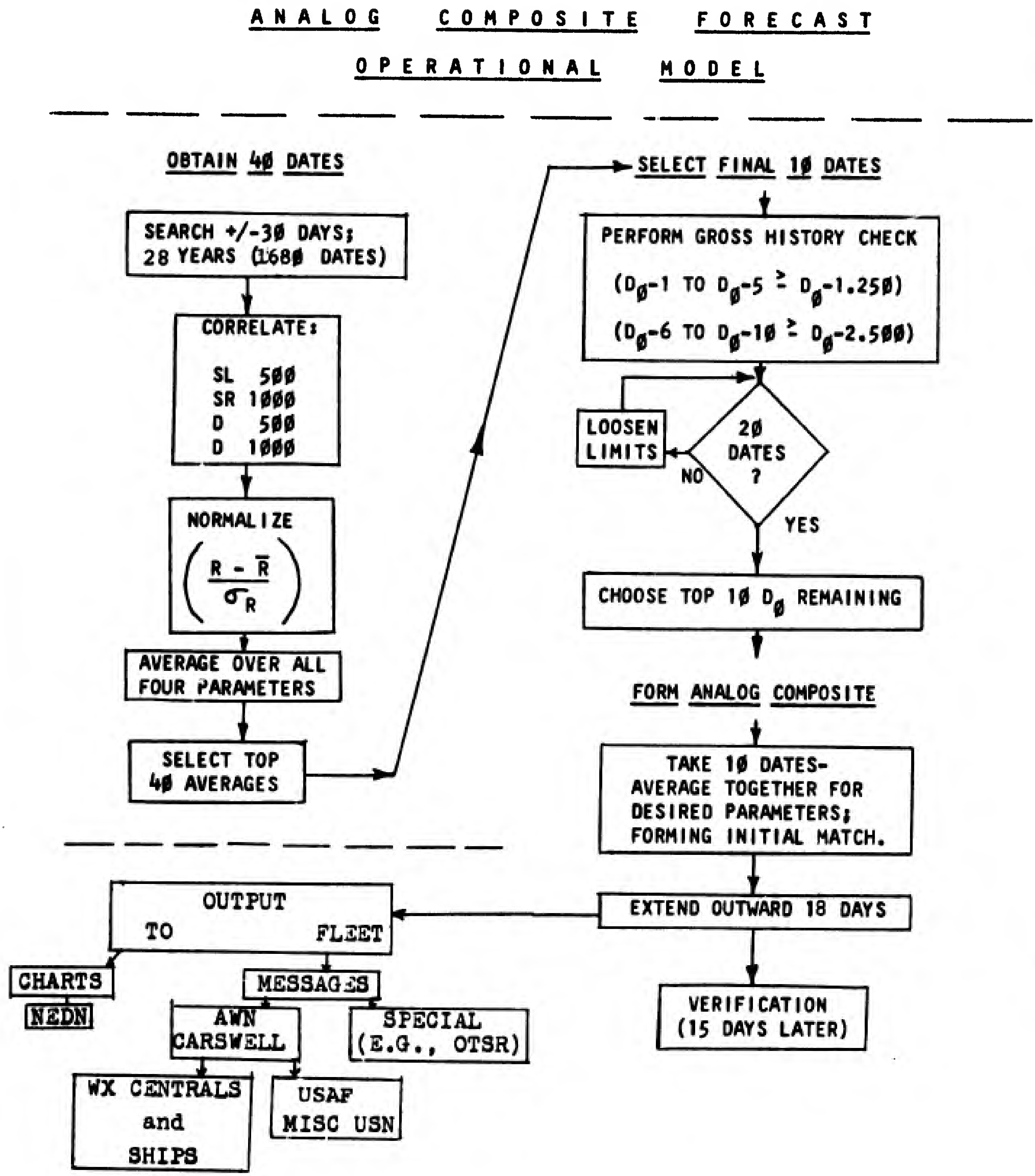


Figure 1 - Flow Diagram of Operational Analog Model

ANALOG DATES MESSAGE - FOR METEOROLOGIST ONLY

1. ANALOGS LISTED BELOW MATCH CURRENT DTG OF 78030712

2. QUALITY OF BEST ANALOGS VS. INITIAL MATCH---
(AVERAGE OF SL 500 AND SR 1000, RANGE CAN BE FROM -1.000 TO +1.000). LAST COLUMN INDICATES THOSE DATES ACCEPTED (YES) FOR COMPOSITE ANALOG.

BEST DATES YYMMDDHH	HEMIS 360 DEG	N. AMER. 40W-130W	PACIFIC 130W-160E	WESTPAC 160E-60E	EI ANT/MED 60E-40W	ACCP COMP
40021312	.6077	.4819	.7288	.5752	.6568	NO
52020512	.8276	.1548	.7349	.8536	.5482	NO
52020612	.5917	-.2213	.8888	.7050	.5032	NO
52022412	.8289	.1506	.8864	.7358	.6746	NO
52022512	.6899	.1307	.9427	.6773	.6529	YES
52022612	.6523	-.1287	.8614	.6736	.6106	NO
58022012	.5877	.5617	.7419	.6323	.6196	NO
59022012	.5899	.6803	.6407	.6426	.4502	NO
57022812	.6588	.2689	.8210	.7990	.4306	NO
58022812	.8338	.2447	.8694	.7643	.0303	NO
68021312	.6037	.4362	.6880	.7069	.6294	NO
68021412	.6536	.4353	.6698	.8218	.6833	YES
68021512	.6615	.2730	.7508	.8325	.7036	YES
68021612	.6880	.3317	.8511	.8532	.6466	YES
68021712	.7295	.5150	.8640	.8404	.6623	YES
68021812	.7336	.6401	.8328	.7917	.6846	YES
68021912	.7095	.6410	.8584	.6591	.5987	YES
68022412	.5895	.4728	.8092	.7356	.3961	NO
68020612	.6600	.5475	.8100	.8144	.2909	NO
68020712	.7175	.7847	.8548	.8857	.3524	YES
68020812	.6468	.7078	.8716	.8154	.3463	NO
68020912	.5914	.6317	.8092	.7433	.1783	NO
71020812	.8245	.3390	.6830	.8124	.6059	NO
71022012	.5832	.4314	.7650	.6620	.3363	NO
71022112	.5665	.3691	.7706	.5155	.4723	NO
40032112	.5778	.4798	.5817	.7995	.5626	NO
40032212	.6323	.7350	.8072	.6772	.2322	NO
40032312	.8828	.8036	.7951	.7224	.2481	YES
50032212	.5749	.4098	.7286	.5880	.5028	NO
50032312	.5796	.3535	.8614	.4140	.5406	NO
52031012	.5755	.6159	.7636	.7419	.5754	NO
57031312	.8240	-.0367	.9190	.7365	.3676	NO
57031412	.8620	.0013	.9677	.7360	.4647	NO
57031512	.6611	.0019	.9406	.6839	.4026	NO
57031612	.8098	-.0175	.9122	.5681	.3918	NO
68030212	.6348	.3809	.8717	.7782	.5598	NO
68030312	.6424	.3062	.9296	.7739	.6146	NO
70031912	.8242	.5782	.8667	.7687	.5166	NO
70032012	.6496	.6043	.8268	.7257	.6120	YES
70032112	.8419	.3476	.8528	.7277	.5329	NO
$\left[\frac{D\ 500 + D\ 1000}{2}\right]$	.8749	.7239	.9481	.8900	.8350	

Figure - 2 Sample Analog Dates Message

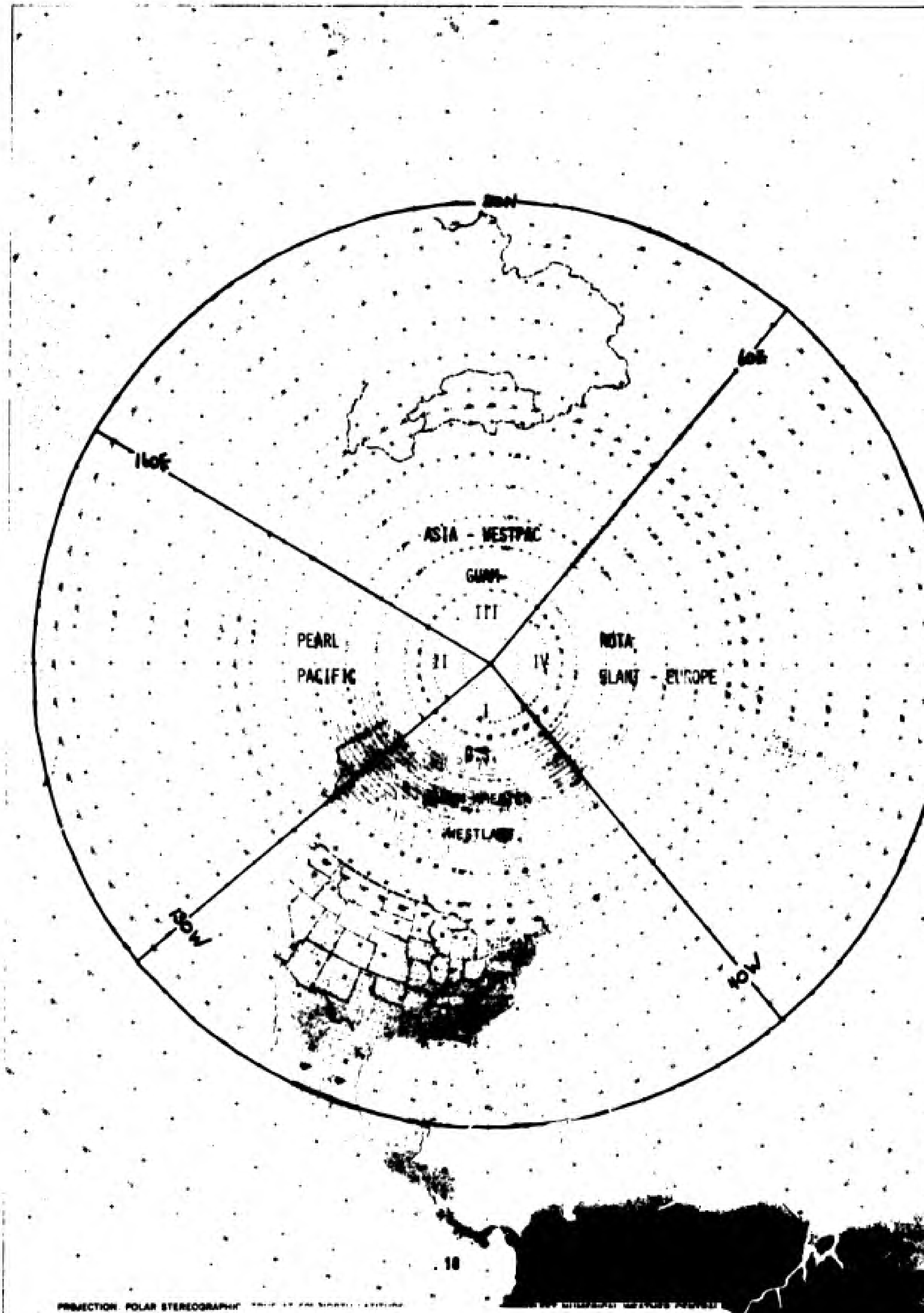


Figure 3 - Sector Boundaries of Hemispheric Analog

CURRENT ANALYSIS DTG, 20 DEC 1971

SELECTED ANALOG DTGS,

11 JAN 1947
12 JAN 1947
13 JAN 1947
14 JAN 1947
23 DEC 1951
24 DEC 1951
6 JAN 1952
6 DEC 1956
28 DEC 1965
29 NOV 1970

15 DAY ANALOG FCST OF DAILY AVERAGE WEATHER FOR ALAMEDA
COMMENCING 0000 LST 20 DEC 1971.

•	SIG WIND BY QUADRANT AND AVG SPEED (KTS)
••	MAXIMUM WIND DIRECTION AND SPEED
*	MAXIMUM TEMPERATURE RANGE (F)
**	MEAN TEMPERATURE (F)
***	MINIMUM TEMPERATURE RANGE (F)
•••	SIGNIFICANT WEATHER (PERCENT PROBABILITY / DAY)

DATE	WIND •	WIND ••	TEMPERATURES *	TEMPERATURES **	TEMPERATURES ***	SIG WX •••
20 DEC	E 6	2723	52/49	46	44/39	R 50 L [illegible] F 50
21 DEC	E 8	1325	53/50	47	44/39	R 60 L [illegible]0 F 20
22 DEC	W 8	1325	55/52	48	44/39	T 10 R 40 L 20 F 30
23 DEC	E10	1323	55/52	4[illegible]	45/38	R 50 L 10 F 40
24 DEC	E 8	1324	55/52	4[illegible]	47/40	R 40 L [illegible]0 F 90
25 DEC	E 7	1328	57/52	49	47/40	R 40 L [illegible] F 70
26 DEC	W 9	2925	56/51	48	45/40	R 30 L 10 F 70
27 DEC	E 5	1838	55/50	49	46/39	R 30 F [illegible]
28 DEC	E 3	2126	56/51	48	46/39	R 30 L [illegible] F 70
29 DEC	W 6	2520	58/53	49	45/40	R 20 F 60
30 DEC	E 5	1321	57/52	49	45/40	R 10 L 10 F 70
31 DEC	E 4	3116	57/52	4[illegible]	44/41	L 10 F [illegible]
1 JAN	E 6	1323	57/54	50	45/42	R 10 L 10 F 80
2 JAN	W10	2929	56/51	49	44/41	T 10 R 50 L 20 F 60
3 JAN	W11	2929	54/49	47	43/40	T 10 R 60 L 20 F 50

