OCEAN DATA SYSTEMS, INC.

odsi

OCEAN DATA SYSTEMS, INC.
2400 GARDEN ROAD, MONTEREY, CALIFORNIA 93940 • 408/649-1133

Submitted to
Goddard Space Flight Center
Greenbelt, Maryland

ATMOSPHERIC ANALYSIS MODELING IN SUPPORT OF SEASAT

Final Technical Report

Prepared under
NASA-GSFC Contract No. NAS5-24469

November 1978

Prepared by
Rodger A. Langland
Pamela L. Stephens
OCEAN DATA SYSTEMS, INC.
Monterey, California

ABSTRACT

ODSI has developed and tested atmospheric objective analysis models for NASA in preparation for assessing the utility of SEASAT data. Of the several discretionary procedures in such computer programs, the effects of three were examined and documented: (1) the effect of varying the weights in the Pattern Conserving Technique; (2) the effect of varying the data influence region; (3) the effect of including wind information in analyses of mass-structure variables. The problem of inserting bogus reports is also examined.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

Ocean Data Systems, Inc. (ODSI) has designed, developed and tested a set of atmospheric objective analysis and prediction models of varying (grid) resolution in support of the SEASAT Program under Contract No. 954668 to the Jet Propulsion Laboratory, and under Contract No. NASW-2558 to Econ, Inc. Some advanced development of these models has been accomplished by ODSI personnel under Contract No. NAS5-24469 to the Goddard Space Flight Center. This Final Technical Report covers the modeling efforts made under NAS5-24469. (Additional advanced development is in progress under Contract NAS5-24468 to GSFC.)

The main objective of all such modeling activities was/is to prepare an appropriate model context: (1) for assessing the utility of SEASAT data in atmospheric analysis and short-range prediction; and (2) for enhancing the usefulness of such data through improved procedures in analysis and prediction models. As a practical matter, this means that one must use procedures to ingest and to distribute SEASAT information into (spatial) scales that are less likely to get "washed out" in the first few hours of a short-range forecast (the model's "adjustment period").

In this specific contractual effort, ODSI identified and tested procedures for enhancing the computational viability of (SEASAT) data. (This development is needed as a prerequisite to the development of a computationally-

economical assimilation model for asynoptic SEASAT information.) There are many discretionary procedures in objective analysis models. Three were singled out for special study herein: (1) an examination of the weights in the Pattern Conservation Technique; (2) an examination of the region of influence for observations; (3) an examination of the effects of wind information in analyses of mass-structure variables.

The analysis technique chosen for this work has been named the Pattern Conservation Technique (PCT). This method preserves specified differential properties of the first-guess field while fitting the latest observations under a system of weights. ODSI selected this scheme because it exhibits many of the desirable characteristics of a good manual analysis procedure. (A closely-related technique is used by Fleet Numerical Weather Central for analyzing the majority of its required meteorological parameters.) A complete description of the PCT technique, including relevant equations and program organization, is available in Volume II of the ODSI Final Report to JPL entitled "Atmospheric Model Development in Support of SEASAT", dated 30 September 1977. Appendices I and II contain replacement pages to Section I and II, respectively, to that Report.

As pointed out earlier, one must allow observations to influence an analysis in spatial scales which are computationally viable in the prediction model -- yet be consistent with the requirements for meteorological validity in such an analysis. In so doing, the important dependencies are:

grid resolution, the number and distribution of observations, the quality of first-guess fields, the dissipative properties of operators in the forecast model, and procedures in analysis for ingesting data to grid points and smoothing the final analysis outputs. A certain amount of "engineering" is necessary to achieve the best balance of all considerations/constraints. It is sometimes necessary to insert "bogus" reports to compensate for the lack of real reports to input non-standard information, or to overcome shortcomings in objective methods. In any case, analysis modeling in an operational context is an on-going proposition requiring daily scrutiny of model outputs in all types of meteorological scenarios.

This Final Report describes work performed on the discretionary procedures in the various analyses. Just as varying the settings for constants in a forecast model can significantly change the resulting output (as described in Volume IV of the aforementioned ODSI Final Report to JPL), "tuning" of discretionary procedures and constrants in an analysis can also alter the output fields drastically.

Sections II through V describe observed deficiencies, rationale for change, and resulting procedures. Appendices I and II contain updated pages of the original Volume II analysis model description as modified by the work in this report.* All of the development effort was expended on the 63x63 models. The concepts developed would be directly applicable to the 187x187 models, but tuning, filtering, etc. would probably be different.

---

* Appendix I replaces pages I-1 through I-26.
Appendix II replaces pages II-1 through II-17.

A substantial coding effort would be required to implement the new assembly concept across partition boundaries in the 187x187 models.

In Section II, the selection of the PCT weights for the different analyses is discussed. The effect of the quality of the first-guess fields on the selection of those weights is also considered. Section III consists of several subsections which come under the general heading of Assembly Procedures. Each subsection contains a description of a procedure used in determining the influence region of a data report. Section IV describes a supplementary assembly procedure, the introduction of bogus information into the analysis. Finally, in Section V, the use of wind observations to better define the height gradients in the upper level height analysis is examined.

II. VARYING THE WEIGHTS IN THE ANALYSIS

The objective analysis programs developed by ODSI employ the Pattern Conserving Technique (PCT), an analysis approach which, through a system of weights, attempts to preserve the differential properties of the first-guess field while incorporating the current observations. With the appropriate selection of weights, control can be exercised over which characteristics will be emphasized in the final analysis.

The testing of the analysis program sequence was carried out with a data set for 12Z, 22 April 1976, available from Fleet Numerical Weather Central (FNWC). Five atmospheric parameters were analyzed -- the sea surface temperature, sea-level pressure, and the upper air winds, temperatures and heights. Of these, the most extensive investigations were carried out with the pressure analysis, and a major portion of the discussion will focus on those tests.

As a secondary objective, ODSI sought to evaluate the effect of the quality of the first-guess field on the final analysis. In an operational context, the quality of the available first-guess field can vary considerably. It is important, then, that a program be designed to handle the problems which might arise from the first-guess field. To test the performance of the NASA-ODSI analysis program, two very different guess fields were used in conjunction with the FNWC data set.

The majority of the tests were carried out using a forecast output field from the previous twelve hours as the first-guess field. These particular forecast fields were not very skillful, and can be considered as examples of the poorer quality guess fields which might be encountered. For comparison, several runs were made using the FNWC operational analyzed fields for that time. These fields represent the best quality that could be expected for first-guess fields. Our aim was to find how the program requirements might change as a consequence of the first-guess field.

A. Background for PCT

As mentioned previously, a complete description of the concepts and implementation of the Pattern Conserving Technique (PCT) is found in Volume II of the OLSI Final Report to JPL.

To review briefly, in PCT, a functional is defined as the sum of the squares of the differences between the characteristics of the analysis and their counterparts in the first-guess field. Each term is multiplied by an assigned weight. The functional is then minimized through an application of the calculus of variations; the numerical solution to which yields a new analysis field.

To demonstrate the role of the weights, let us assume that ϕ is one property included as a term in the PCT equations. Its contribution to the functional at a gridpoint is given by:

$$\rho(\phi r - \phi g)^2_{I,J}$$

where ϕg = value of the property in the guess field

ϕr = value of the property in the resulting analysis

ρ = PCT weight assigned to property ϕ

I,J = gridpoint indices

Contributions by the other properties to the functional can be similarly expressed. All the terms are summed over the entire grid and the equations are solved for the new analysis field; i.e., for the ϕr values. The relative importance of a particular characteristic is dependent upon its weight.

Clearly, in the expression above, when the weight ρ is large, a term's contribution to the functional will be small only when the analyzed field value, ϕr, approaches the guess field value of the property, ϕg. A smaller value of ρ would allow a greater difference between ϕr and ϕg for the same total contribution to the functional. (Recall that the functional is to be minimized.) Since each property or characteristic has its own weight, its influence on the final analysis can be changed with an adjustment of the weights.

Three types of information comprise the terms in the PCT equations -- the gradients, the Laplacians and the new data. The gradients in eight directions from a gridpoint are calculated as well as the Laplacian at the gridpoint. The ϕg values for these differential properties are determined from the first-guess field and remain constant through the entire program sequence. The new data, on the other hand, is incorporated via the assembled field. On each scan, the new assembled field replaces the field from the previous scan as the constraining terms, the ϕg's, in the PCT equations.

Two major changes have been made in the PCT program described in the report to JPL. The first is a simplification of the Laplacian term in the equations. In the original programs, the information from a report was assembled only to the nearest gridpoint. From there, the information was distributed through the field by the gradient and Laplacian terms when the PCT equations were solved. To do this in

large spatial scales required not only the Laplacian term at a gridpoint, but at the four surrounding gridpoints as well. The resulting pentadiagonal matrices made the solution of the PCT equations quite time consuming. With the development of a better assembling procedure in which a report influences the assembled values at more than one gridpoint location, the four extra Laplacian terms could be removed from the PCT equations. A significant reduction in the time required to solve the equations (at least 25%) is realized in using the simpler tridiagonal form.

The second major change in PCT was to modify the weights on the basis of the local data density. That is, the PCT weights are altered according to the relative density of reports in the vicinity of a gridpoint. Where reports are closely spaced, a fairly accurate approximation of the state of the atmosphere can be made from the assembled field. In these locations, the analysis should favor the assembled field over the first-guess field. On the other hand, where reports are widely scattered, it is probably better to rely more heavily on the first-guess field properties.

The modification of the weights is done via INFOFAC, the information density factor. INFOFAC is calculated in a separate subroutine, INFODEN, which is described fully in a later section. In INFODEN, an attempt is made to determine how many reports will contribute to the assembled field value at a particular gridpoint. This constitutes the information density. INFOFAC is a measure of the density at a point relative to the densities at other locations. The

term or factor is applied simultaneously to increasing the weights on the assembled field and reducing the weights on the differential properties. The degree to which the PCT weights are altered is controlled through REDUCE, a fraction between zero and one. The smaller the value of REDUCE, the less the variation in the information density is allowed to affect the weights and, hence, the analysis. Thus, at a gridpoint:

$$ASMBLWT_{I,J} = (1. + REDUCE * INFOFAC_{I,J}) * ASMBLWT$$

$$GRADWT_{I,J} = (1. - REDUCE * INFOFAC_{I,J}) * GRADWT$$

$$PCTLAPL_{I,J} = (1. - REDUCE * INFOFAC_{I,J}) * PCTLAPL$$

where ASMBLWT = PCT weight on the assembled field

GRADWT = PCT weight on the gradients of the first-guess field

PCTLAPL = PCT weight on the Laplacian of the first-guess field

INFOFAC = information density factor at location I,J

REDUCE = fraction; $0 \leq REDUCE \leq 1$

I,J = gridpoint indices

In the following sections, the examples will be limited to those for which REDUCE was equal to either zero or 0.75. This permits a comparison of cases in which the information density does not affect the PCT weights and cases for which those weights are altered.

B. Sea-Level Pressure Analysis

The sea-level pressure analysis differs significantly from the upper air analyses in several respects. First and most importantly, the number of observations available to the analysis is much greater, often an order of magnitude greater. Where the height analysis might have five hundred observations to use, the surface pressure analysis will have five thousand. Secondly, the PCT weights for the sea-level pressure analysis are modified by the terrain gradients.

Observational reports from mountainous regions often reflect the local effects due to elevation rather than general conditions; consequently, the field gradients in these regions tend to be unreliable. For this reason, the PCT gradient weights are significantly reduced where the terrain gradients are strong. The modification of the gradient weights is given by:

$$GRADWT_{I,J} = [1. - (TERGRAD_{I,J}/TERMAX)] * GRADWT_{I,J}$$

where GRADWT = PCT gradient weight
TERGRAD = terrain gradient
TERMAX = maximum terrain gradient for the field
I,J = grid indices

The first group of tests which will be discussed is the sea-level pressure analyses which used a 12-hour forecast output field as the first guess. Figure II-1 shows a portion of this first-guess field on the Northern Hemisphere 63x63 grid. The discussion will focus on three features -- the

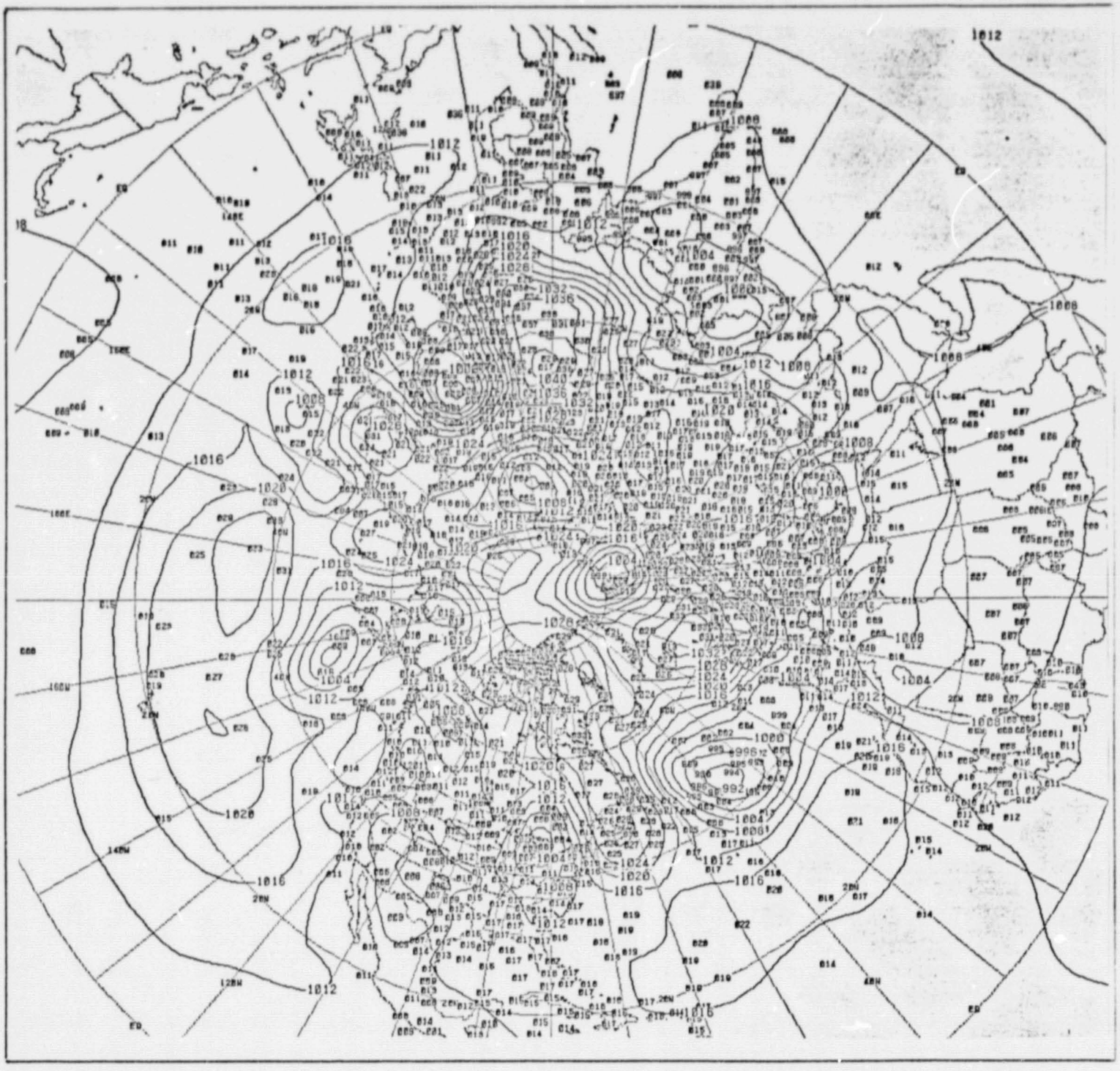


FIGURE II-1: SEA LEVEL PRESSURE ANALYSIS, FIRST-GUESS FIELD FROM FORECAST OUTPUT.

developing cyclone in the North Atlantic, the broad anticyclone in the North Pacific, and a smaller cyclone, also in the North Pacific, which is not readily apparent in this first-guess field. This second cyclone appears in the analyses at 40°N between 160°E and 170°E. A careful examination of the pressure reports shows that the forecast underestimated the strength of the major synoptic systems. For instance, the two 1033mb reports between 170°E and 180°E and near 37°N suggest that a 1032mb contour should appear in the North Pacific anticyclone, but the largest contour shown is 1024mb. In the North Atlantic cyclone, a report of 986mb to the west of the center of the 992mb contour again indicates that the cyclone is much deeper than predicted.