Figure 4 - Sample 15-Day Forecast of Daily Average Weather

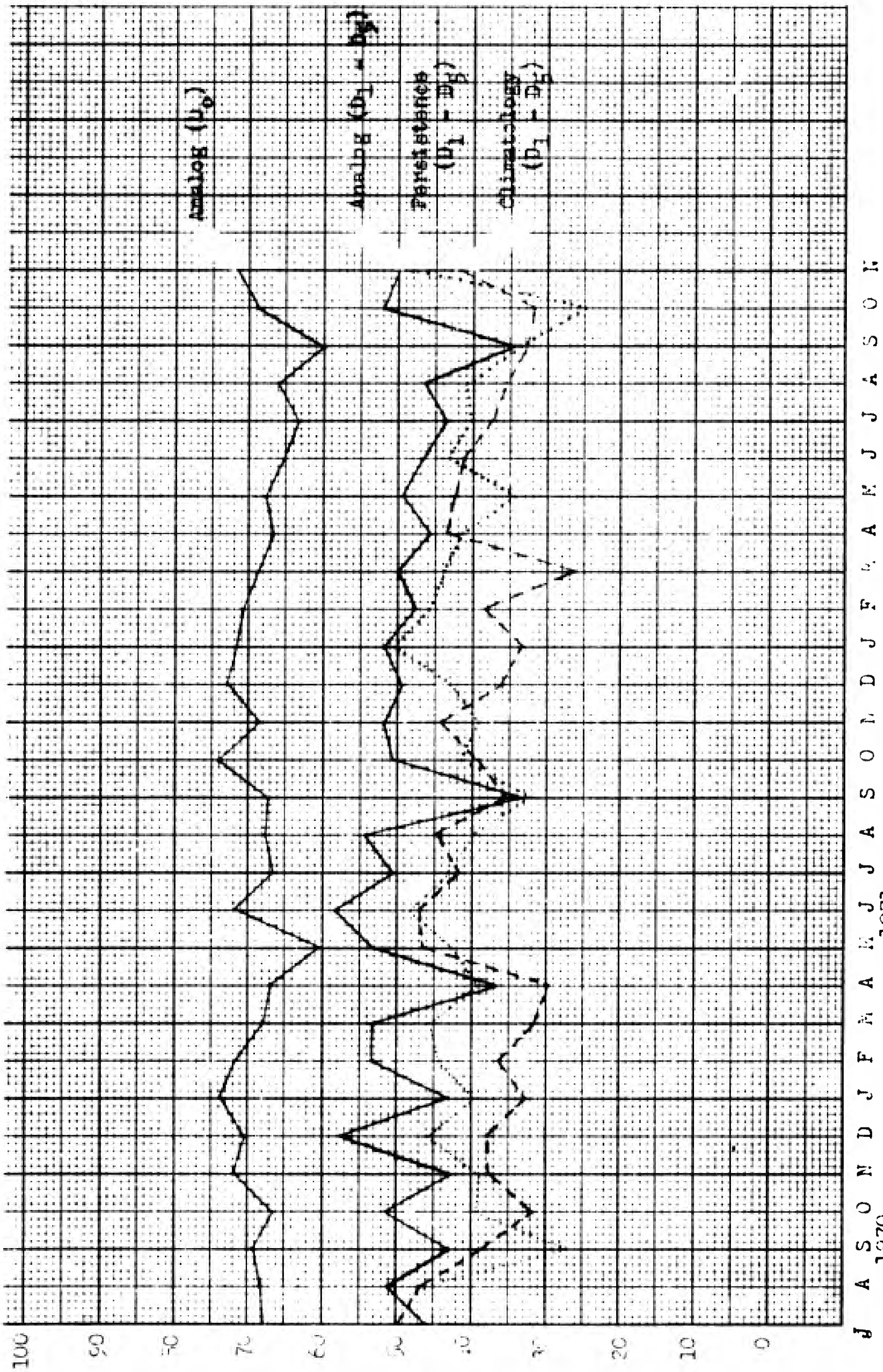


Figure 5a - Monthly Comparison of Analog Initial Match with Analog, Climatology and Persistence 1-5 day Forecast (July 1970 - November 1972)

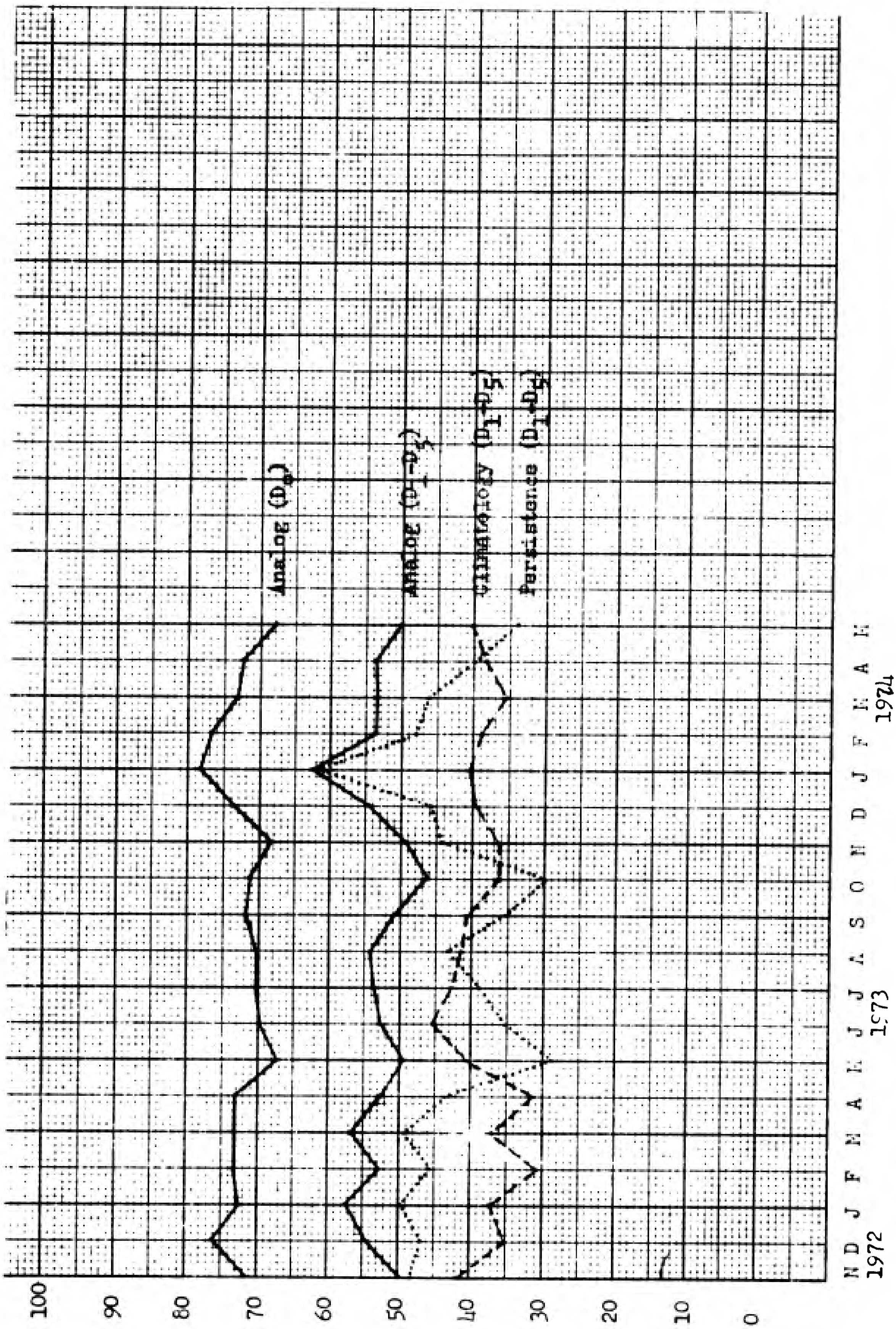


Figure 5b - Monthly Comparison of Analog Initial Match with Analog, Climatology and Persistence 1-5 day Forecast (November 1972 - May 1974)

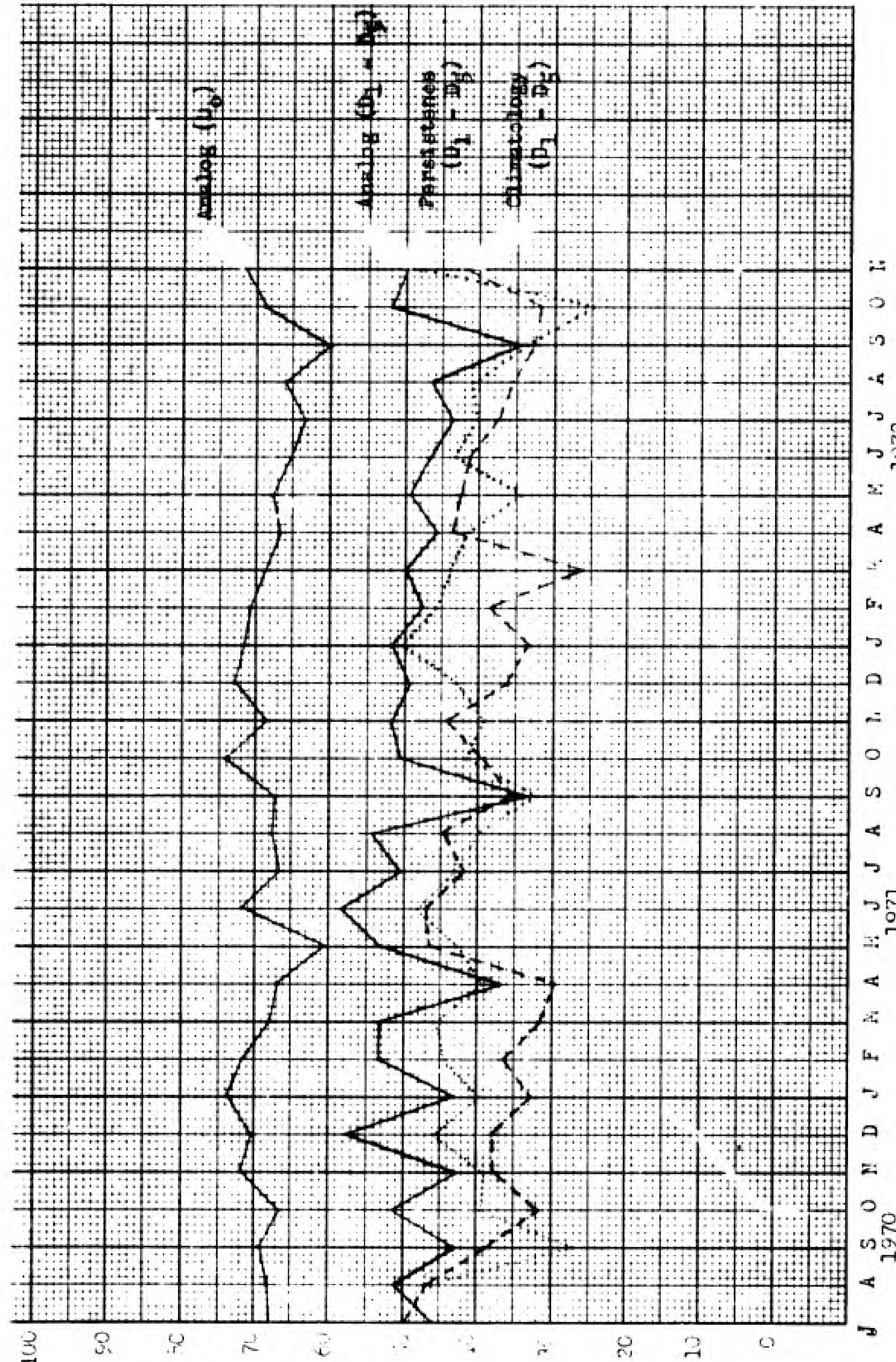


Figure 5a - Monthly Comparison of Analog Initial Match with Analog, Climatology and Persistence 1-5 day Forecast (July 1970 - November 1972)

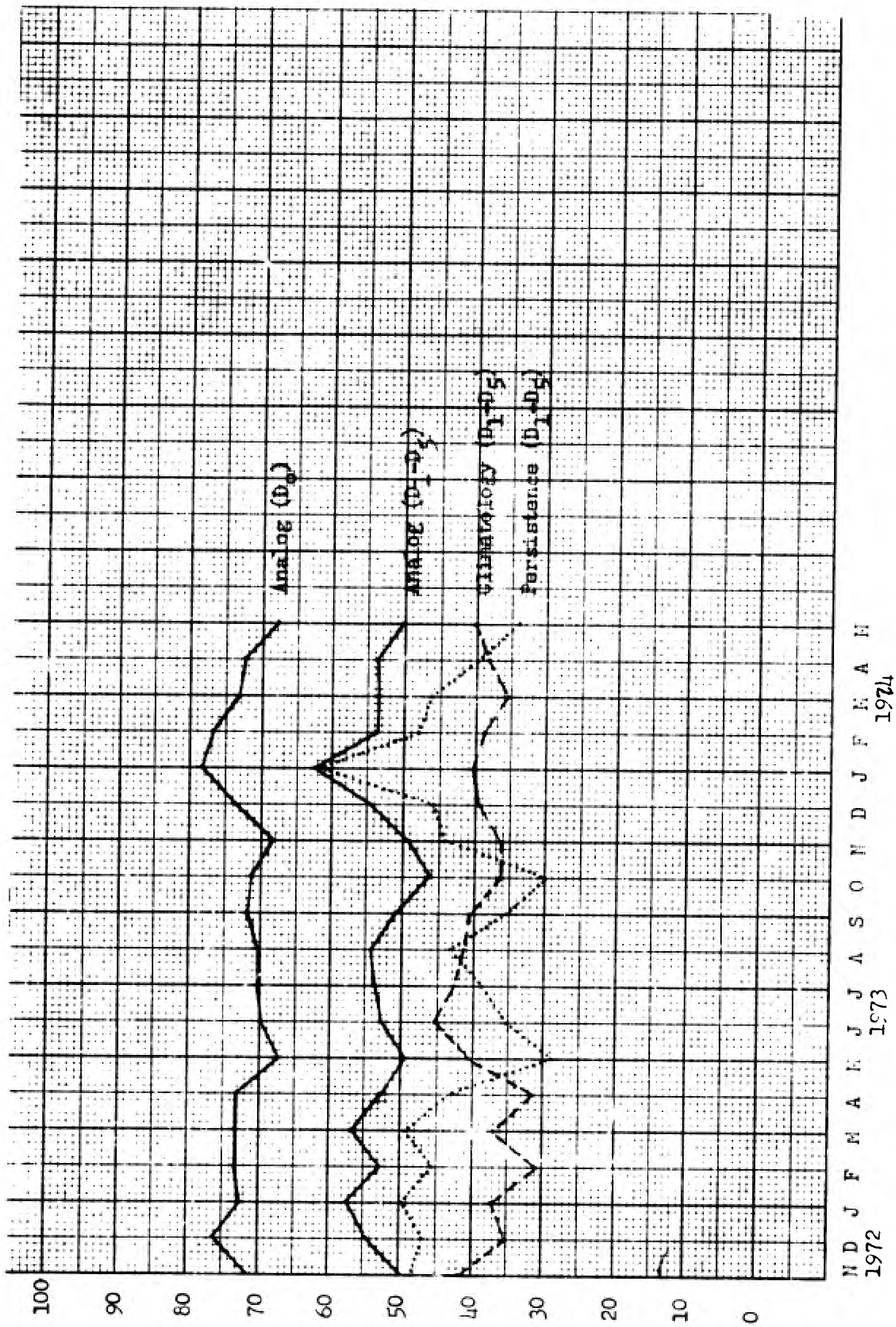
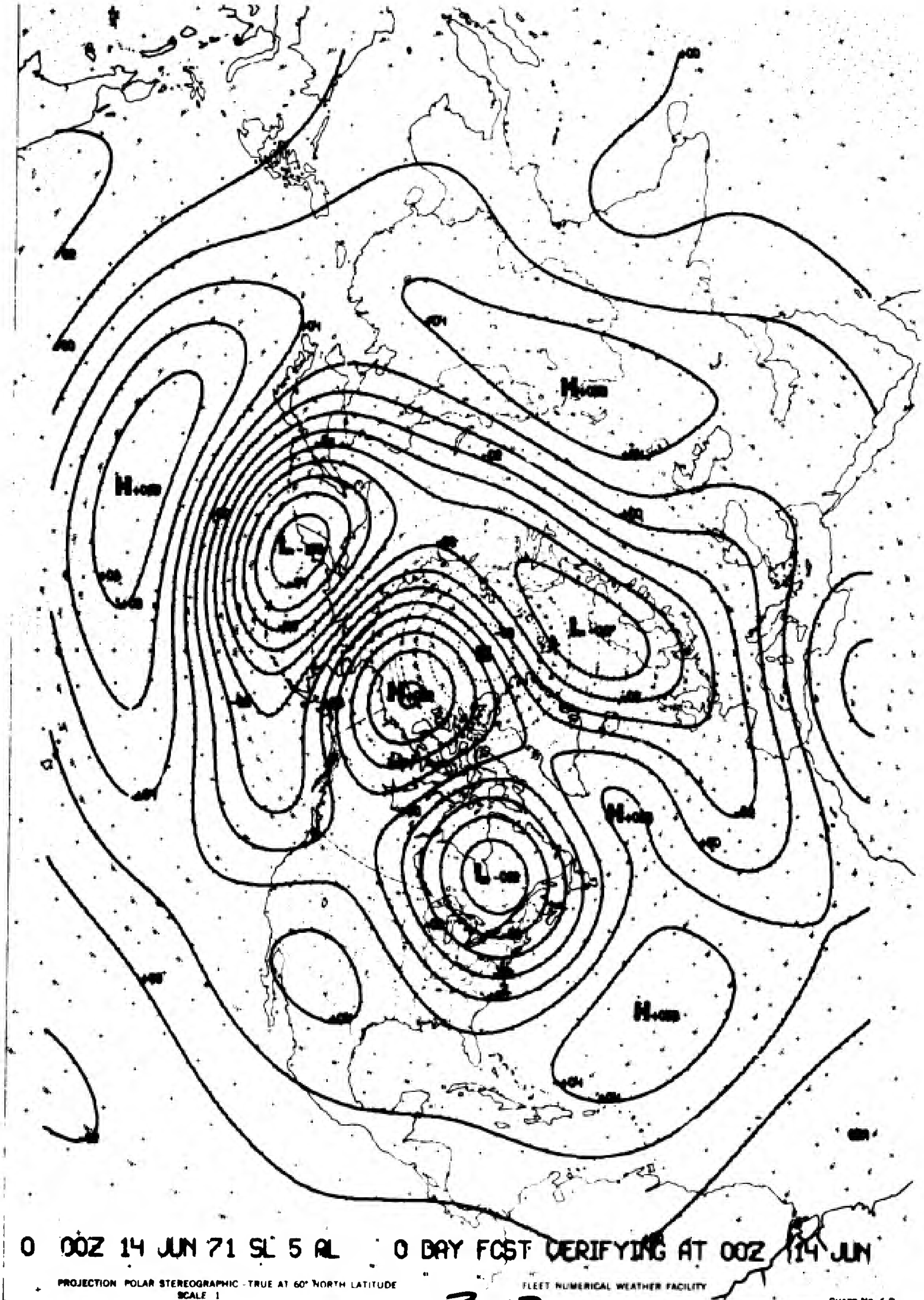


Figure 5b - Monthly Comparison of Analog Initial Match with Analog, Climatology and Persistence 1-5 day Forecast (November 1972 - May 1974)