The analyzed fields which are obtained under different PCT weights using this guess field vary significantly. The next four figures demonstrate the effect of varying the PCT weight on the assembled field. The PCT weights for the gradient and Laplacian terms were held constant (see Table II-1) as the assembled field weight was increased by a factor of ten on each succeeding run, from one to one thousand. (Note that the information density was allowed to modify the PCT weights; REDUCE is set to 0.75.)

When the weight on the assembled field was small (1.0), the analysis (Figure II-2) shows little change from the first-guess pressure field. A large number of observations were rejected. (In the figures, a rejected observation is enclosed in a black box. The total number of rejected reports in each case is included in Table II-1.) The analysis

TABLE II-1: SEA LEVEL PRESSURE ANALYSIS STATISTICS FOR SELECTED RUNS - FORECAST FIELD AS FIRST GUESS.

ENTRY	ASSEMBLED FIELD WEIGHT	GRADIENT WEIGHT	LAPLACIAN WEIGHT	REDUCE	RMS (MB)	ITERATIONS	REJECTED OF 4706
1	1.0	10.0	2.5	0.75	1.43	39	439
2	10.0	10.0	2.5	0.75	1.19	12	229
3	100.0	10.0	2.5	0.75	0.88	7	185
4	1000.0	10.0	2.5	0.75	0.81	7	177
5	100.0	10.0	2.5	0.00	1.11	8	205
6	10.0	10.0	2.5	0.00	1.40	15	367
7	100.0	10.0	25.0	0.75	1.02	10	191
8	100.0	10.0	25.0	0.00	1.29	12	262
9	100.0	10.0	250.0	0.75	1.32	47	291
10	10.0	100.0	2.5	0.75	1.41	56	411
11	10.0	1.0	2.5	0.75	1.02	8	191
12	--	--	--	--	0.81	--	177

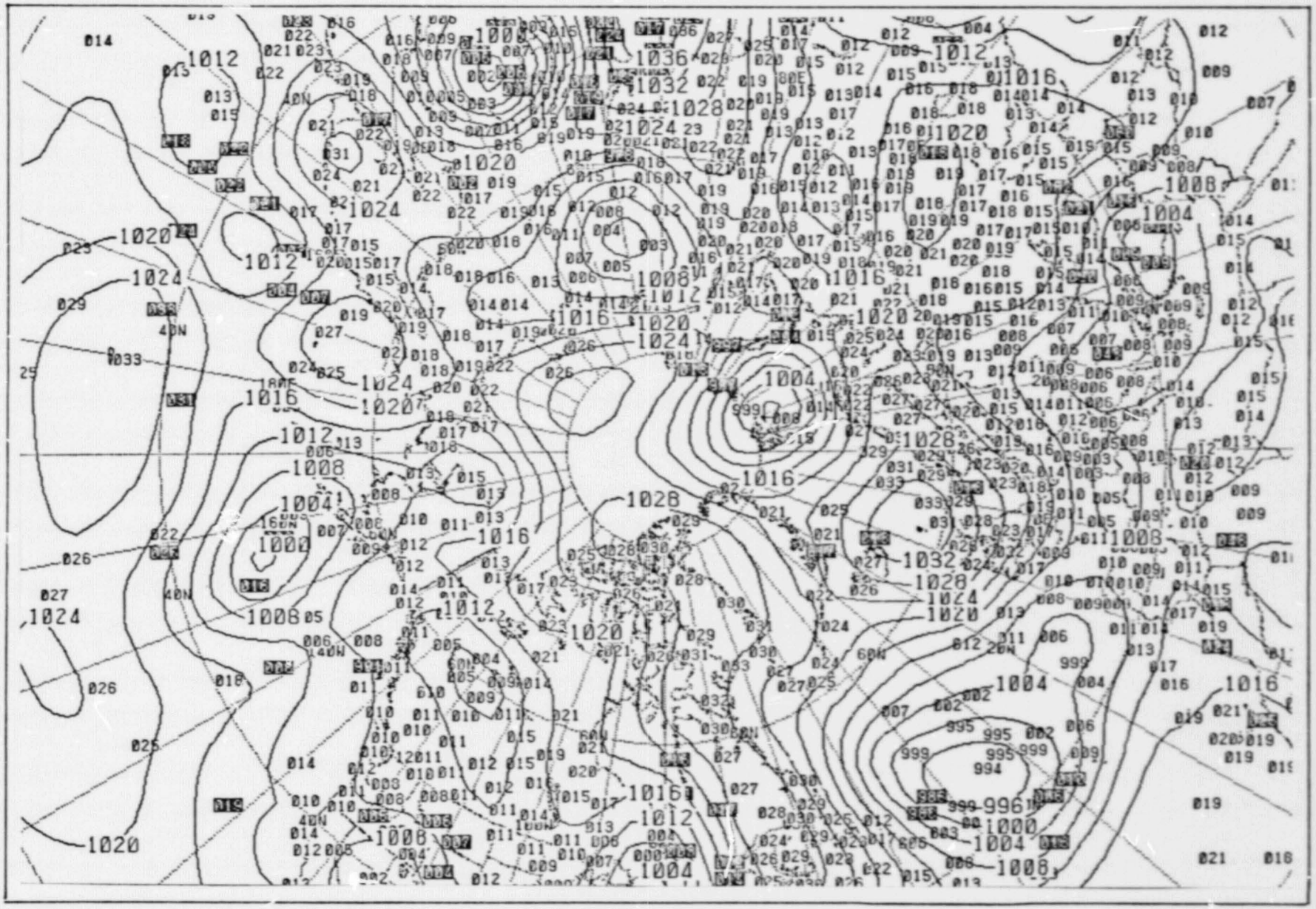


FIGURE II-2: SEA LEVEL PRESSURE ANALYSIS USING FORECAST FIELD
ASMBLWT=1., REDUCE=0.75, GRADWT=10., PCTLAPL=2.5

better accommodates the observation when the PCT assembled field weight is increased to ten. Note that the central contour in the Pacific high pressure center has reached 1028mb (Figure II-3). In the same analysis, a 1004mb contour is now included in the low pressure center over the western United States, as suggested by the observations. With a further increase in the assembled field weight (to one hundred) a 1032mb contour appears in the Pacific high, seen in Figure II-4. Other important changes include the analysis of a cyclone at approximately 46°N and 162°W. While the cyclone in the North Atlantic is still not as deep as might be expected, the analysis is at least including a 992mb contour where the lowest value contour in the other two analyses was 996mb. The last figure in this group, Figure II-5, shows the analysis when the PCT assembled field weight is one thousand. The analysis has come even closer to the observations. The 1032mb contour of the Pacific high covers a wider area. The Pacific low just to the northwest of it has been deepened to include a 1004mb contour. In the Atlantic cyclone, the 992mb contour encloses a larger area which includes the 986mb observation.

In addition to looking at the plotted analyzed fields, one can appreciate the effect of varying the weights by noting some of the program statistics which have been included in Table II-1. These first four cases are given as the first entries in Table II-1. Three quantities can be compared from one run to the next -- the root mean square (RMS) departure, the number of iterations required to solve the