B-13

Figure 6 - Initial Composite Analog Match for SL 500 (SL 5 AL) for 14 JUN 71

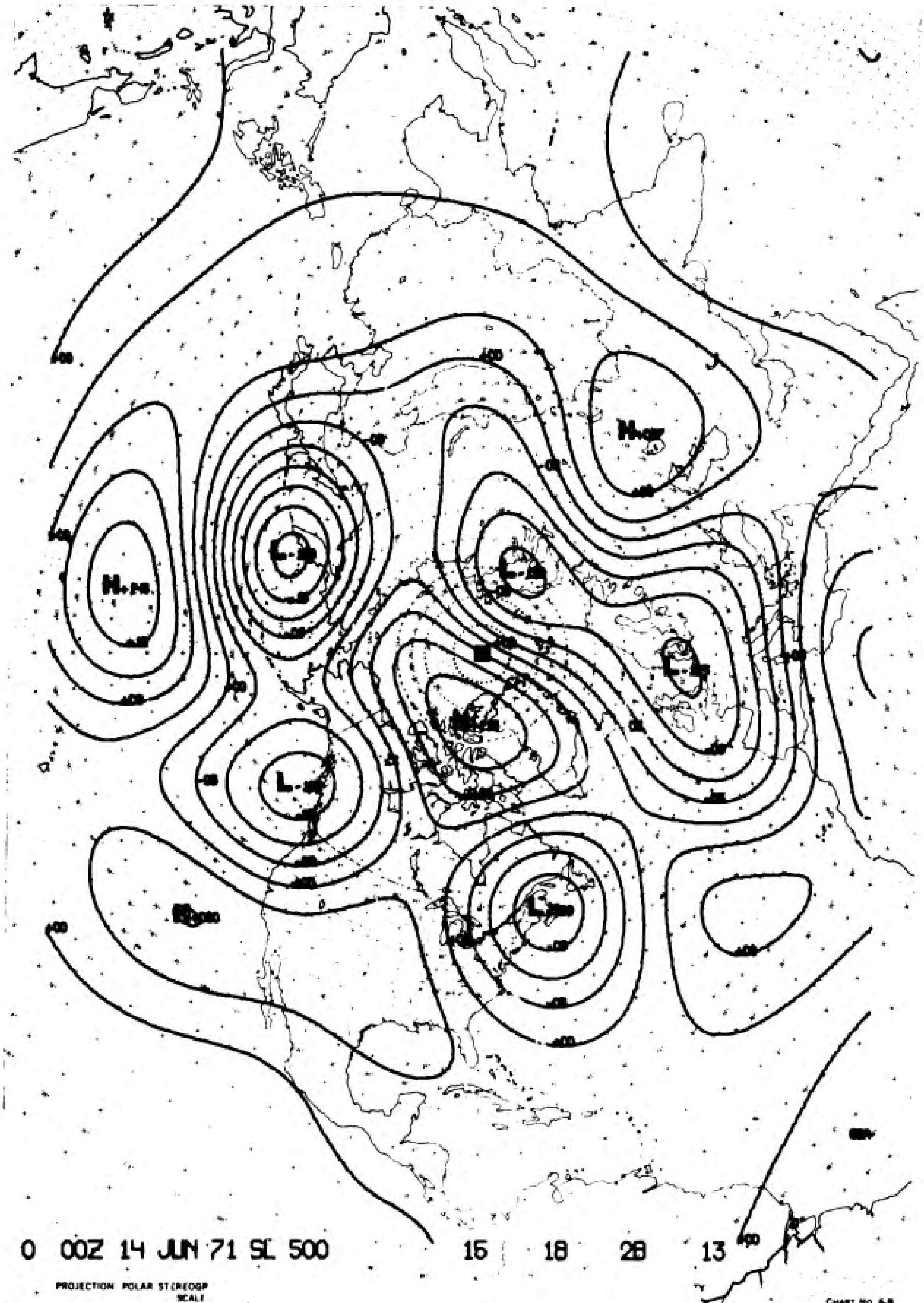


Figure 7 - SL 500 Analysis for 00Z
14 JUN 71

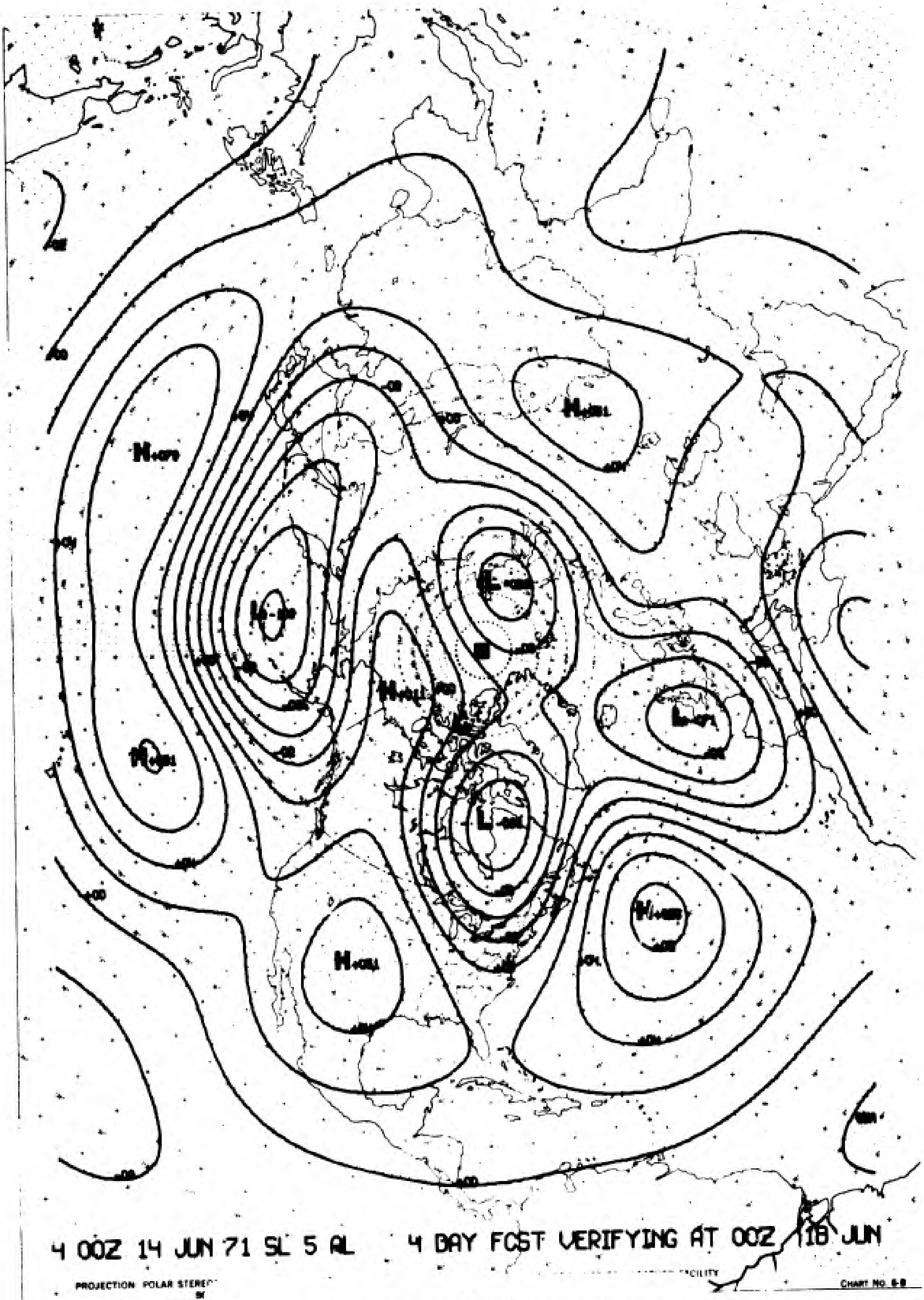


Figure 8 - 4-DAY SL 500 Analog Forecast
Verifying at 00Z 18 JUN 71

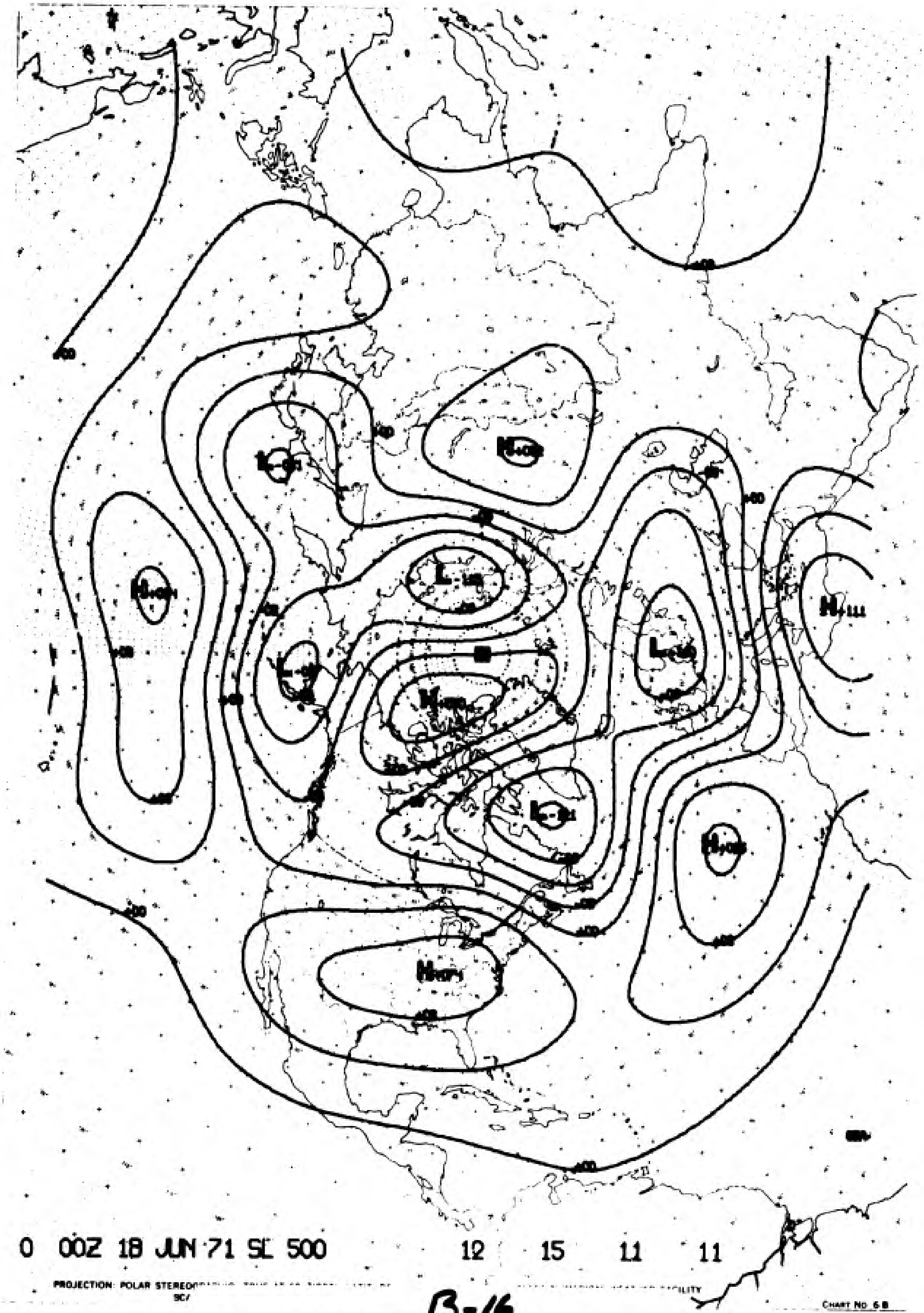
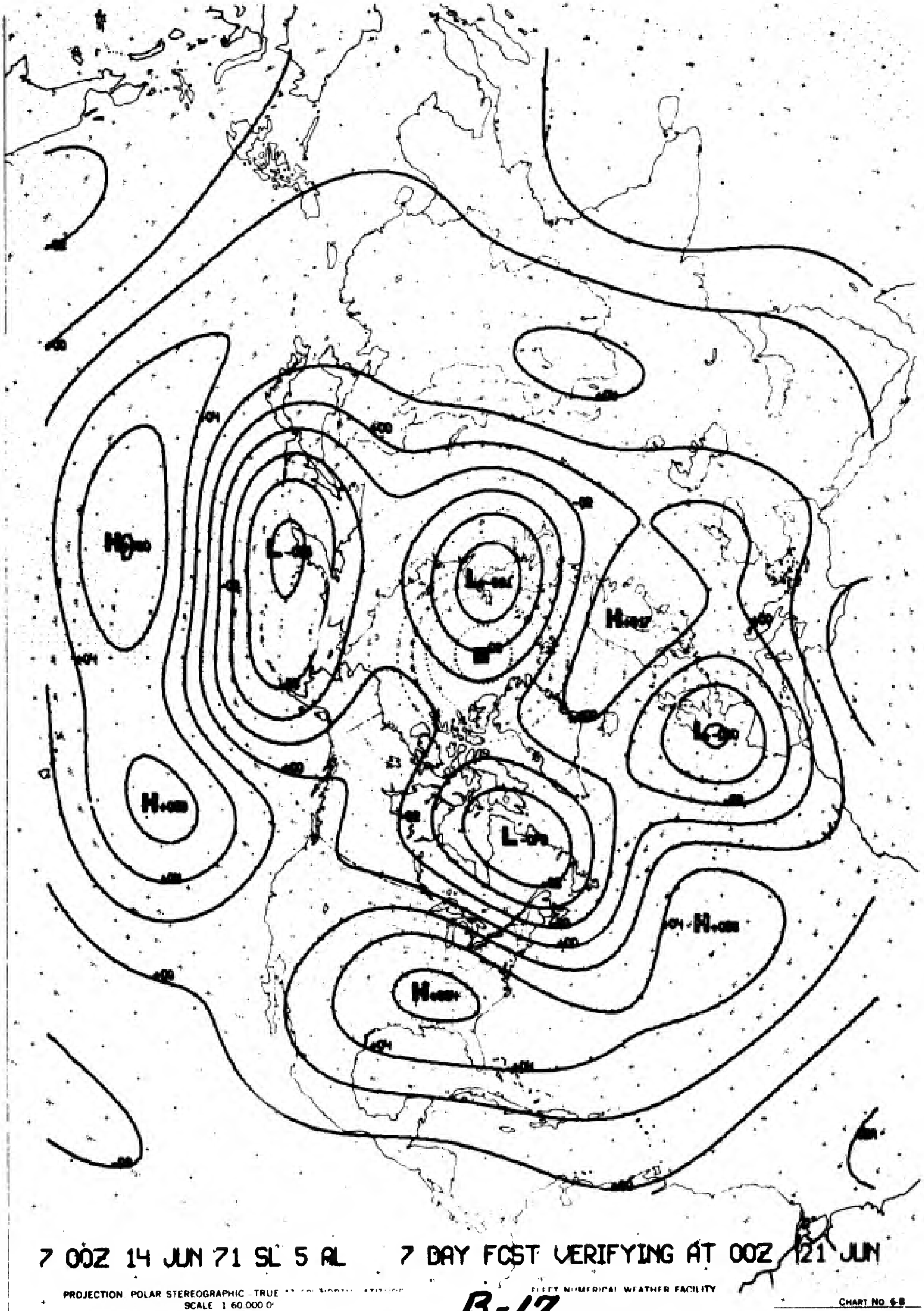
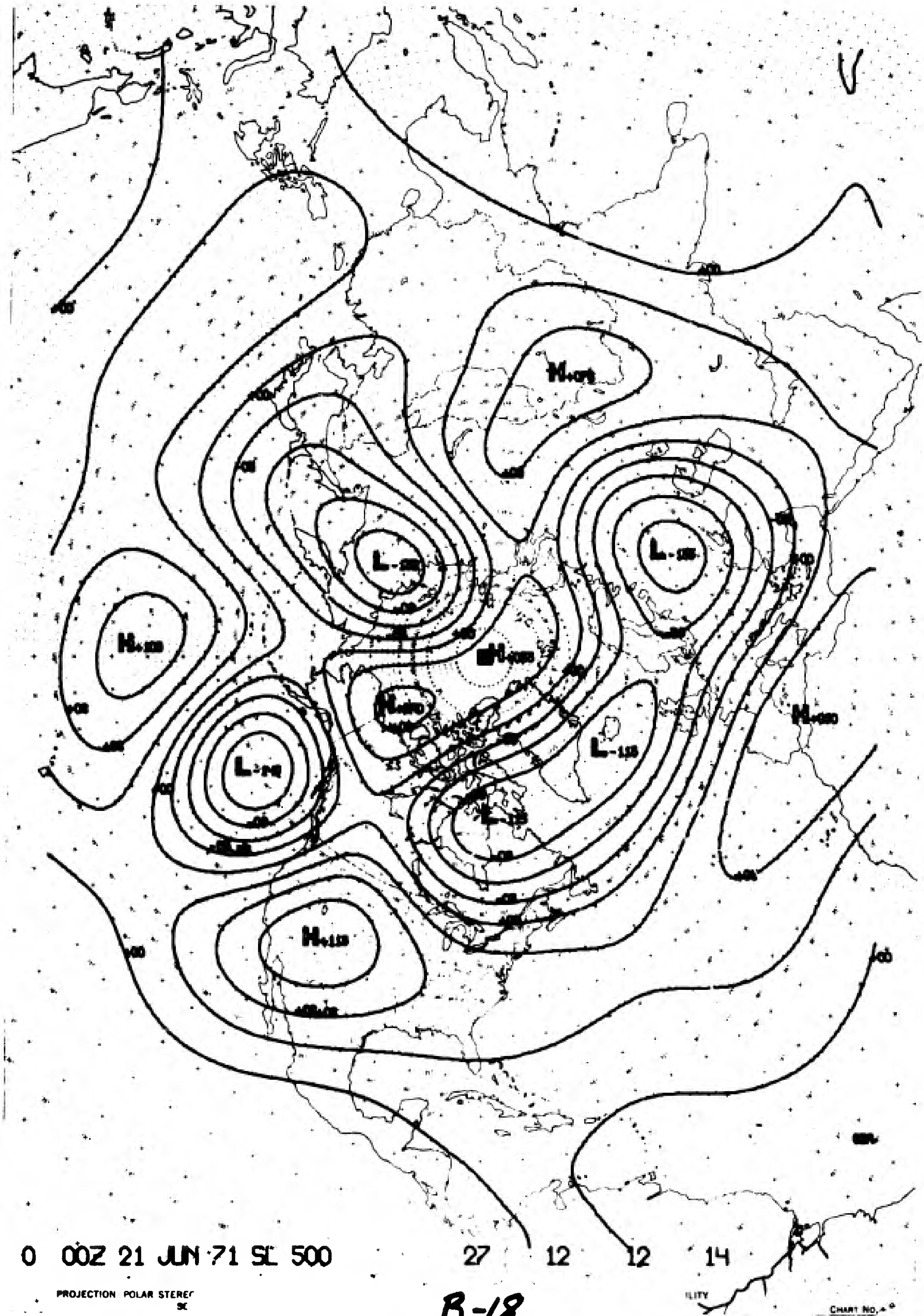


Figure 9 - SL 500 Analysis for 00Z 18 JUN 71 (Verifying Chart for Figure 8)



B-17

Figure 10 - 7-DAY SL 500 Analog Forecast Verifying at 00Z 21 JUN 71



B-18

Figure 11 - SL 500 Analysis for 00Z 12 JUN 71 (Verifying Chart for Figure 10)

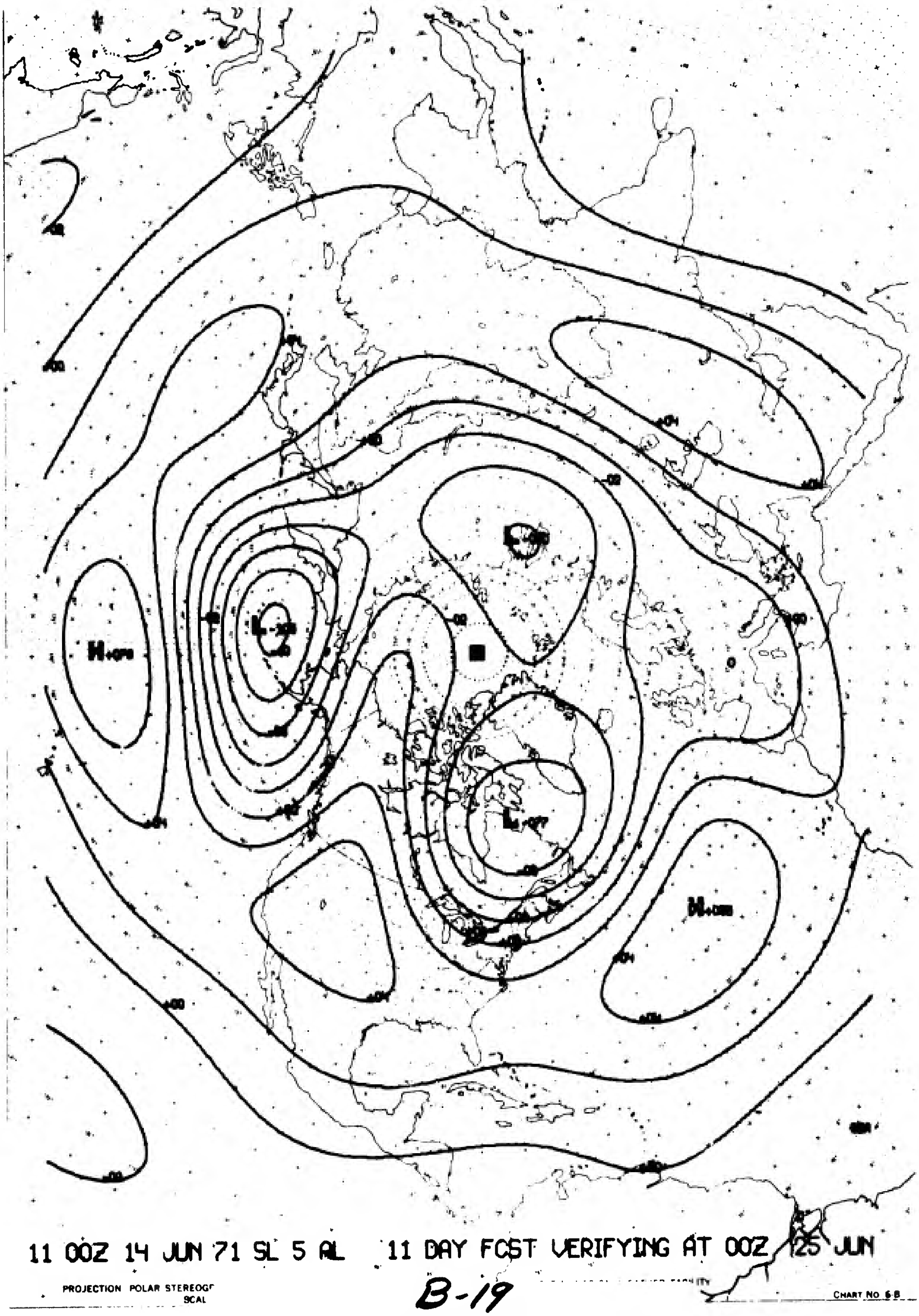
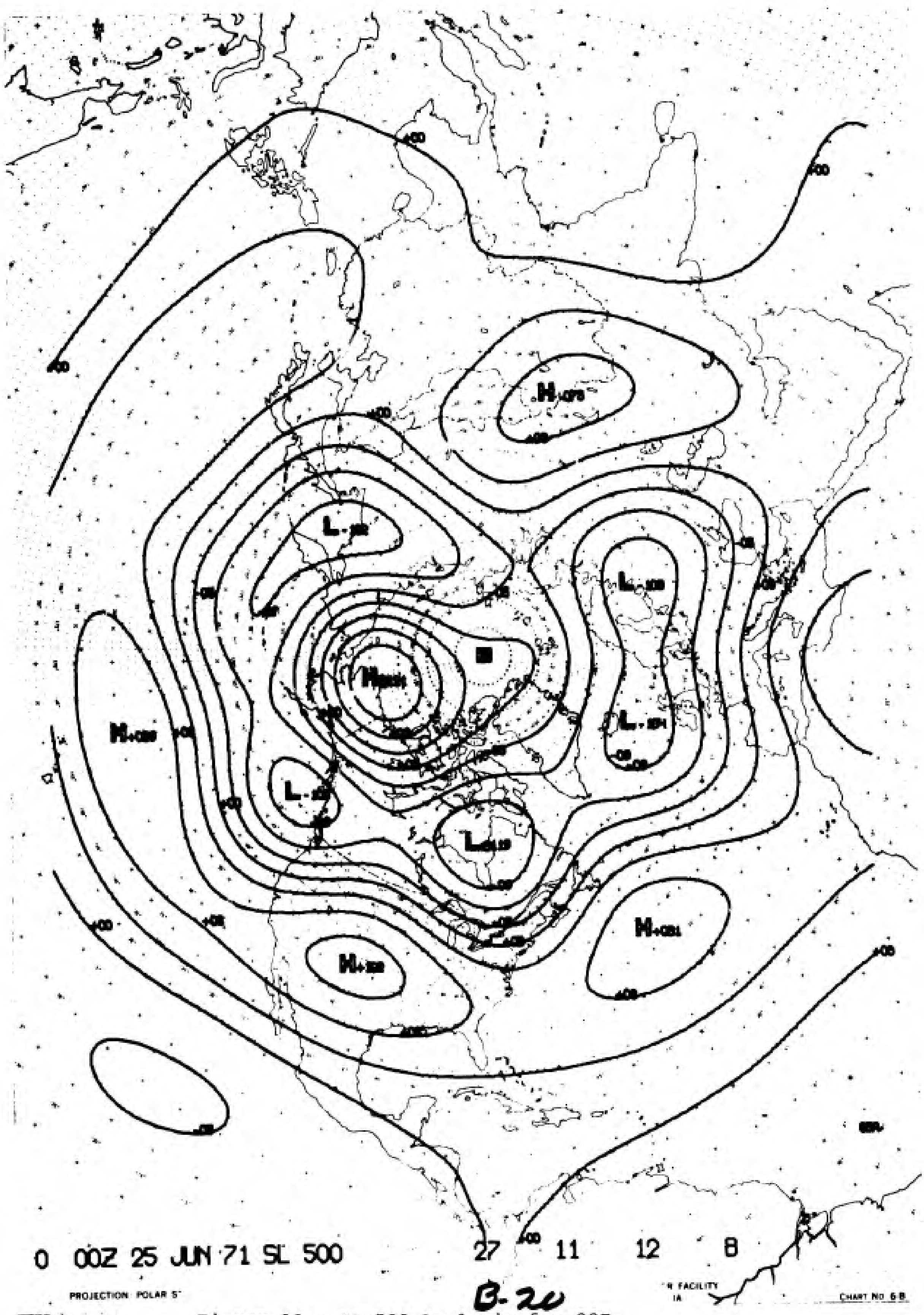


Figure 12 - 11-DAY SL 500 Analog Forecast Verifying at 00Z 25 JUN 71



B-20

Figure 13 - SL 500 Analysis for 00Z
25 JUN 71 (Verifying Chart
for Figure 12)

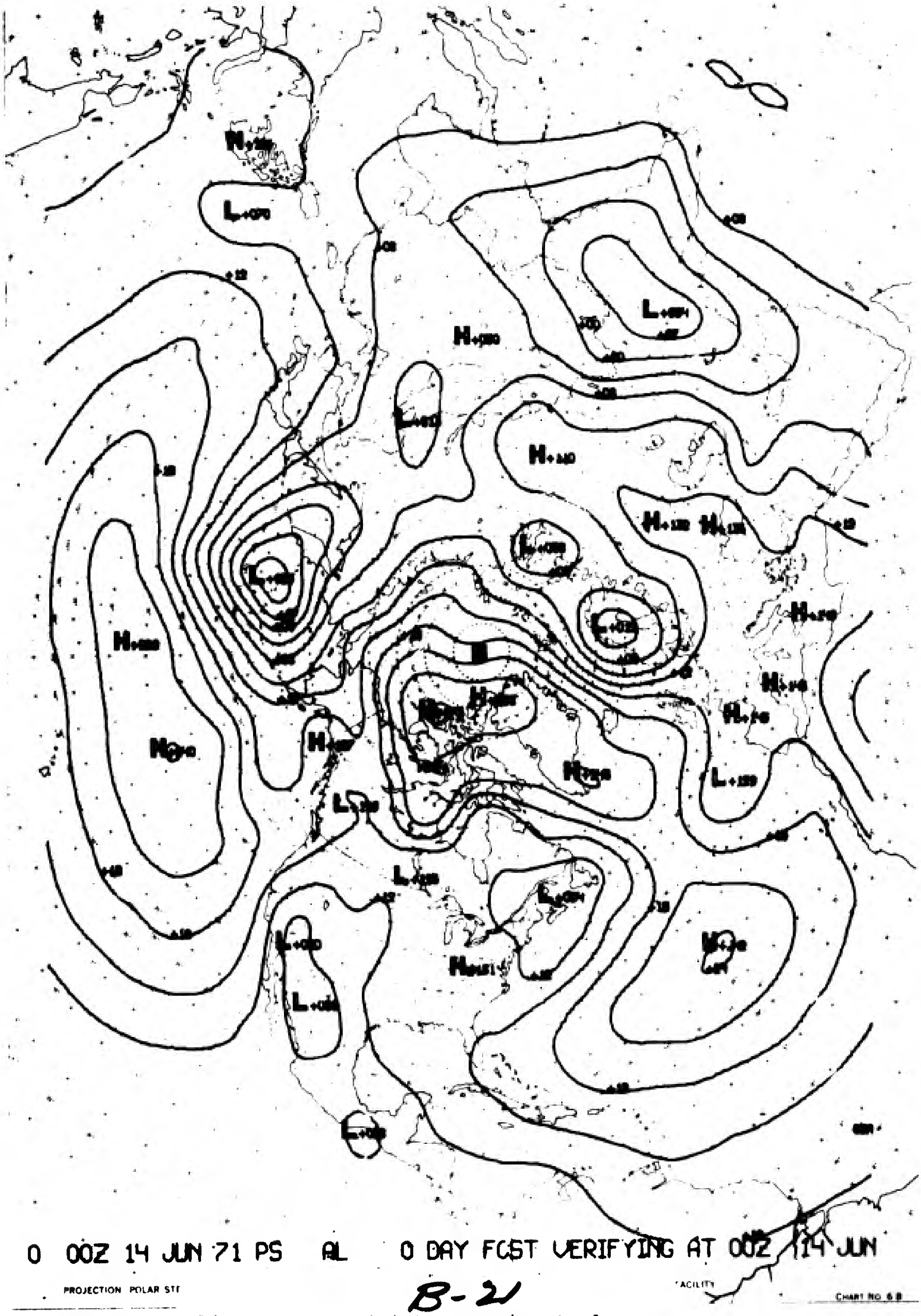


Figure 14 - Initial Composite Analog Match at the Surface (PS AL) for 00Z 14 JUN 71