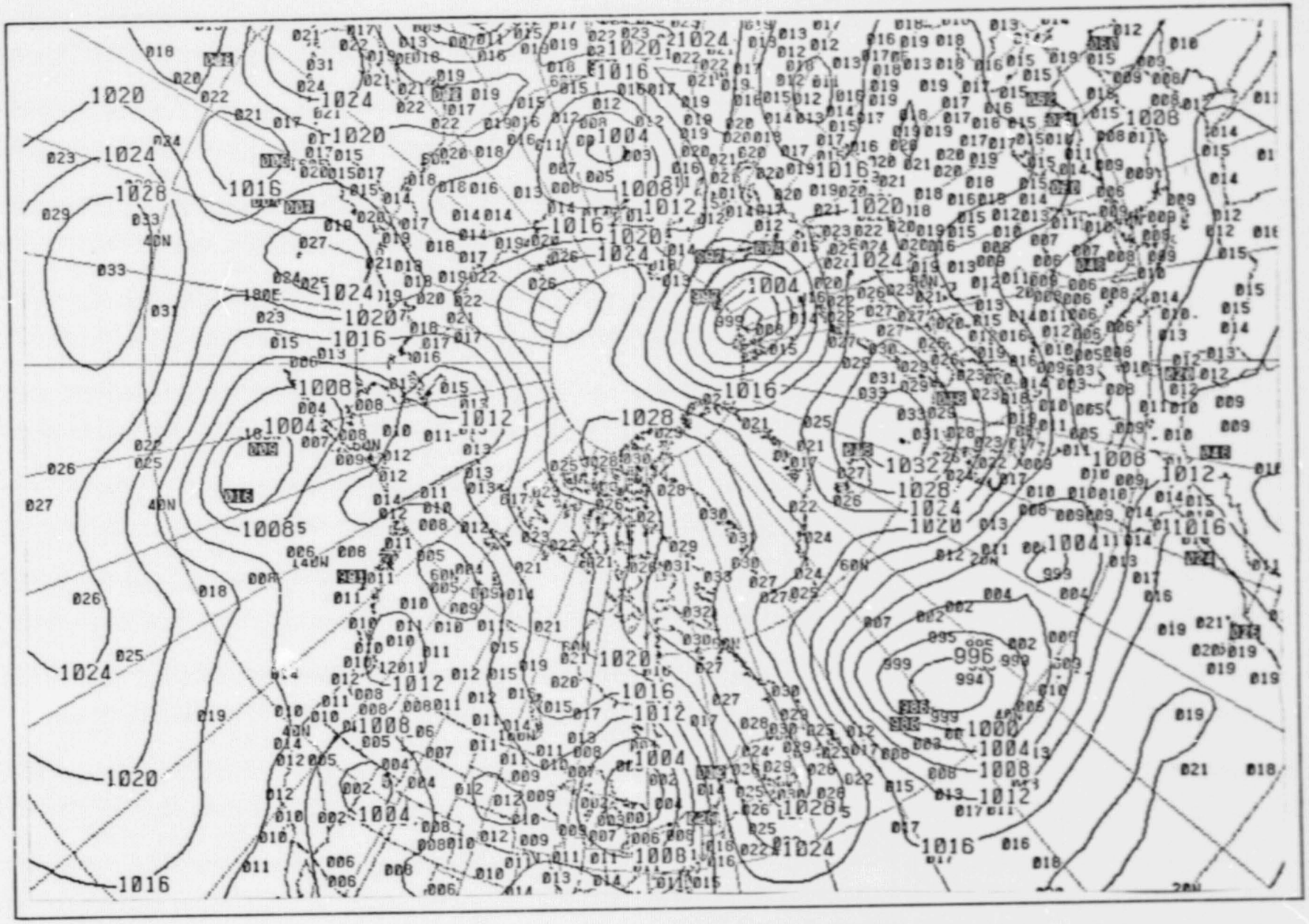


FIGURE II-3: SEA LEVEL PRESSURE ANALYSIS USING FORECAST FIELD
ASMBLWT=10., REDUCE=0.75, GRADWT=10., PCTLAPL=2.5

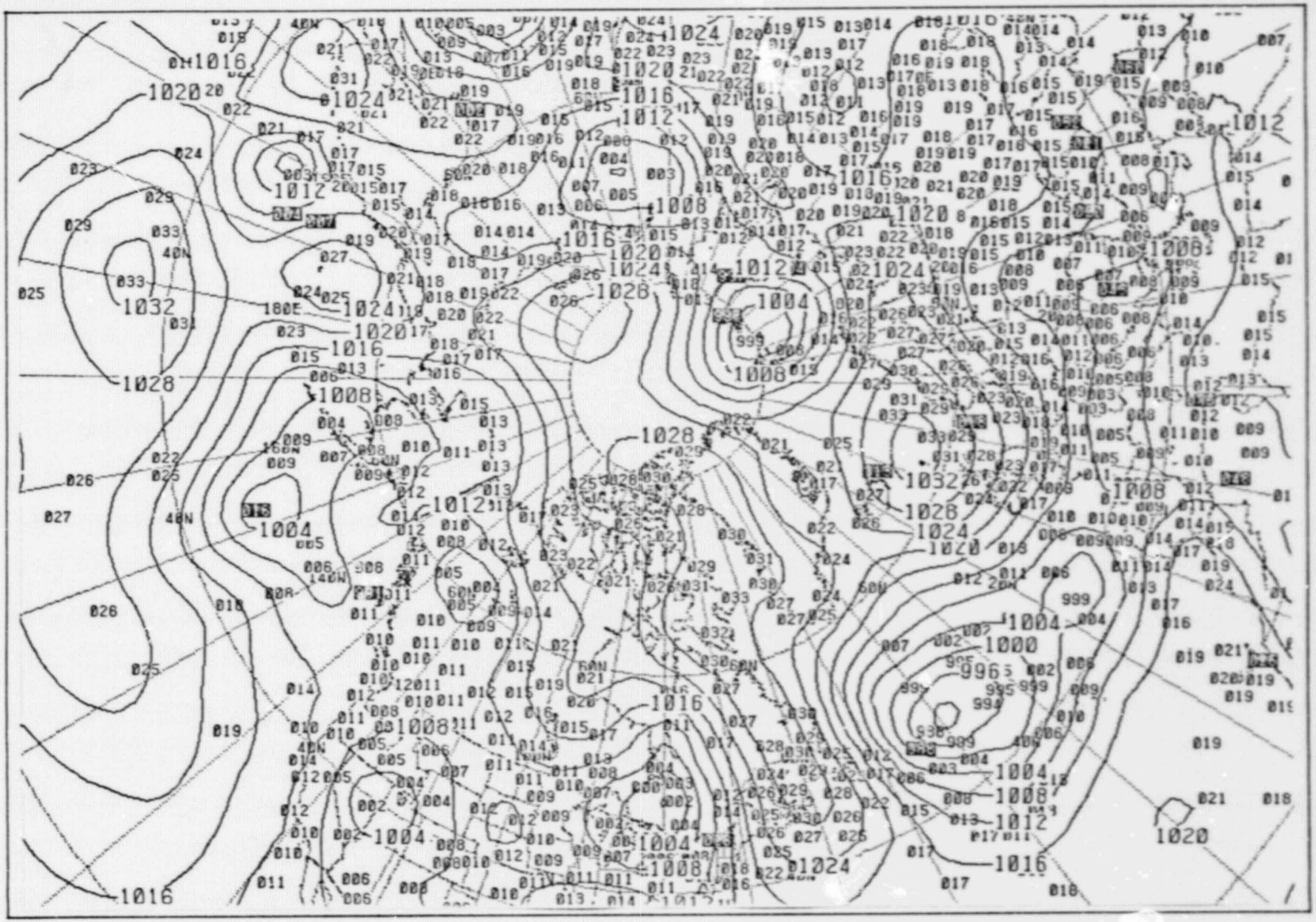


FIGURE II-4: SEA LEVEL PRESSURE ANALYSIS USING FORECAST FIELD
ASMBLWT=100., REDUCE=0.75, GRADWT=10., PCTLAPL=2.5

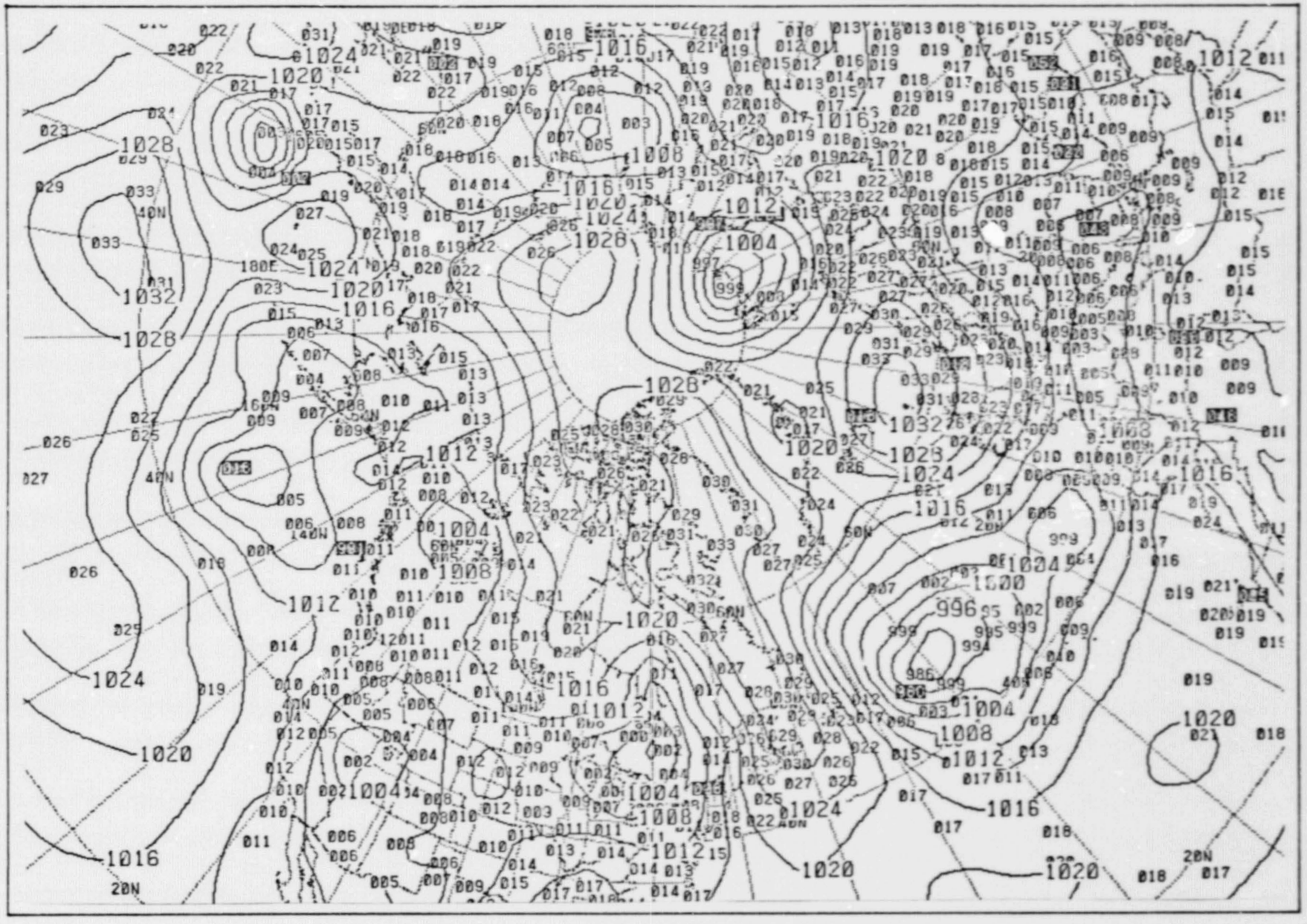


FIGURE II-5: SEA LEVEL PRESSURE ANALYSIS USING FORECAST FIELD
ASMBLWT=1000., REDUCE=0.75, GRADWT=10., PCTLAPL=2.5

PCT equations, and the number of reports rejected in the analysis. (The RMS departure is a weighted difference between an observation and the analysis value interpolated to the location.)

As the emphasis on the assembled field is increased, the RMS values become smaller, as would be expected, indicating that the analysis is approaching the form suggested by the observations. The PCT equations are solved by an iterative over-relaxation technique. It is advantageous, in terms of total program time, to have as few iterations as possible. Table II-1 shows the sharp reduction in the number of iterations that occurs (from 39 to 12) when the assembled field weight becomes ten versus one. Similarly, the number of rejected reports drops to about half for the same increase in the assembled field weight. The number of iterations again decreases when ASMBLWT is raised to 100 but further increases appears not to further reduce the number of iterations. However, the number of rejected observations continues to decrease, but at a much less rapid rate.

From the preceding discussion, it is apparent that when the initial guess field is not especially good, a relatively large PCT assembled field weight is required to obtain an analysis which closely fits the observations. How large an ASMBLWT is needed when the first-guess field is more accurate? To answer this question, several runs of the surface analysis were made with the FNWC analyzed sea-level pressure field serving as the first guess. Of course, this field, Figure II-6, shows most of the synoptic features which the

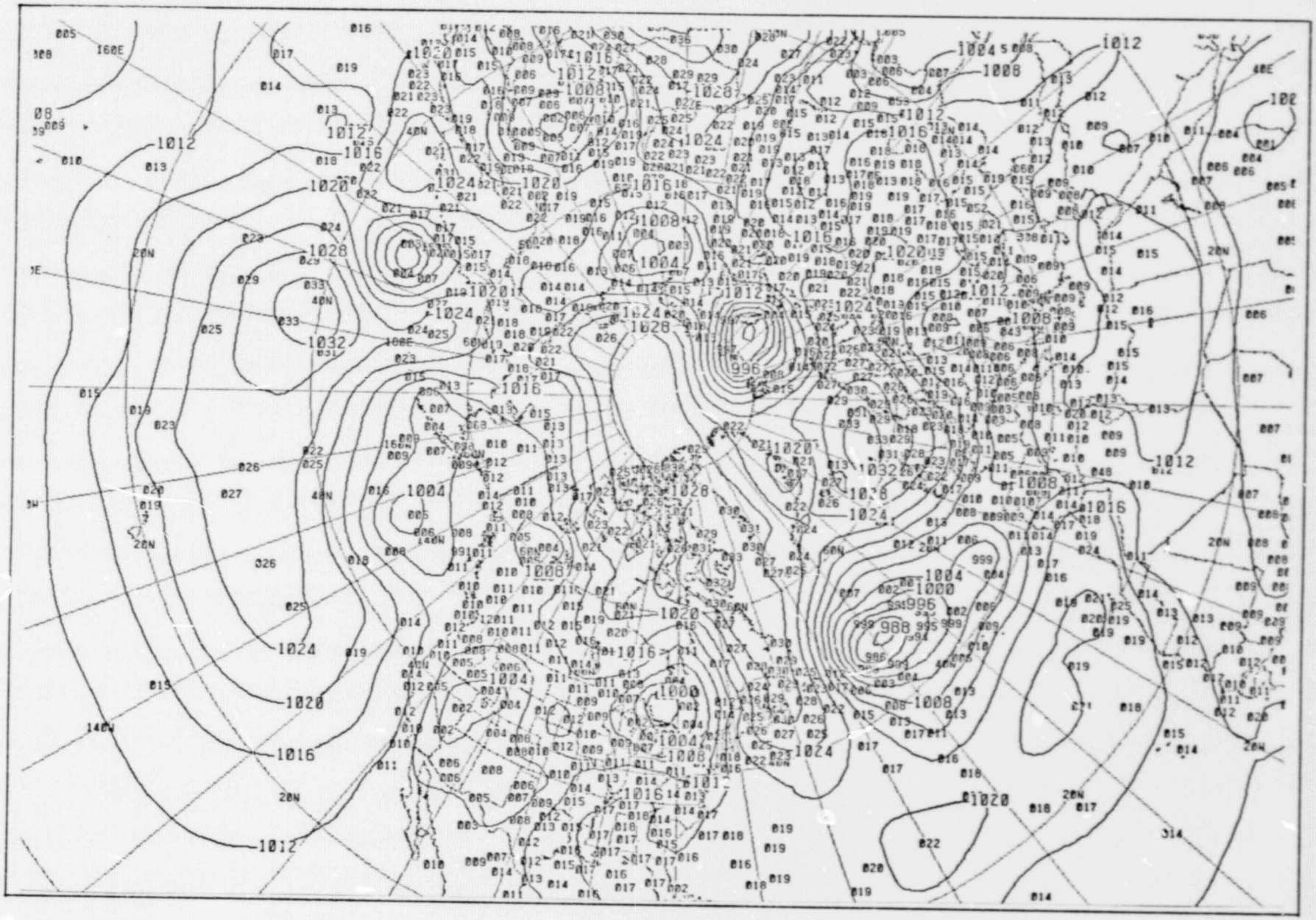


FIGURE II-6: SEA LEVEL PRESSURE ANALYSIS, FIRST-GUESS FIELD FROM FNWC ANALYSIS.

observations suggest, such as the previously discussed cyclone in the North Pacific and the 1032mb contour in the Pacific high. Note also that the Atlantic cyclone has a 988mb contour.

Starting with this better first guess, the analysis was run again varying the assembled field weights by a factor of ten, from one to one hundred. The gradient and Laplacian weights remained the same as before. In the case of the better guess field, when the value of the assembled field weight is relatively small, the resulting analysis closely resembles the first-guess field (compare Figures II-6 and II-7). Figures II-8 and II-9 show the resultant analyzed pressure fields when ASMBLWT takes on the values of ten and one hundred, respectively. There is little difference, however, between them and the analysis where the assembled field weight is one, with the exception of a slightly better depiction of the two cyclones under discussion. The statistics in Table II-2, entries 1-3, further support the conclusion that there is relatively little difference between them, as evidenced by the RMS values which range from 0.82 to 0.77mb. When the first-guess field already fits the observations, this is to be expected. However, it is important to compare the number of iterations required for convergence in the solution of the PCT equations. If the value of the assembled field weight is less than the other two weights, i.e., the gradient and Laplacian weights, the number of iterations required for convergence is significantly larger, in this case twice the number needed in the other two runs.

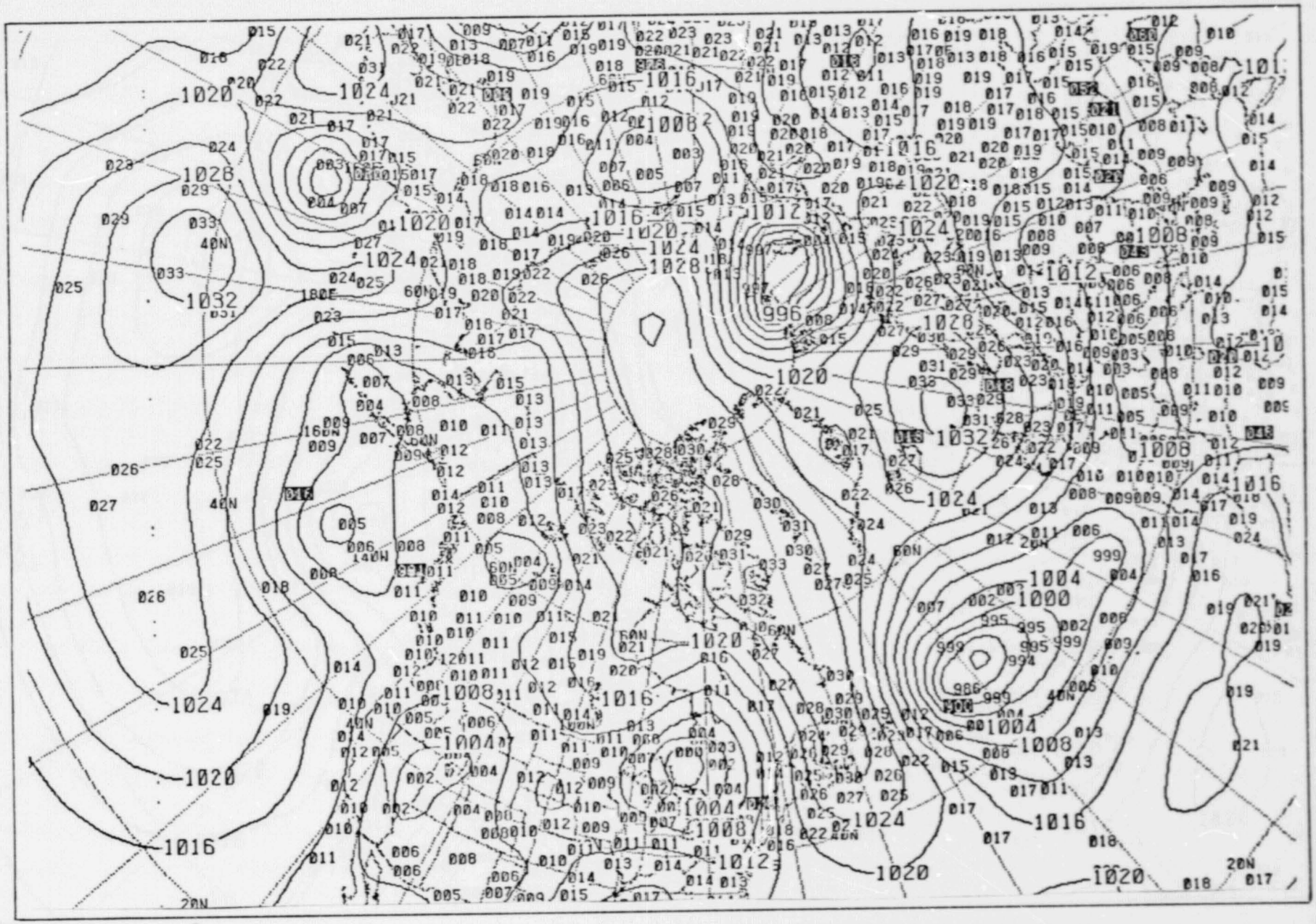


FIGURE II-7: SEA LEVEL PRESSURE ANALYSIS, USING FNWC FIELD, ASMBLWT=1, REDUCE=0.75, GRADWT=10, PCTLAPL=2.5

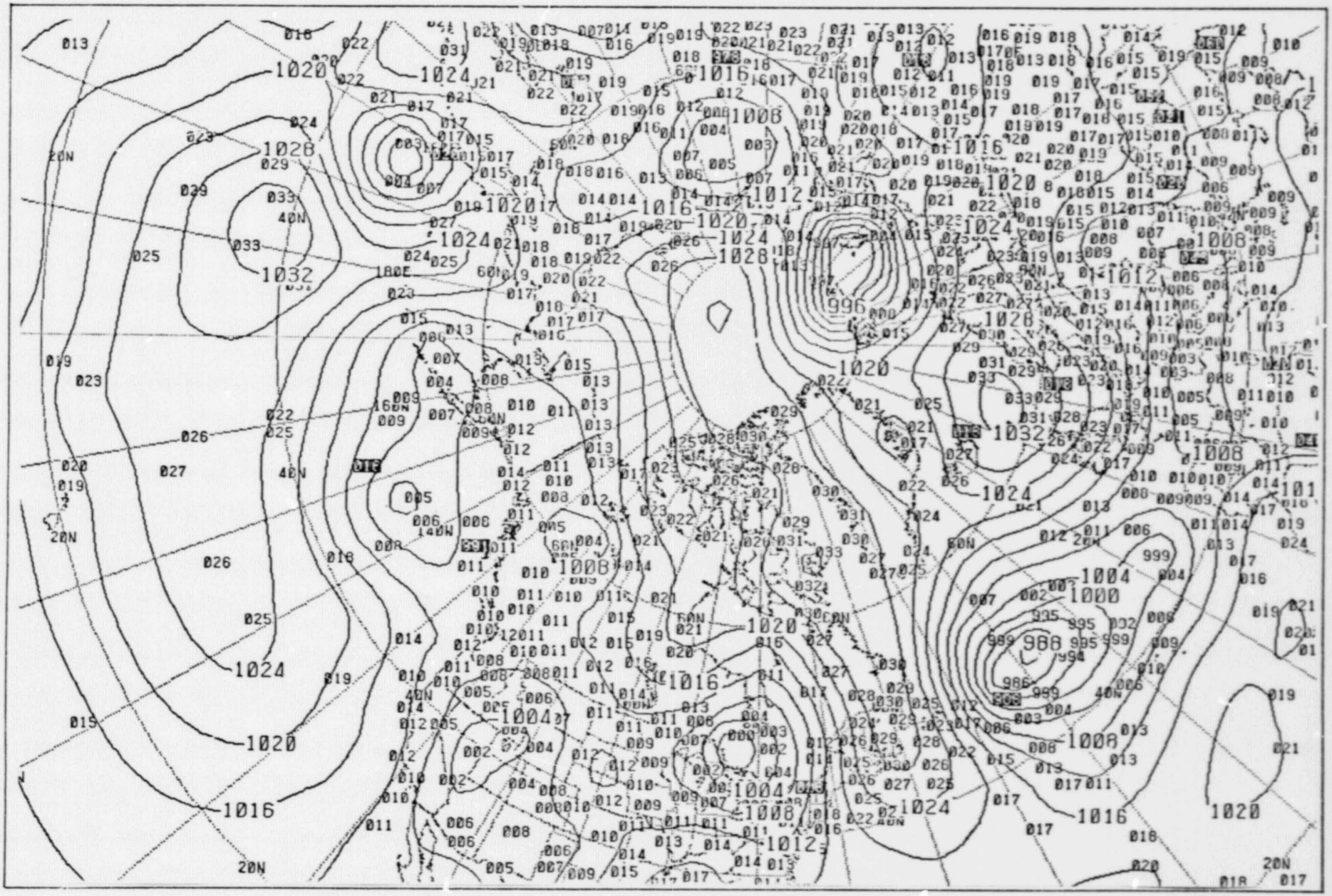


FIGURE II-8: SEA LEVEL PRESSURE ANALYSIS, USING FNWC FIELD
ASMBLWT=10, REDUCE=0.75, GRADWT=10, PCTLAPL=2.5

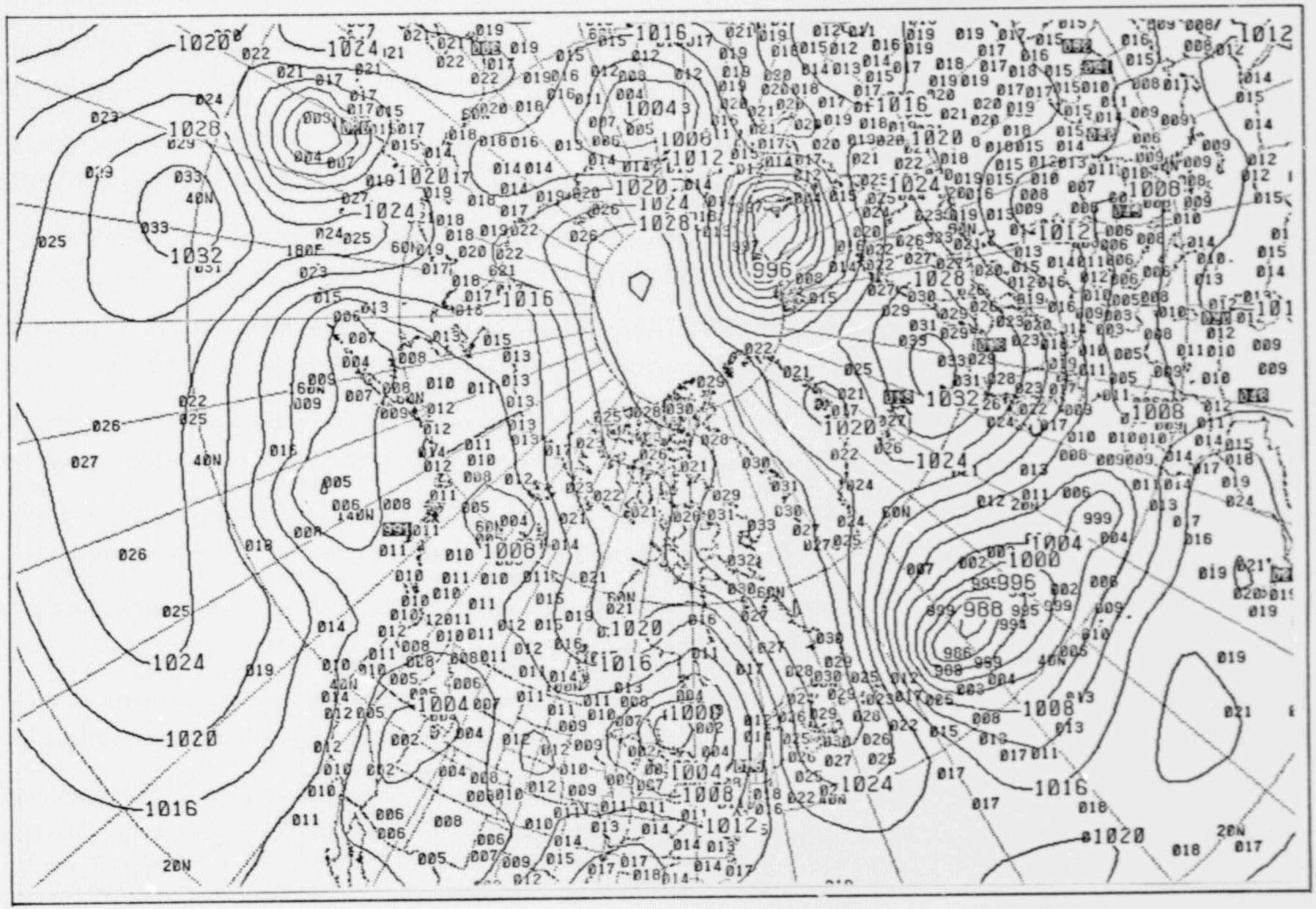


FIGURE II-9: SEA LEVEL PRESSURE ANALYSIS, USING FNWC FIELD,
ASMBLWT=100, REDUCE=0.75, GRADWT=10, PCTLAPL=2.5

TABLE II-2: SEA LEVEL PRESSURE ANALYSIS STATISTICS -- FNWC FIELD AS FIRST GUESS.

ENTRY	ASSEMBLED FIELD WEIGHT	GRADIENT WEIGHT	LAPLACIAN WEIGHT	REDUCE	RMS (MB)	ITERATIONS	REJECTED OF 4706
1	1.0	10.0	2.5	0.75	0.82	14	175
2	10.0	10.0	2.5	0.75	0.80	8	175
3	100.0	10.0	2.5	0.75	0.77	7	167
4	100.0	10.0	2.5	0.00	0.79	7	170

In an operational context, this can be a major factor in deciding how to assign weights, especially if a finer grid requiring a much greater number of computations were to be used instead of the 63x63 grid.

Further tests were conducted in which the Laplacian and gradient weights were varied while holding the assembled field weight constant. Comparing entries 3, 7, and 9 in Table II-1, cases for which the gradient and assembled field weights were held at ten and one hundred, respectively, it can be seen that increasing the Laplacian weight from 2.5 to 250 actually increases the RMS values from 0.88 to 1.32mb. The number of iterations for the solution of the PCT equations jumps from 7 to 47. Similarly, when the gradient weight is increased, a very large number of iterations is required to solve the PCT equations. As shown in entries 2, 10, and 11 of Table II-1, when the assembled field and Laplacian weights are 10 and 2.5, respectively, the number of iterations required rises as does the RMS error (from 1.02 to 1.41) when the gradient weight increases.

Figure II-10 shows the analysis field which results when the gradient weight is set to 100. Comparing this figure with the initial field in Figure II-1, it can be seen how effectively the analysis is constrained to resemble the original guess field. Reports that increase the gradients are rejected. A 1008mb contour in the low over the western United States is found in both figures. The analysis field in Figure II-10 also shows the Pacific high with a maximum contour of only 1028mb. The lowest reported pressures in the Atlantic cyclone have been rejected.