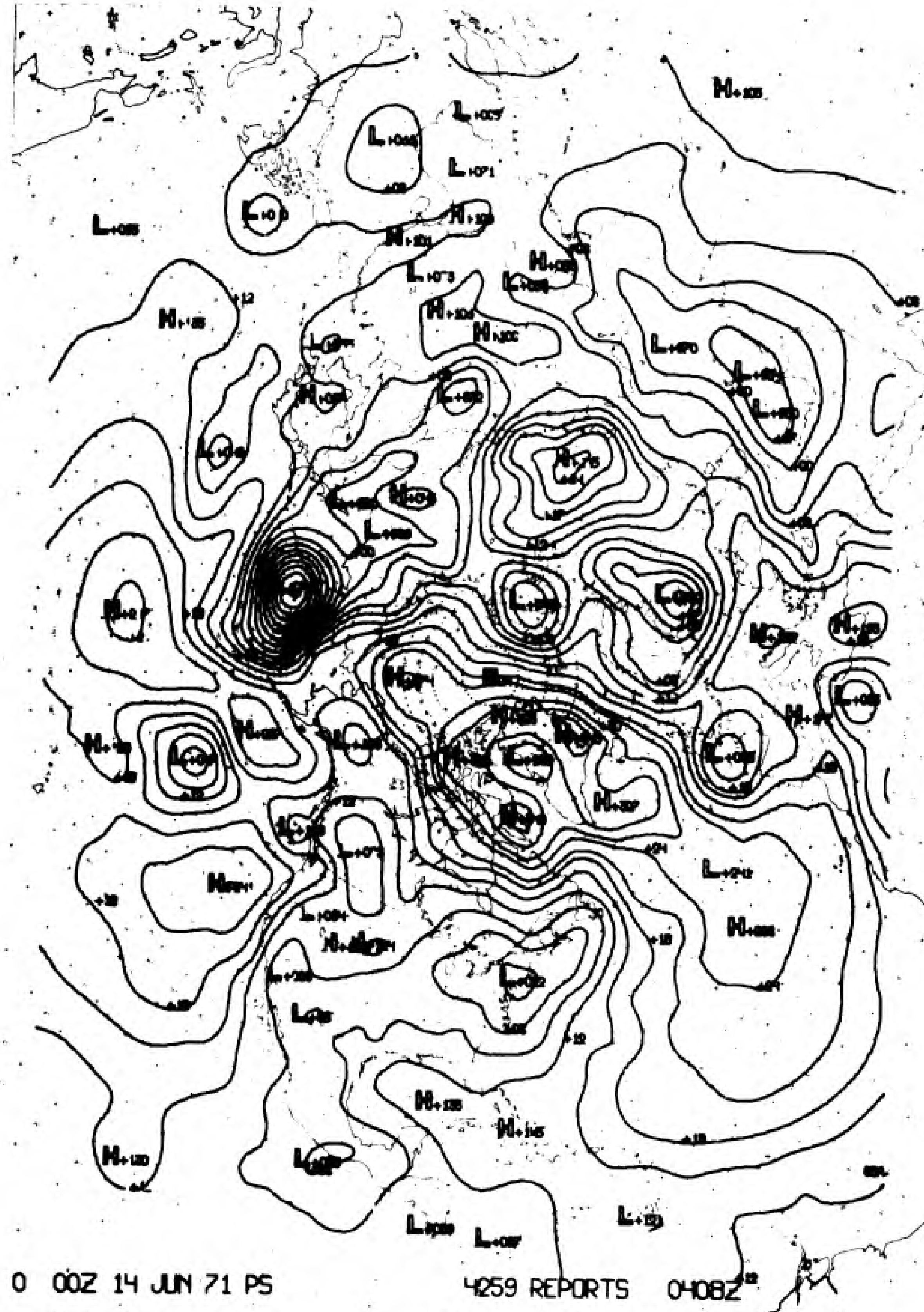
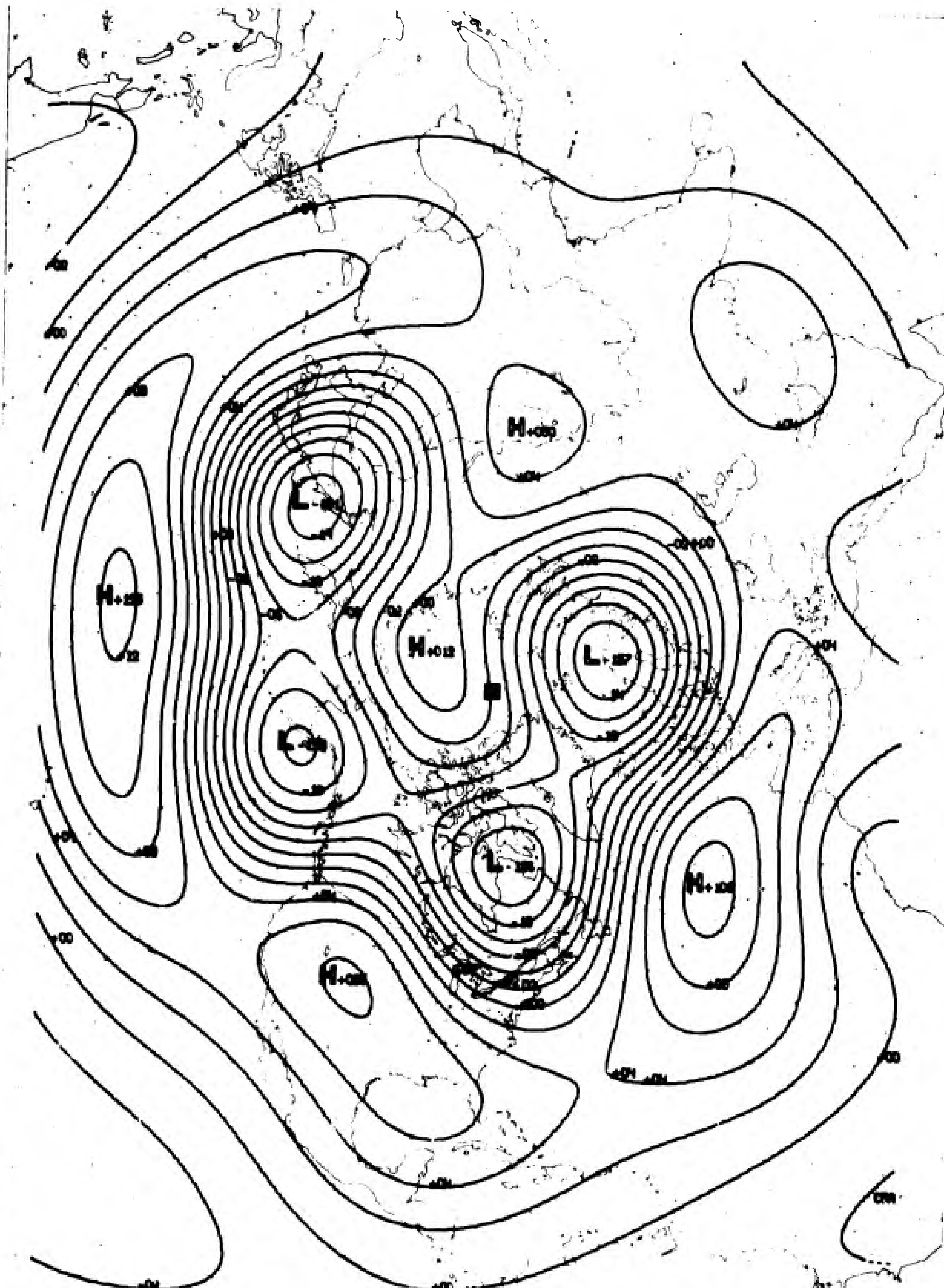


Figure 15 - Surface Analysis (PS) for
00Z 14 JUN 71



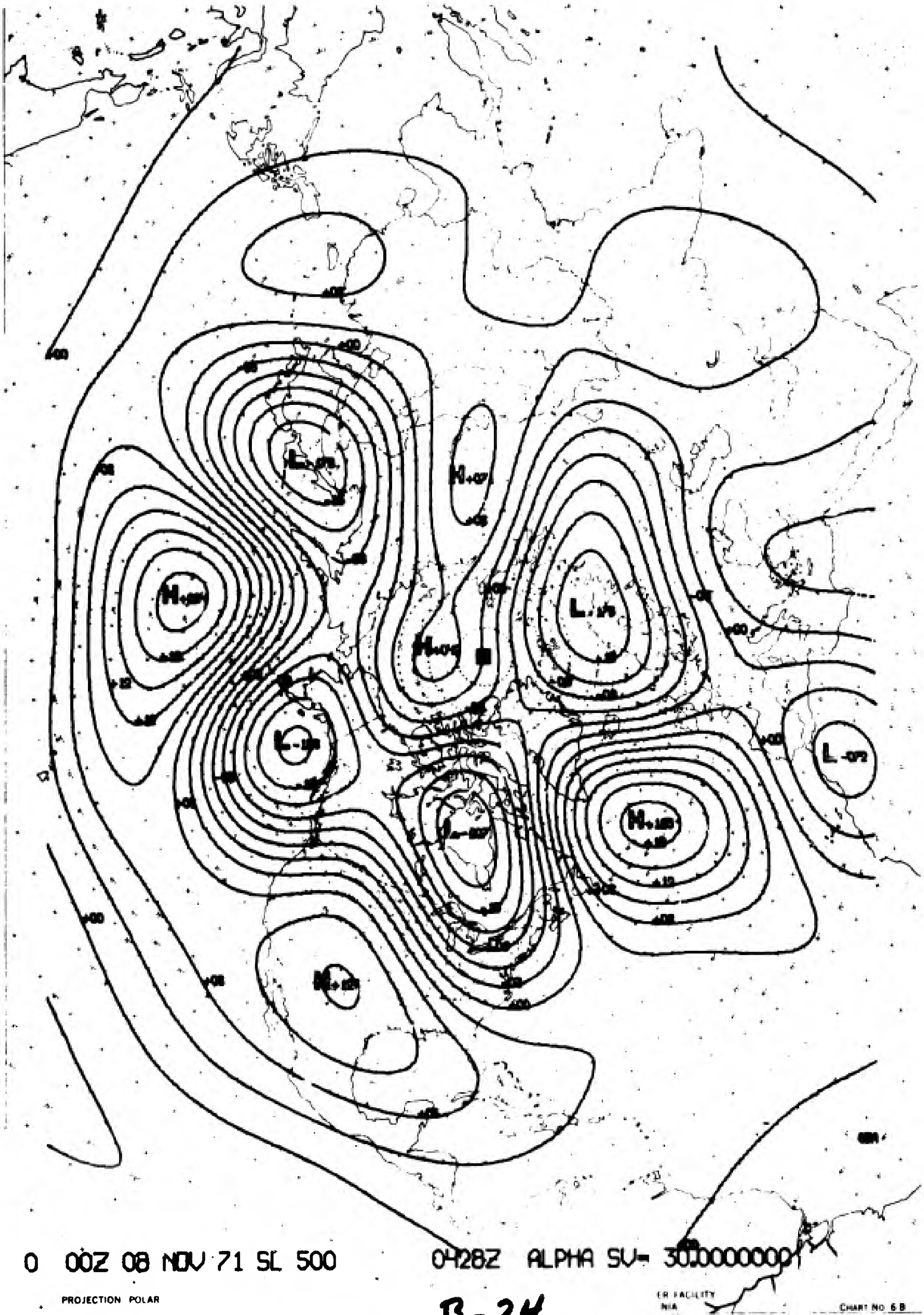
0 00Z 08 NOV 71 SL 5 AL 0 DAY FCST VERIFYING AT 00Z 08 NOV

PROJECTION POLAR STEREO

CHART NO 6 B

B-23

Figure 16 - Initial Composite Analog
for SL 500 at 00Z 8 NOV 71



B-24

Figure 17 - SL 500 Analysis for 00Z
8 NOV 71

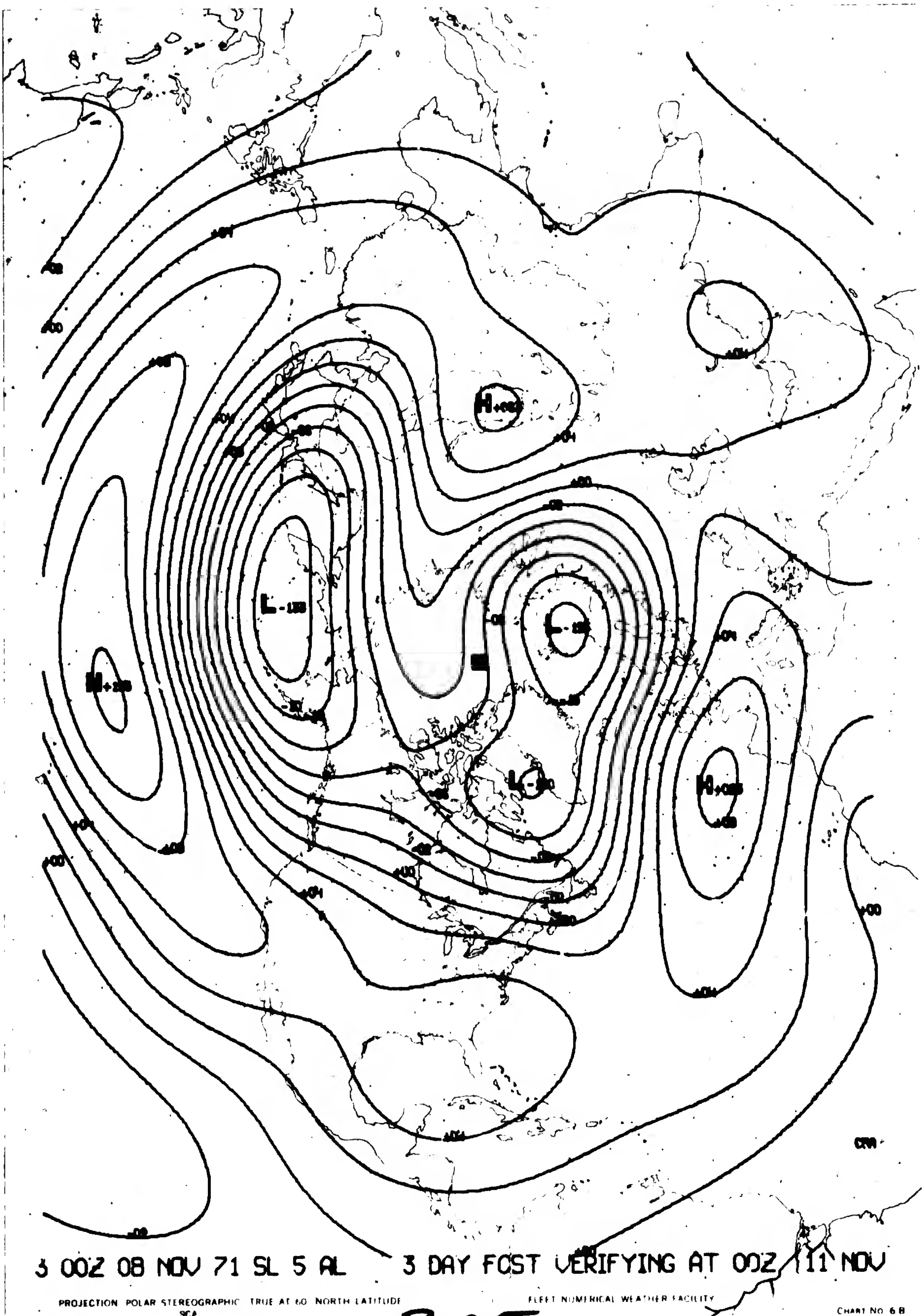
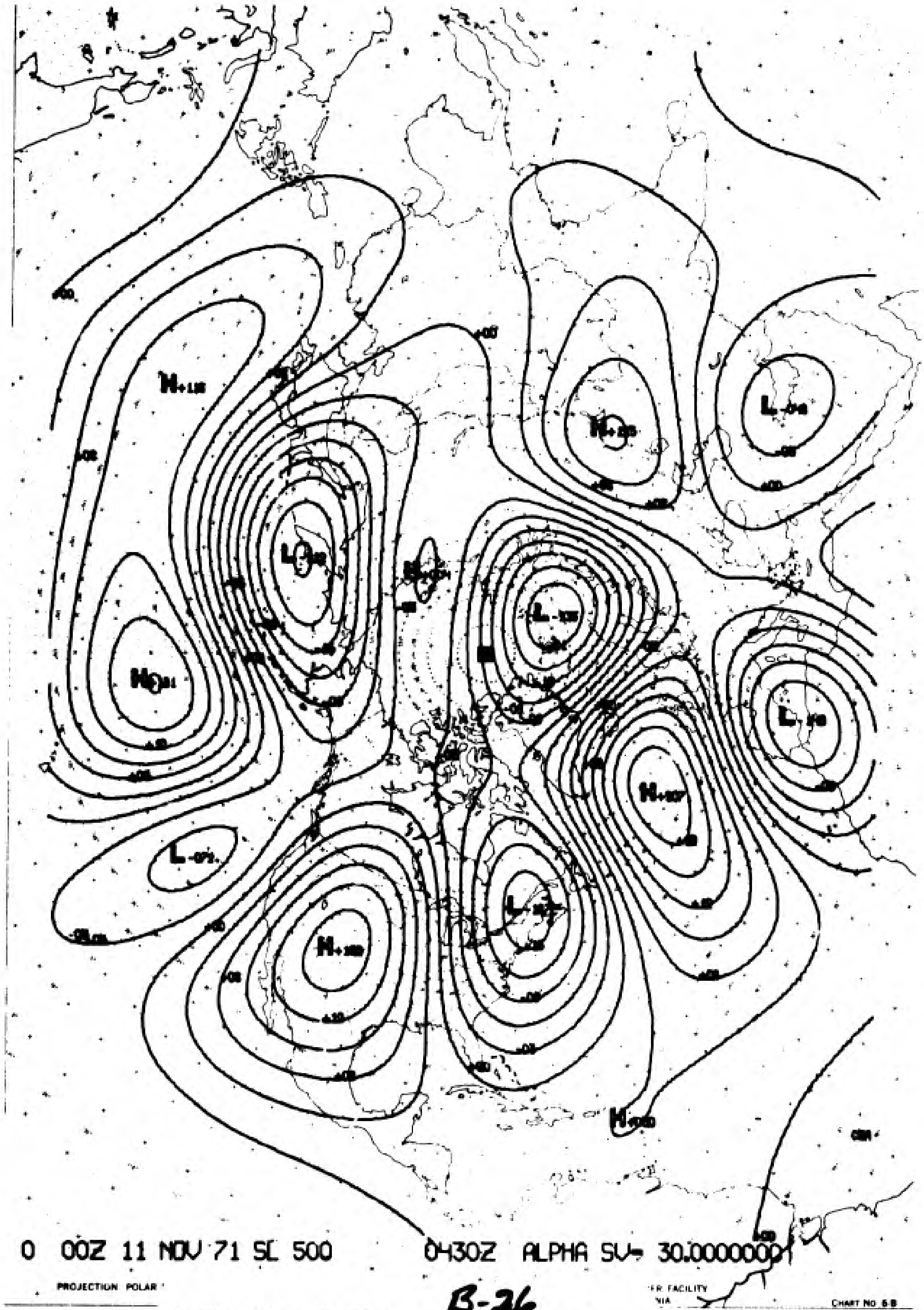