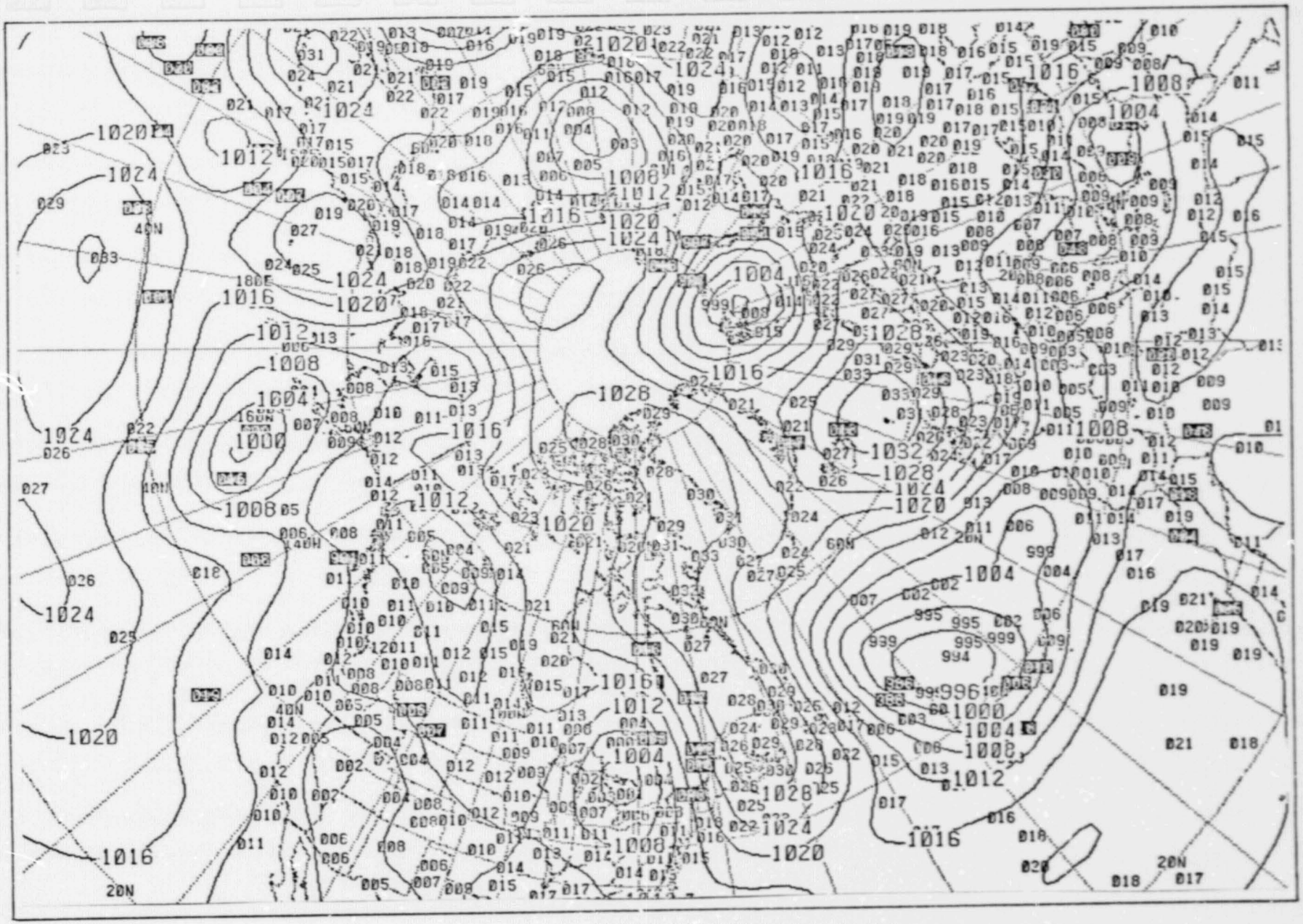


FIGURE II-10: SEA LEVEL PRESSURE ANALYSIS, USING FORECAST FIELD, ASMBLWT=10, REDUCE=0.75, GRADWT=100, PCTLAPL=2.5

The other factor which must be taken into consideration in these tests is the effect of including the information density modification to the PCT weights. Recall that where the observations are closely spaced, the weights on the assembled field are relatively higher and the weights on the differential properties are lower than in areas with few reports. Several runs were made without the information density factor. These cases are the entries in Tables II-1 and II-2 for which REDUCE equals zero. Regardless of the combination of PCT weights, the RMS values are better when the information density is included. This result could be anticipated since the inclusion of the information density increases the weights on the assembled field (which most clearly approximates the data). The decreased emphasis on the gradient and Laplacian terms in these same areas further supports an analysis which accommodates the observational information. In the data sparse regions, the field is more constrained to preserve the characteristics of the guess field, but the assembled field weights are still slightly larger than they would be without INFOFAC. It is not as obvious why fewer iterations are required to solve the PCT equations, but it appears that this is always true.

As a final step in the examination of the PCT weights, a test was made in which the PCT equations were eliminated. The analysis field, Figure II-11, represents the smoothed output after three cycles of the assembly procedure. If this figure is compared with the field obtained with strong assembled field weights (ASMBLWT, GRADWT, and PCTLAPL are

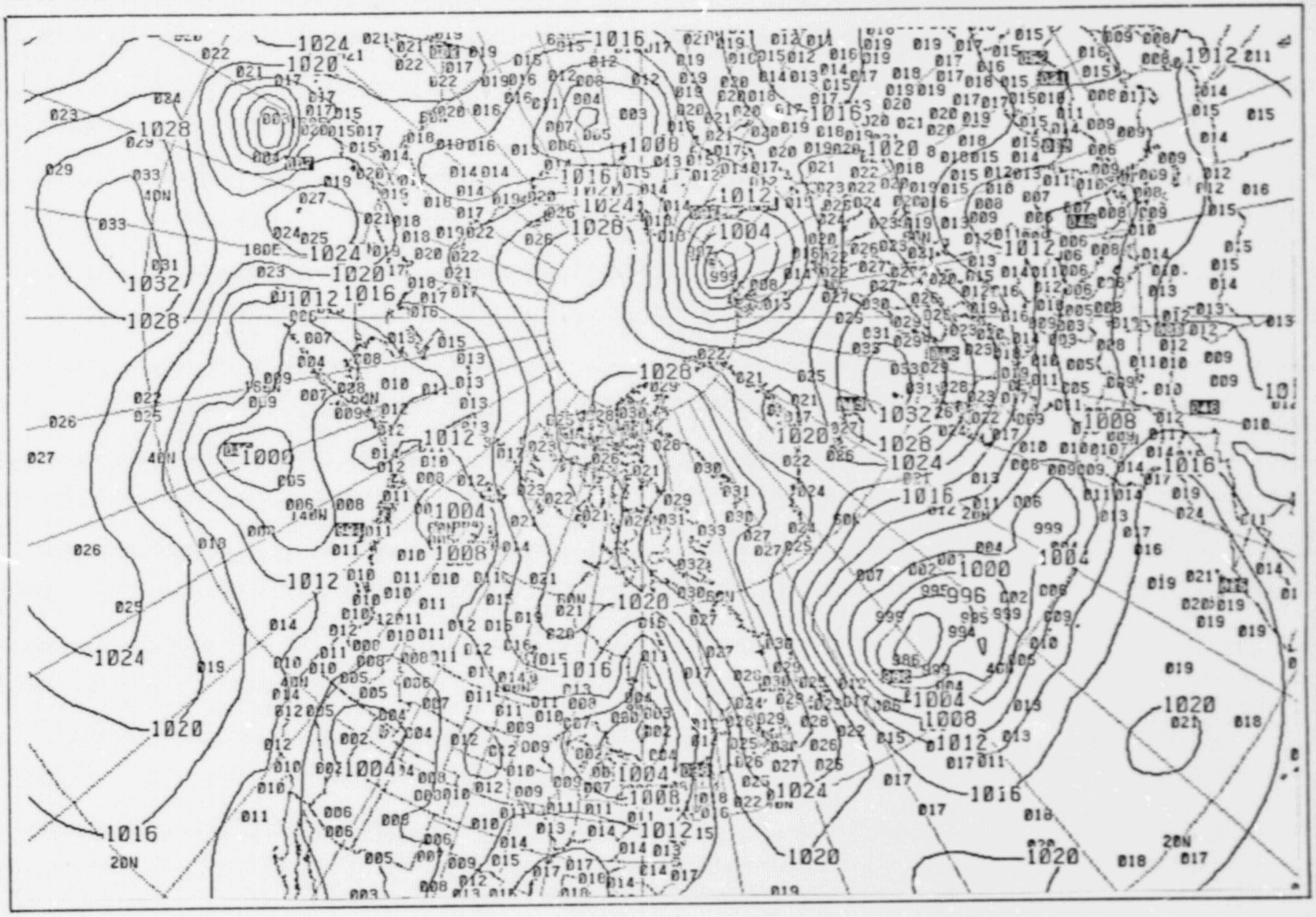


FIGURE II-11: SEA LEVEL PRESSURE ANALYSIS USING FORECAST FIELD, NO PCT.

1000., 10., and 2.5, respectively) shown in Figure II-5, it is difficult to find differences between them. The similarity between the two is further supported by the statistics in Table II-1. Entries 4 and 12 show the same RMS and the same number of rejected reports. Thus, it would appear that where the PCT assembled field weight is very much larger than the weights on the differential properties of the guess field, there may not be much advantage in using the PCT equations, at least for the surface pressure analysis.

C. Sea Surface Temperature and Upper Air Analyses

The four remaining analyses -- the sea surface temperature, and upper air temperature, height and wind analyses -- will be discussed as a unit. They share a common problem in that, relative to the surface pressure analysis, few observations are available. This increases the importance of the first-guess field properties in the PCT equations. These analyses are likely to become "noisy" if large weights are assigned to the assembled fields because the reports are widely spaced and the observational input at a gridpoint is generally from one rather than several reports.

To illustrate, we will compare several runs of the sea surface temperature (SST) analysis. The SST first-guess field appears in Figure II-12. The isotherms in the Pacific Ocean are generally parallel. This configuration changes according to the relative sizes of the PCT weights. In the following SST analyses, the weights were varied as in the first set of pressure analyses. Weights on the gradient and Laplacian terms were held constant (10.0 and 2.5, respectively) while the weights on the assembled field increased by a factor of ten, from one to one thousand. Figures II-13 and II-14 represent the analyses for which ASMBLWT was one and ten, respectively. The amount of curvature in the eastern Pacific isotherms is quite different between the two analyses. The strongly curved 22°C contour (Figure II-14) occurs when the analysis tries to accommodate two widely-spaced 22°C observations. With the smaller ASMBLWT (one), there is more emphasis on the original gradients in the PCT equations

"Page missing from available version"

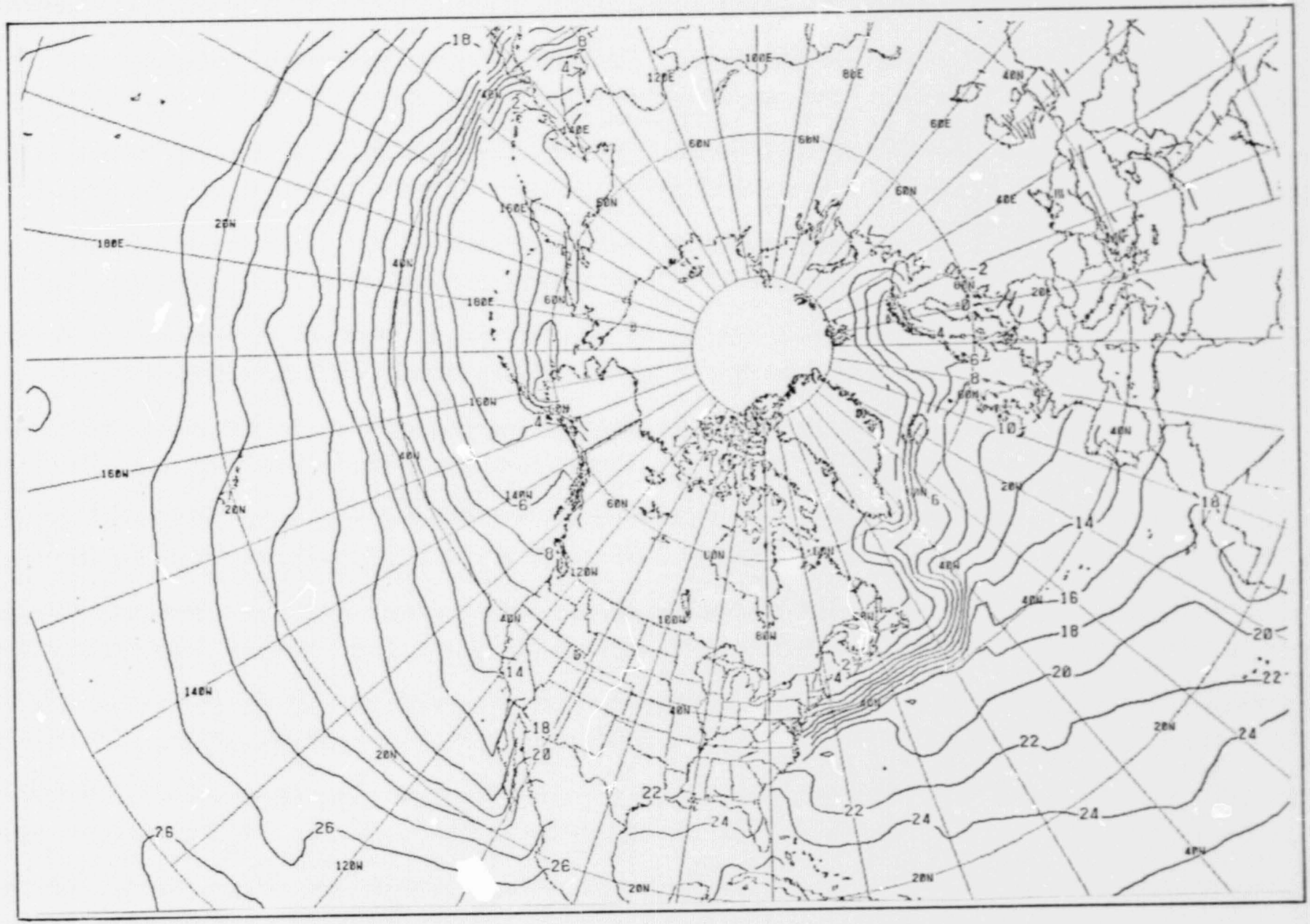


FIGURE II-12: FIRST-GUESS FIELD FOR SEA SURFACE TEMPERATURE ANALYSIS.

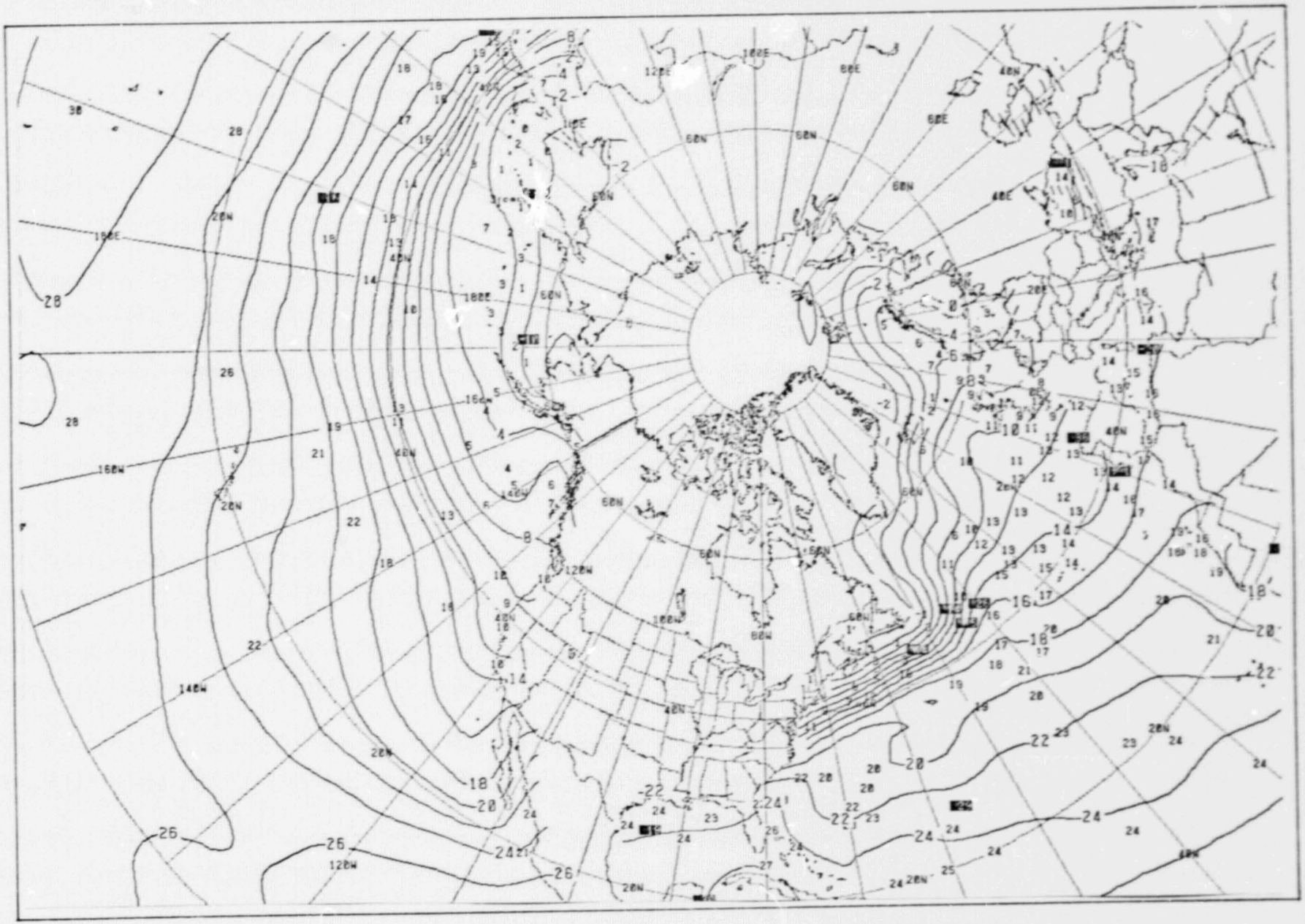


FIGURE II-13: SEA SURFACE TEMPERATURE ANALYSIS, ASMBLWT=1., GRADWT=10., PCTLAPL=2.5

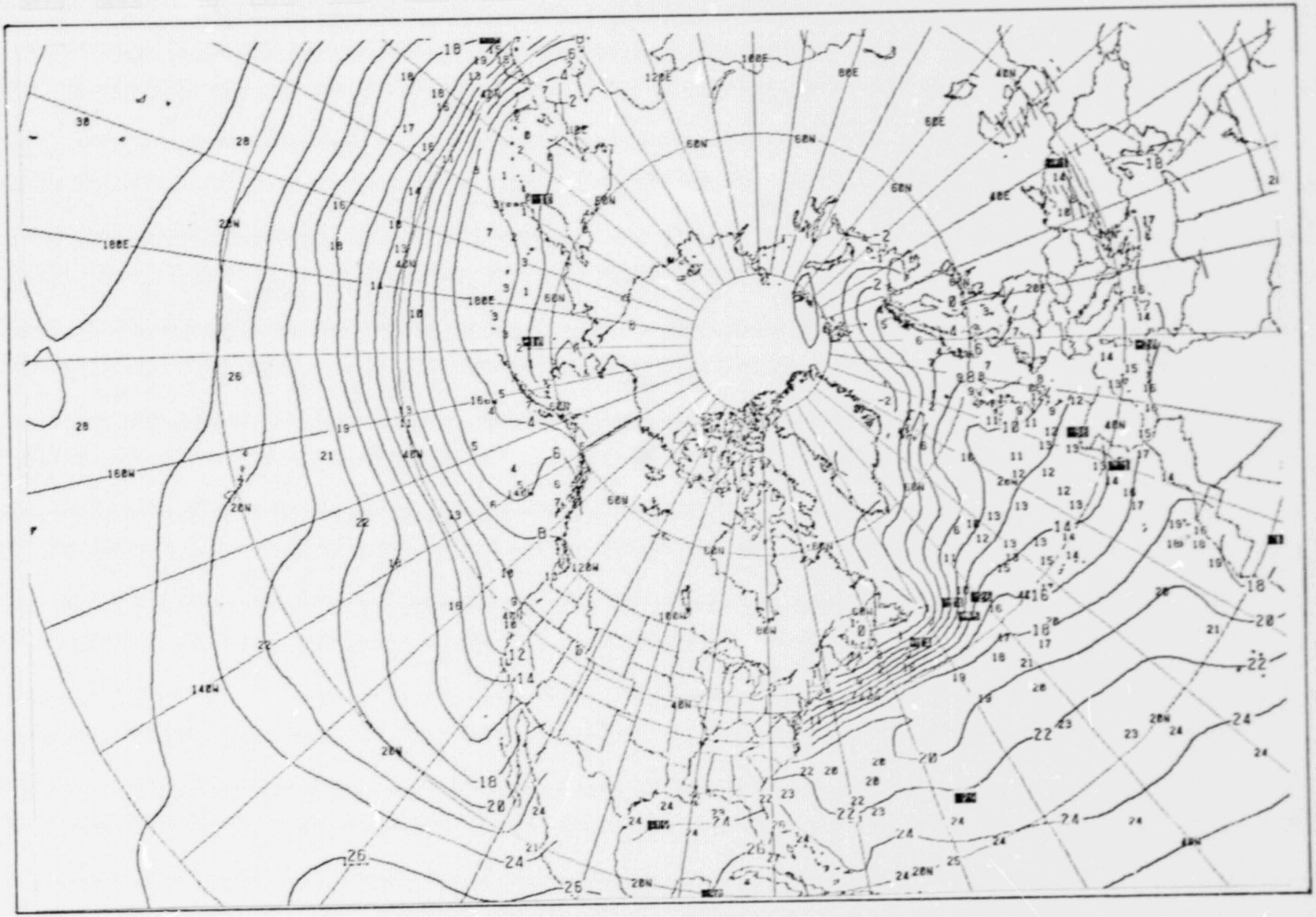


FIGURE II-14: SEA SURFACE TEMPERATURE ANALYSIS, ASMBLWT=10., GRADWT=10., PCTLAPL=2.5

which results in a smoother field (Figure II-13). The 24°C isotherm near the West Indies displays a similar change in configuration between the two analyses. As the weight on the assembled field becomes larger, the effect of the isolated reports becomes more pronounced. Compare the SST analysis in Figure II-15, where the assembled field weight is one hundred, with the previous two analyses.

It could be argued that a human analyst would smooth these isotherms because the sea surface temperature measurements are prone to errors. Some regions tend to be smoothly-varying while others are not. In the Gulf Stream and Kuroshio currents are warm, relatively narrow currents with strong gradients. However, most of the world's oceans are characterized by weak gradients and smooth distributions of the sea surface temperature. For this reason, either of the first two analyses (Figures II-13 and II-14) might be considered more acceptable than the last despite their higher RMS values. See Table II-3 for the statistics for those runs.

To achieve the relative smoothness of the field for SST analysis, the weights in the differential properties must be relatively large. Since it is also desirable to emphasize the observations as much as possible, the best balance appears to result when the assembled field weights are approximately the same order of magnitude as the other PCT weights. In considering the time required to solve the PCT equations, it is found that this relative size relationship keeps the number of iterations small. For instance, for the

TABLE II-3: SEA SURFACE TEMPERATURE STATISTICS

ENTRY	ASSEMBLED FIELD WEIGHT	GRADIENT WEIGHT	LAPLACIAN WEIGHT	REDUCE	RMS (°C)	ITERATIONS	REJECTED of 442
1	1.0	10.0	2.5	0.75	0.95	24	29
2	10.0	10.0	2.5	0.75	0.90	9	27
3	100.0	10.0	2.5	0.75	0.87	7	28
4	100.0	10.0	2.5	0.00	0.88	8	28
5	1000.0	10.0	2.5	0.75	0.87	6	28
6	--	--	--	--	0.87	--	28

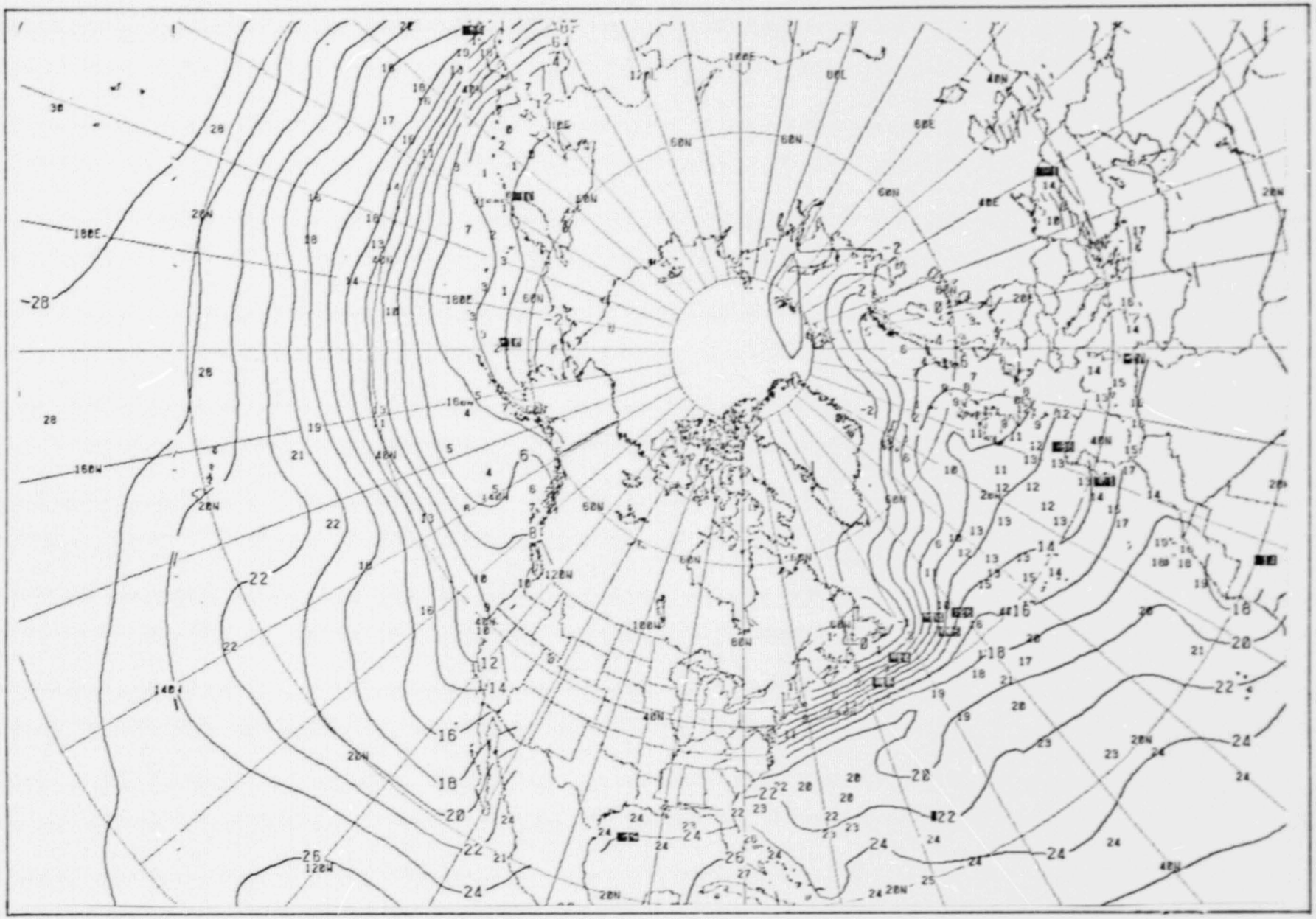


FIGURE II-15: SEA SURFACE TEMPERATURE ANALYSIS, ASMBLWT=100., GRADWT=10., PCTLAPL=2.5

SST analysis, nine iterations were required to solve the PCT equations when the assembled field weight was ten versus the twenty-four required when ASMBLWT was one (see Table II-3).

The examples presented thus far have included the modification of the weights based on the information density, as explained in Section II-A. By comparing entries 3 and 4 in Table II-3, it is evident that the use of the information density affects this analysis in much the same way as it does the surface pressure analysis. That is, the RMS value is smaller and fewer iterations are required for the solution of the equations when the information density modifies the PCT weights.

Two additional examples appear in Table II-3. Entries 5 and 6 show the statistics for the analysis when ASMBLWT was one thousand and the analysis was done omitting the PCT equations. As was true for the corresponding pressure analyses, these two analyses are quite similar both statistically and in the fields which are produced (compare Figures II-16 and II-17). This supports the statement made earlier that the time required to solve the PCT equations may not be justified if the assembled field weights are several orders of magnitude larger than the differential property weights.