Figure 18 - 3-DAY SL 500 Analog Forecast Verifying at 00Z 11 NOV 71



B-26

Figure 19 - SL 500 Analysis for 00Z
11 NOV 71 (Verifying Chart
for Figure 18)

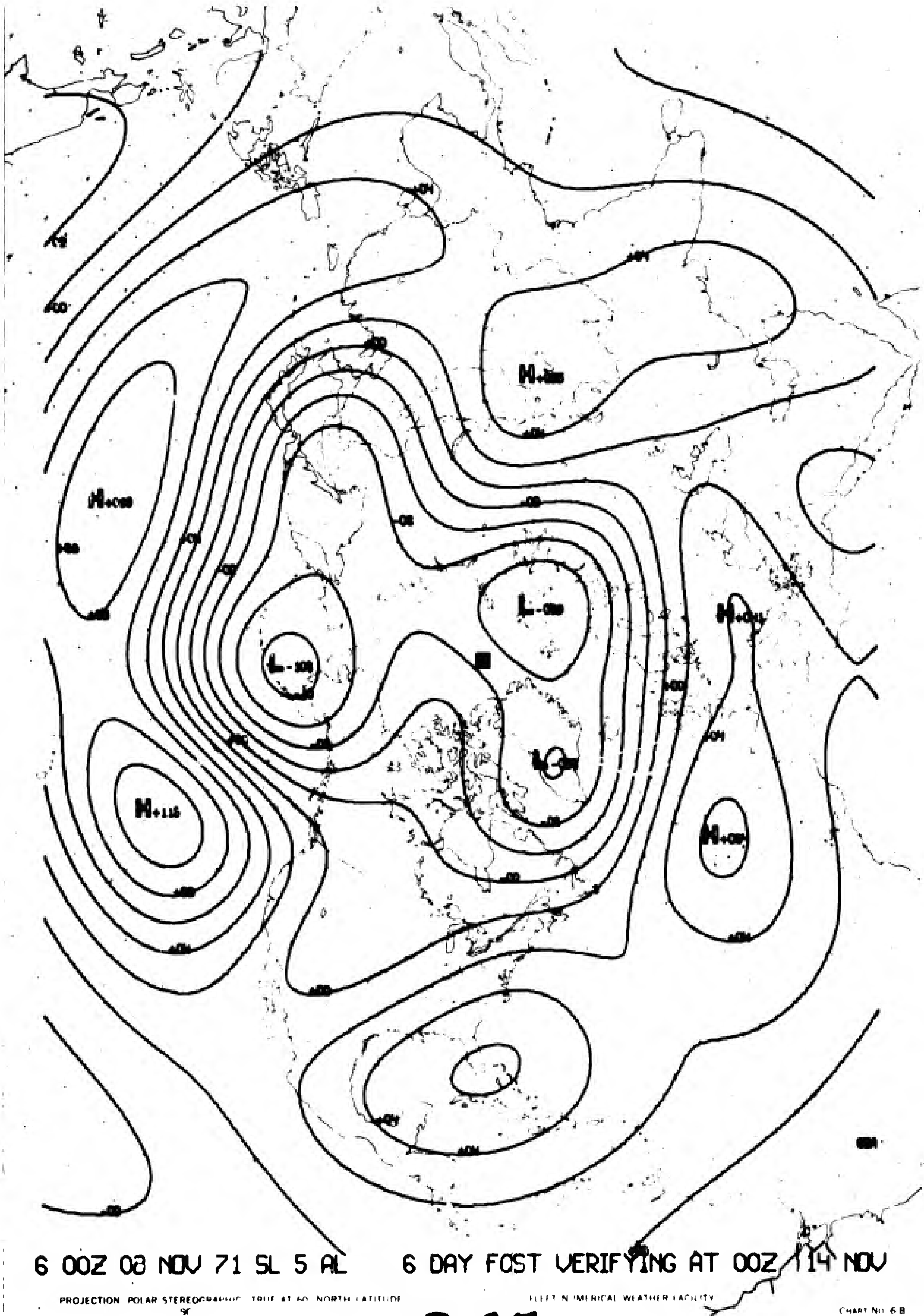


Figure 20 - 6-DAY SL 500 Analog Forecast
Verifying at 00Z 14 NOV 71

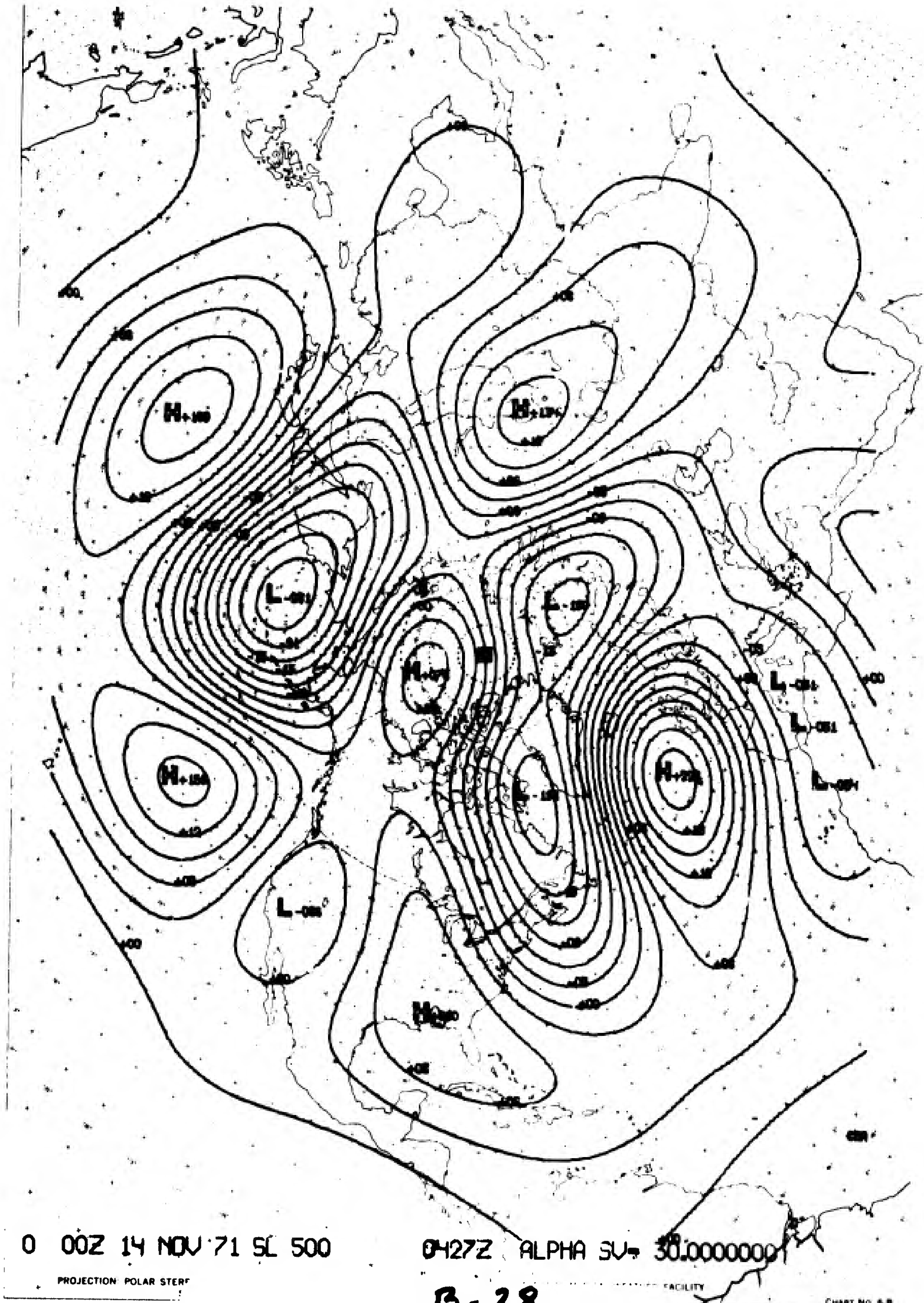
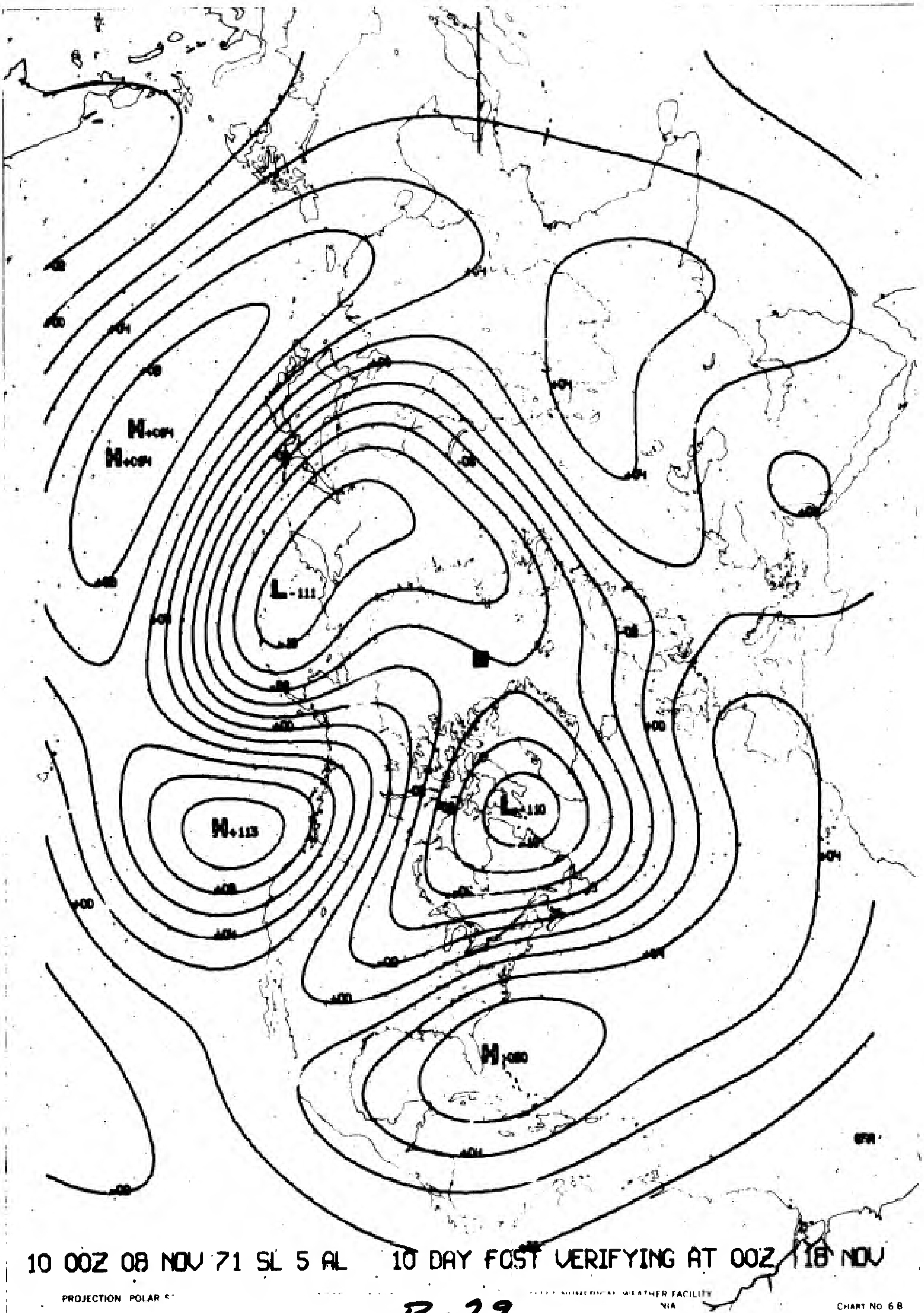
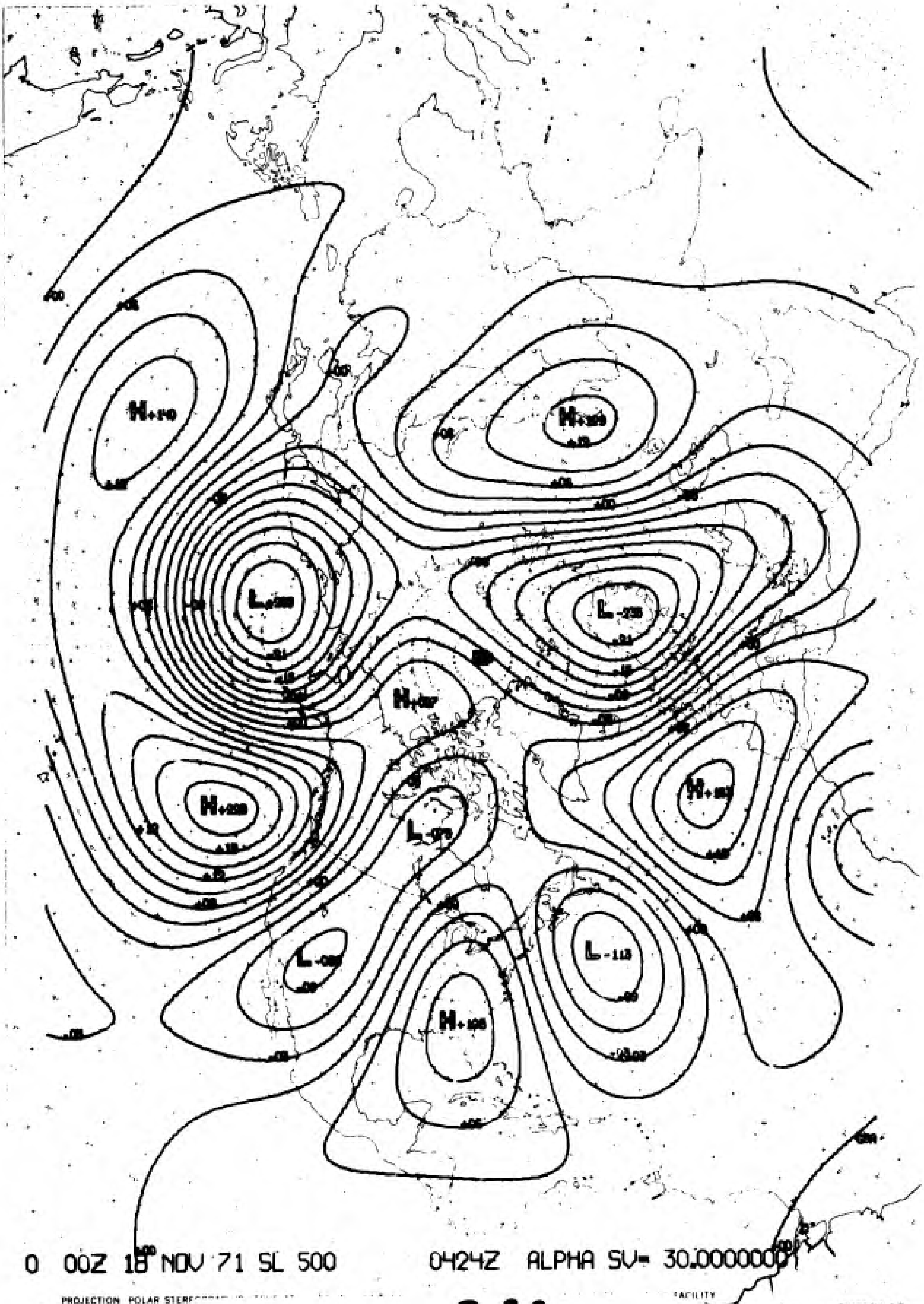