In many respects, the results from the upper air analyses parallel those for the sea surface temperature analysis. In studying the statistics for several runs of the temperature analysis under the same series of weight combinations, it becomes clear that several generalizations can be made (Note: this following discussion applies also to

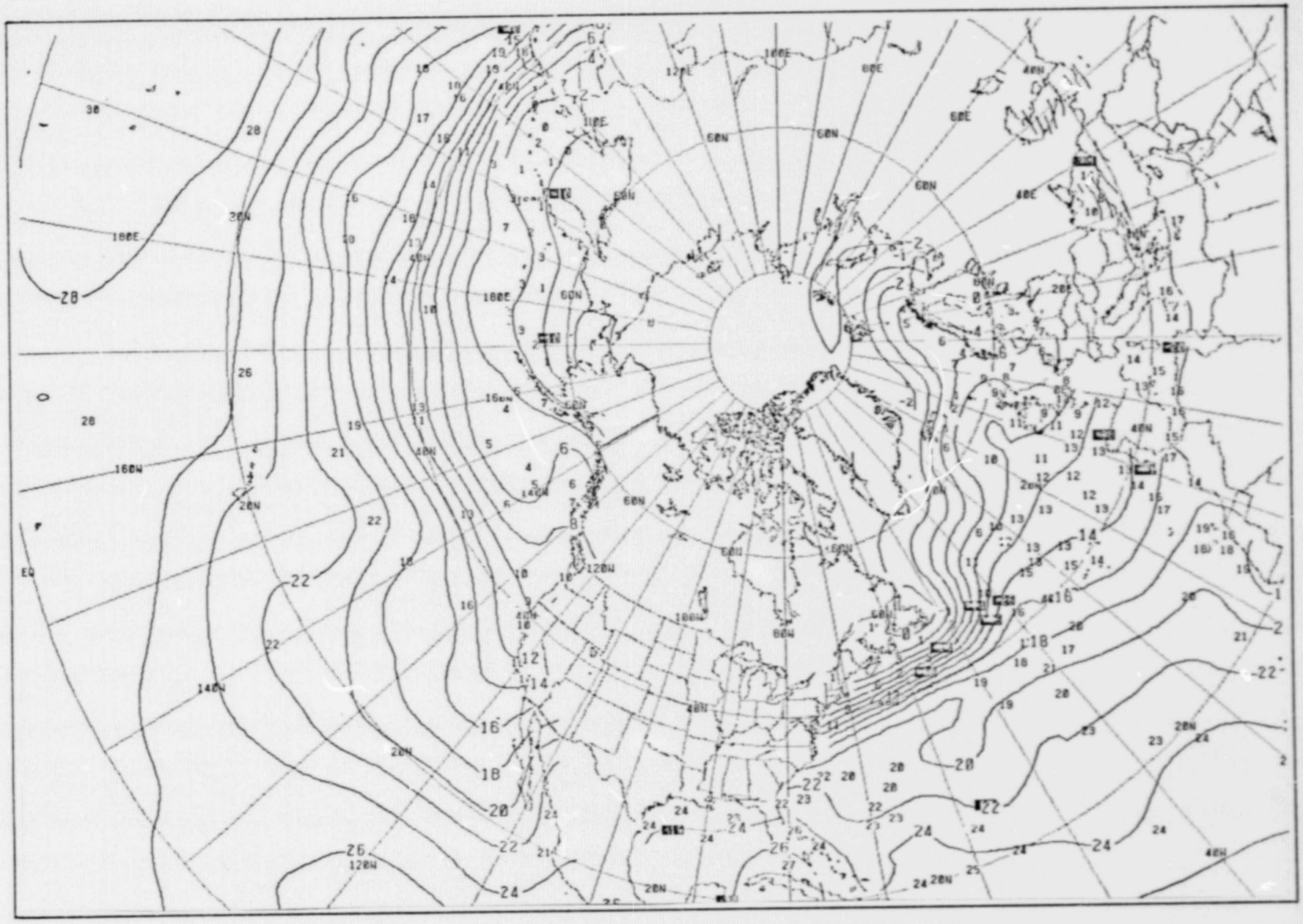


FIGURE II-16: SEA SURFACE TEMPERATURE ANALYSIS, ASMBLWT=1000., GRADWT=10., PCTLAPL=2.5

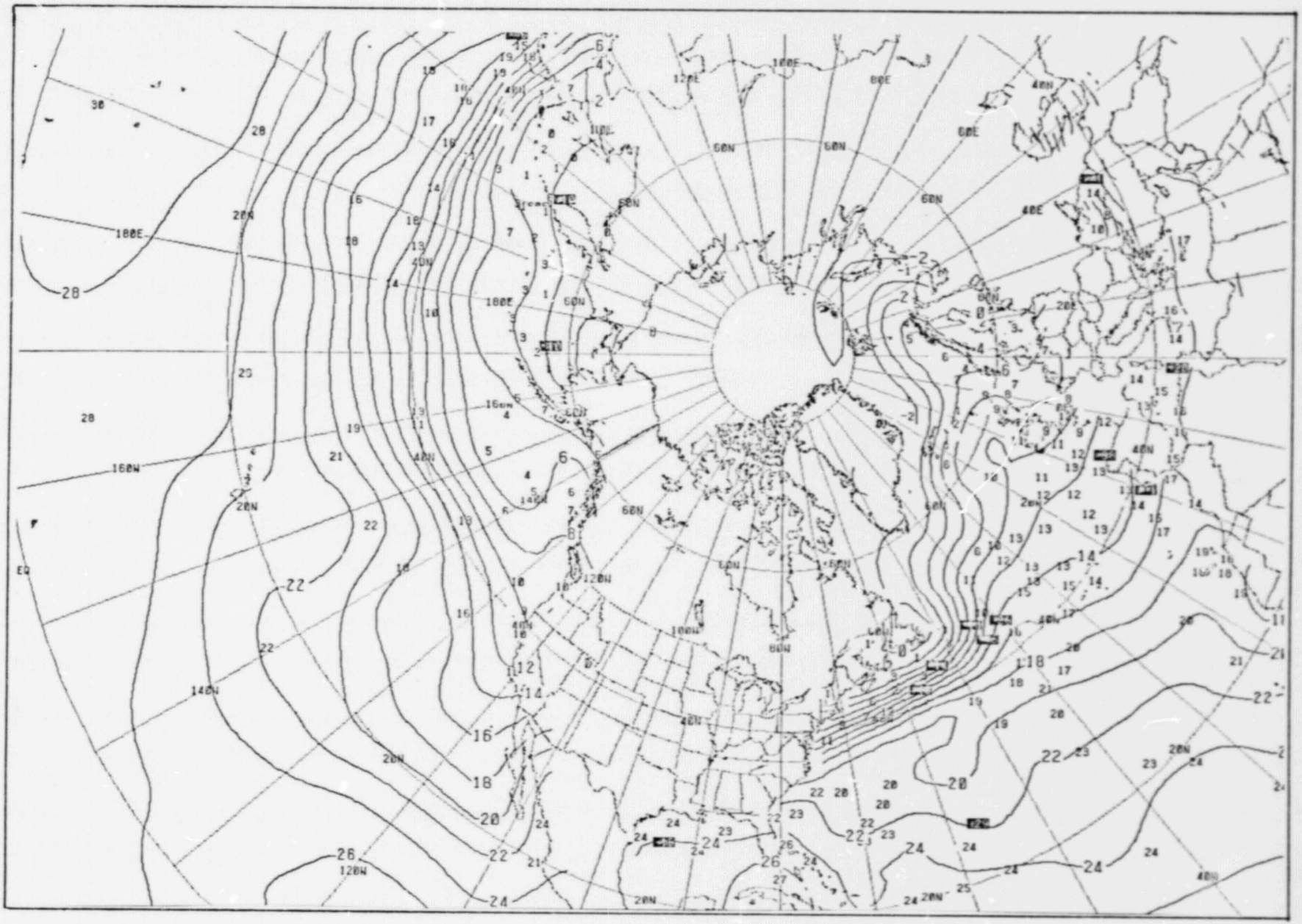


FIGURE II-17: SEA SURFACE TEMPERATURE ANALYSIS, NO PCT.

the height analysis). Table II-4 shows the statistics for several of the tests of the temperature analysis. For simplicity, the results from only two of the twelve levels have been included, the 1000mb and 500mb levels. The number of iterations required for the convergence of the PCT equations decreases as the relative size of the weights on the assembled field increases, here dropping from thirty-one iterations when the assembled field weight was one to six iterations when the weight was one thousand. When the PCT equations are not used (see Entry 5), the statistics vary little from the analysis with very large assembled field weights (Entry 4).

One additional technique was tested which involved the filtering of the first-guess field before the analysis was done. For the temperature analysis, this somewhat improved the smoothness of the analysis and improved the statistics for a case where the guess field contains a lot of noise. In using a better guess field, however, little overall improvement is achieved with an initial filtering of the field. Compare Figures II-18 and II-19, the 1000mb level analyses for which the first-guess field was filtered and not filtered, respectively. The filter has smoothed the 20° contour in the Atlantic, removing a feature which was found in the first-guess field but which is not supported by data. However, the temperature analysis over the southwestern portion of the United States is quite different. In the filtered version, the smoothing of the field apparently caused more reports to be rejected (Figure II-18) than in

TABLE II-4: TEMPERATURE ANALYSIS STATISTICS -- 1000 AND 500 MB LEVELS.

ENTRY	ASSEMBLED FIELD WEIGHT	GRADIENT WEIGHT	LAPLACIAN WEIGHT	REDUCE	RMS (°C)	ITERATIONS	REJECTED
1	1.0	10.0	2.5	0.75	1.27	31	6.00%
					1.12	19	1.62%
2	10.0	10.0	2.5	0.75	1.09	10	4.55%
					0.94	9	1.62%
3	100.0	10.0	2.5	0.75	0.83	6	3.82%
					0.70	6	1.62%
4	1000.0	10.0	2.5	0.75	0.76	6	3.64%
					0.64	6	1.62%
5	--	--	--	--	0.75	--	3.64%
					0.63		1.62%
6*	1000.0	10.0	2.5	0.75	0.75	6	3.09%
					0.64	6	1.62%

* Filtered First-Guess Field

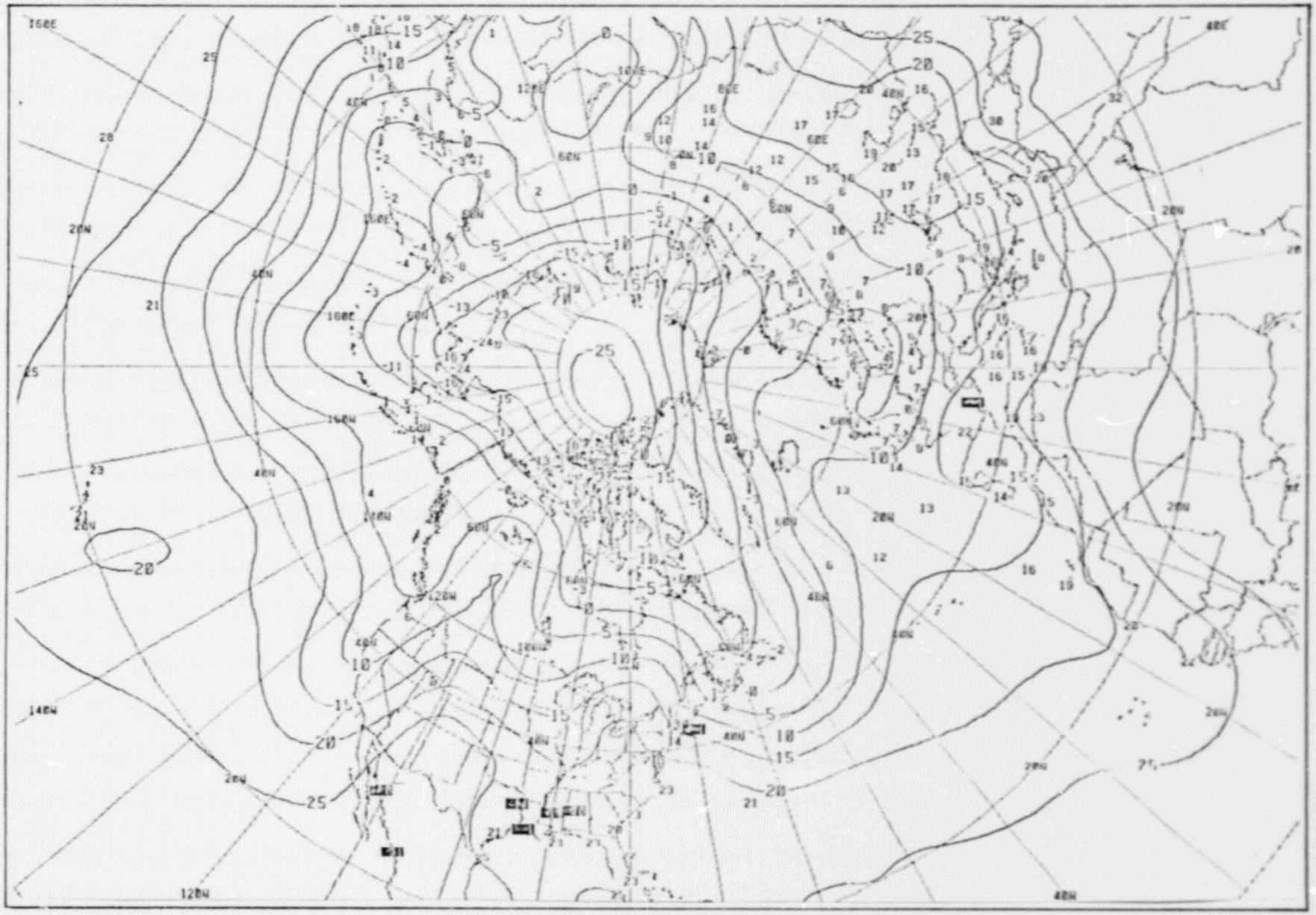


FIGURE II-18: TEMPERATURE ANALYSIS, 1000MB LEVEL, USING FILTERED FIRST-GUESS FIELD.

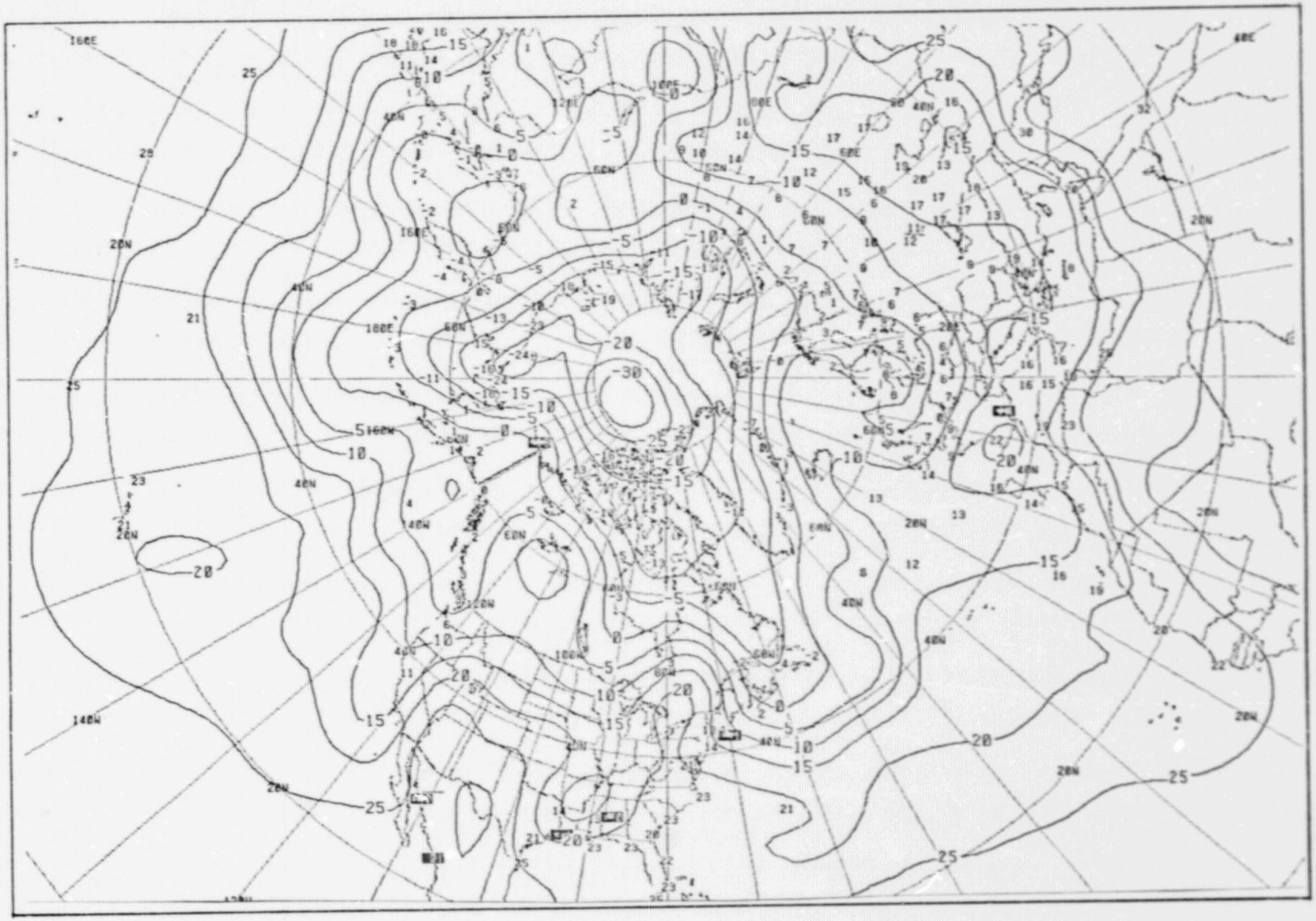


FIGURE II-19: TEMPERATURE ANALYSIS, 1000MB LEVEL, USING UNFILTERED FIRST-GUESS FIELD.

the other analysis (Figure II-19). Whether all of these reports should have been rejected is difficult to answer. Of course, the final analyzed temperature field is generally smoother and this may be an important criterion for the forecast model.