Figure 21 - SL 500 Analysis for 00Z 14 NOV 71 (Verifying Chart for Figure 20)



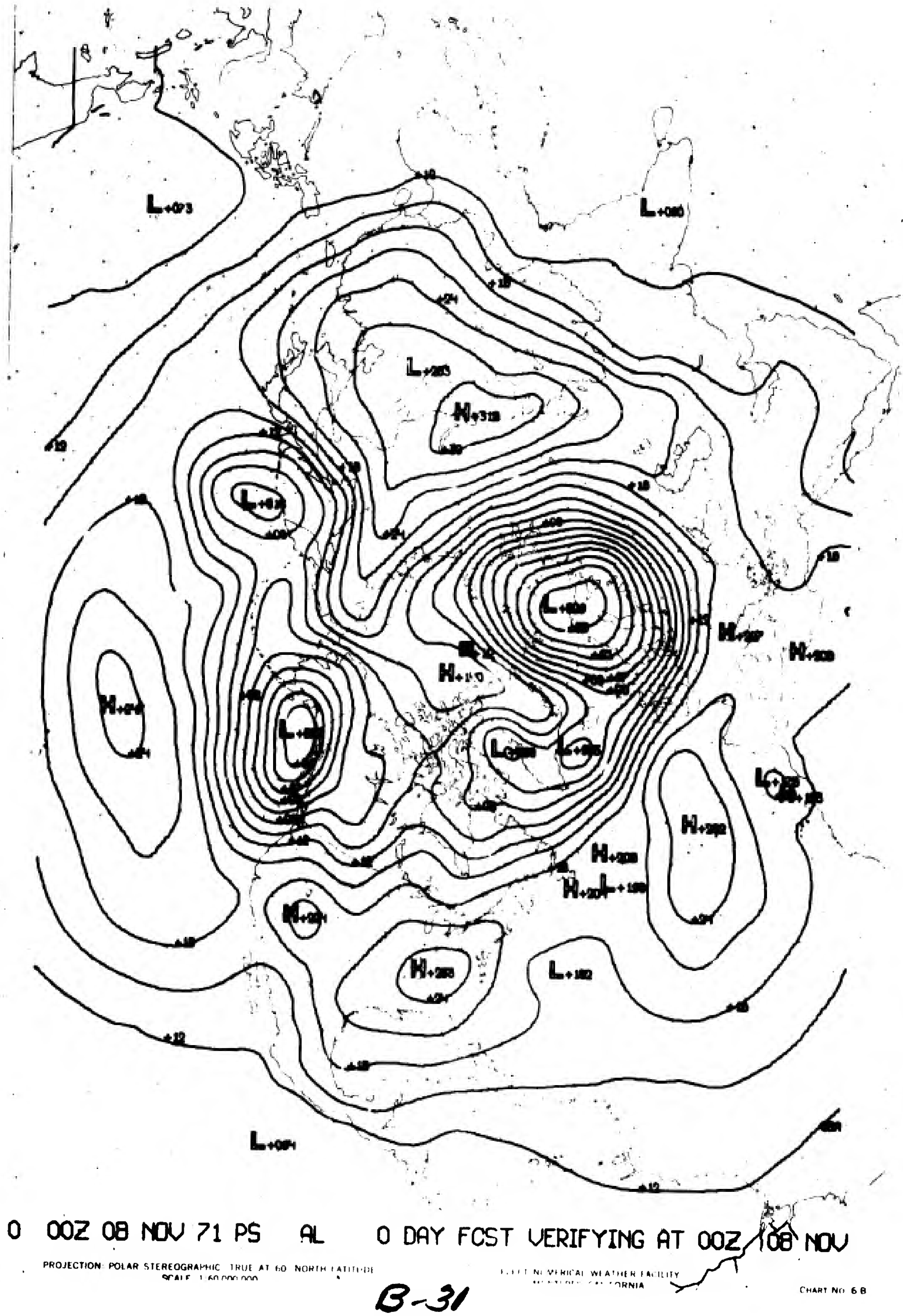
B-29

Figure 22 - 10-DAY SL 500 Analog Forecast Verifying at 00Z 18 NOV 71



B-30

Figure 23 - SL 500 Analysis for 00Z 18 NOV 71 (Verifying Chart for Figure 22)



B-31

Figure 24 - Initial Composite Analog Match at the Surface for 00Z 08 NOV 71

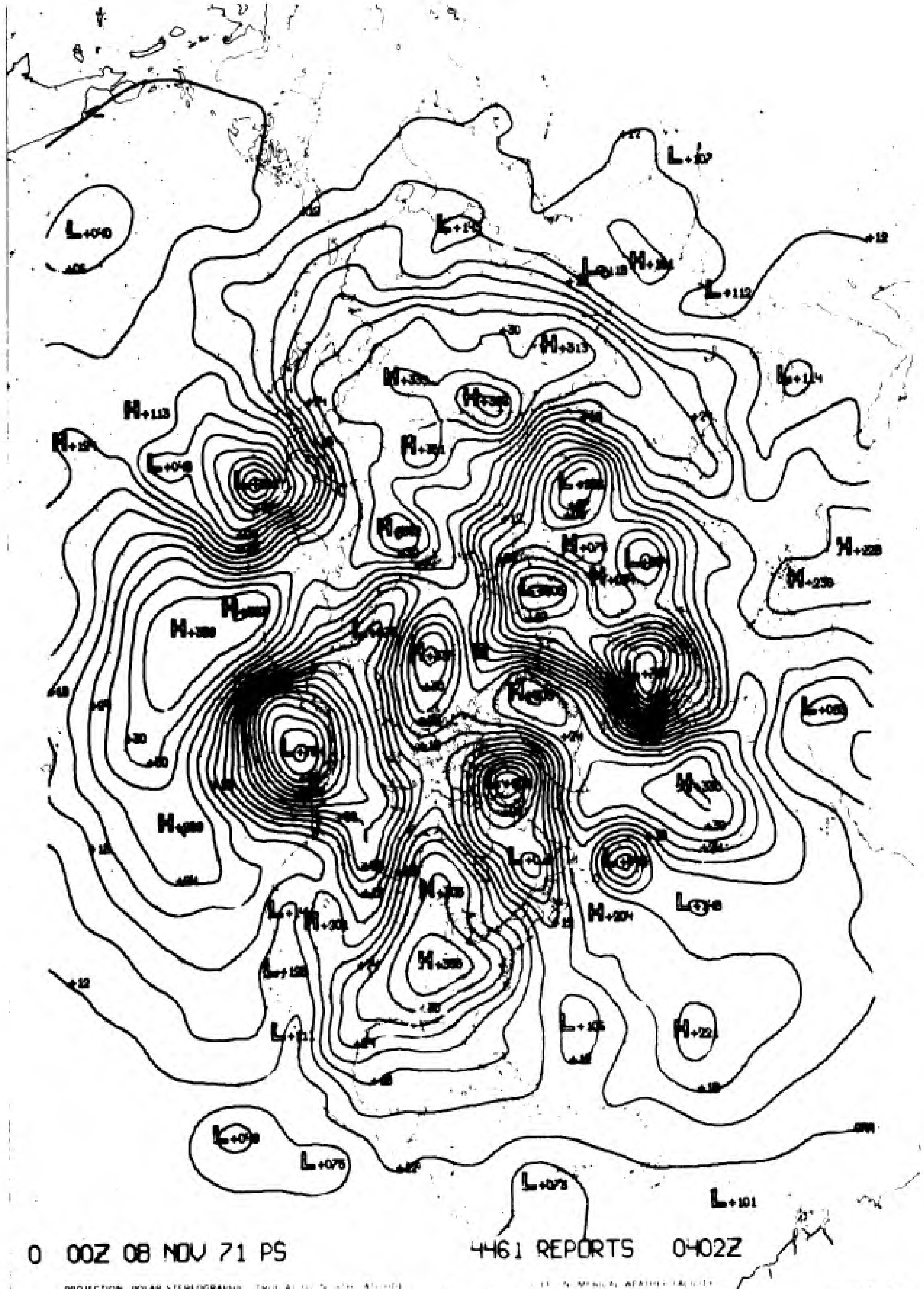


Figure 25 - Surface Analysis for 00Z 08 NOV 71