As an additional note, the wind analysis first-guess field is derived geostrophically from the height analysis. As a result, the weights on the observations (via the assembled field) are set relatively larger than the other weights.

In place of the Laplacian constraint, a divergence and vorticity term are used in the PCT equations, although the relative weights on these properties are small. In the course of this work, the observed winds were incorporated into the height analyses (see Section V). With a relatively large assembled field weight in the wind analysis, consistency is maintained between the two analyses.

Table II-5 contains the values of the PCT weights for each type of analysis as currently employed in the 63x63 NASA-ODSI model.

TABLE II-5: PCT WEIGHTS.

	SST	SURFACE PRESSURE	TEMPERATURE	HEIGHT	WIND
Assembled Field	5.0	25.0	10.0	25.0	10.0
Gradient	5.0	1.0	1.0	5.0	0.02
Laplacian	0.5	0.25	0.25	5.0	--
Vorticity	--	--	--	--	0.02
Divergence	--	--	--	--	0.10

D. Summary

The results for selected tests of the PCT weights have been described from which several conclusions can be made:

1. Convergence of the PCT equations occurs more rapidly as the relative size of the PCT assembled field weight increases.

2. The first-guess field has a very significant affect on the quality of resulting analyzed fields. However, regardless of the quality of the guess field, it is advantageous to have the assembled field weights at least the same order of magnitude as the weights on the differential properties both in terms of response to data input and the time required to do the analysis.

3. If the assembled field weights are to be several orders of magnitude greater than the other PCT weights, it is better to omit the use of the PCT equations. The time for the analysis is reduced with very little difference in the resulting analyzed field.

4. The use of the Information Density to modify the PCT weights appears to improve the analysis.

5. In the surface pressure analysis, the much greater availability of data allows the use of much larger assembled field weights relative to the other PCT weights. The wind analysis should emphasize the observations, especially with the use of the observations in the height analyses. To avoid the introduction of very small-scale features in the other analyses, the assembled field weight must be approximately the same order of magnitude as the gradient weight.

III. INFLUENCE OF DATA - ASSEMBLY PROCEDURES

The variation in the number, quality and distribution of observations presents the most basic problem in objective analysis. Questions arise such as how to determine the accuracy of a given report, whether the report reflects general conditions or local effects, and to what extent a report should influence the surrounding area. These problems are especially acute over the oceans generally and over certain land regions as well where the paucity of data limits the quality of an objective analysis.

ODSI has introduced into the assembly procedure of the models several new techniques which improve the analyses. In the following sections, these techniques and their role in determining the impact of the data reports are discussed. The topics include:

1. The basic area influenced by a report - Region of Influence.
2. The proximity of other data reports - Information Density.
3. The properties of synoptic distributions near the location of the report - Gradient Factor and Laplacian Dependent Weights.
4. The degree to which a report departs from the surrounding field - Reevaluation of Data Weights and Reject Criteria.

A. Region of Influence

The ODSI objective analysis scheme involves several scans or cycles in which data are assembled to grid locations and the PCT equations solved. With each successive scan, the region influenced by a report decreases in size. Presently, three scans are made in each of the analysis programs.

The basic influence region is a circle whose radius is the product of a radius factor (FACT), the basic scan unit (RAD), and the map or image plane factor (AMAP). This product is normalized by the standard meshlength at the reference latitude on a 63x63 hemispheric polar stereographic grid (AMESH) so that the radius of influence is defined in terms of meshlengths. That is:

$$\text{RADIUS} = \frac{\text{FACT} * \text{RAD} * \text{AMAP}}{\text{AMESH}}$$

where RADIUS = assembly radius, or radius of influence.

FACT = the radius factor, a function of the maximum radius, the information density, and local gradients.

AMAP = map factor = $\frac{1 + \sin \phi_o}{1 + \sin \phi}$

ϕ_o is the reference latitude on a polar stereographic grid (60° in this case) and ϕ is the latitude of the location of the data report.

AMESH = standard meshlength at ϕ_o = 381 kilometers (205.74 nm).

RAD = basic scan unit, a multiple of AMESH, its value varies according to the parameters being analyzed.

Tests were made to investigate the effect of varying the basic scan unit, RAD. The results suggested that, in the sea-level pressure analysis, RAD should equal the standard meshlength (AMESH) or 205.74 nm. However, the much smaller number of observations available to the other analyses necessitated the use of a larger value of RAD to achieve a smoother distribution of the information. Values up to three times AMESH were tried, but very large values did not further improve the smoothness of the analyses. Consequently, for the sea-surface temperature analysis and the upper-air analyses, the current value of RAD is twice the standard meshlength.

A comparison of Figures III-1 and III-2 demonstrates the greater smoothness which is achieved with the larger RAD value. These figures represent the final assembled fields in the sea-surface temperature analysis. For Figure III-1, RAD was set to one standard meshlength; for Figure III-2, it was doubled. With the larger scan unit, much of the "localized" influence is eliminated. This is especially evident in the North Pacific Ocean.

The other term in the algorithm for the assembly radius, FACT, represents a composite of several factors, all of which vary independently and influence the size of RADIUS. FACT is computed as:

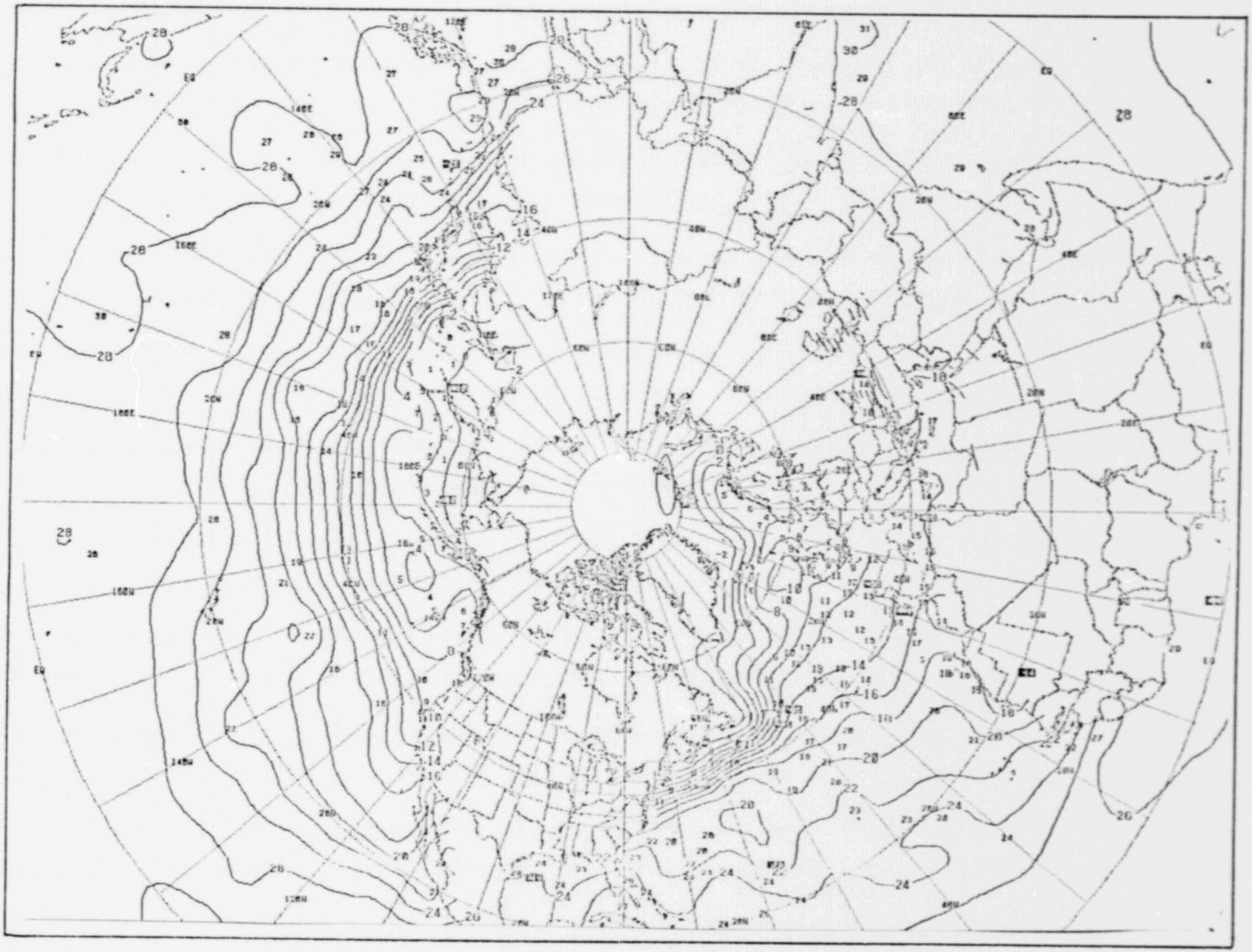


FIGURE III-1: SEA SURFACE TEMPERATURE ANALYSIS, RAD=205.74NM.

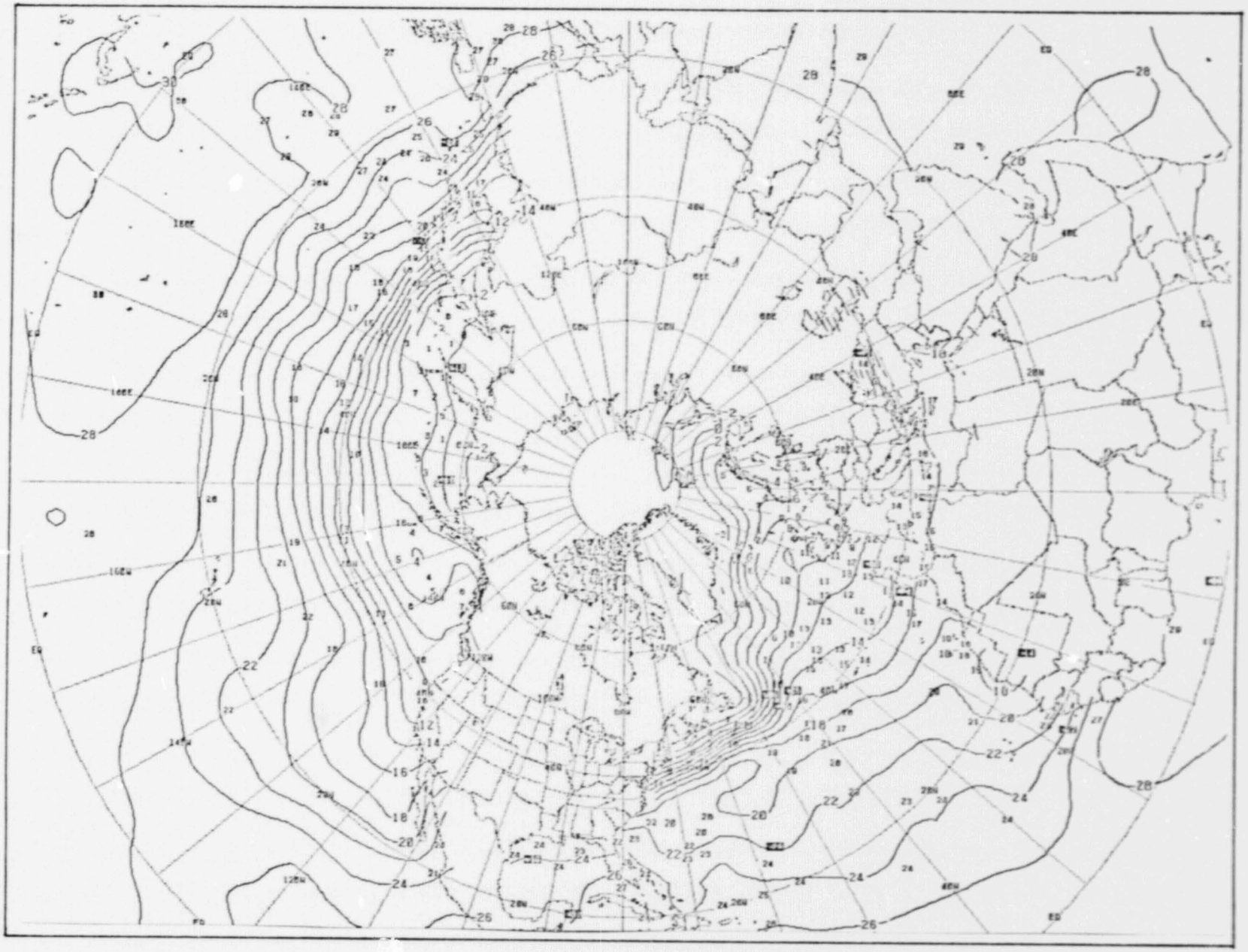


FIGURE III-2: SEA SURFACE TEMPERATURE ANALYSIS, RAD=411.48NM.

$$FACT_{I,J} = GRADFAC_{I,J} * [RADMAX - INFOFAC_{I,J} * (RADMAX - RADMIN)]$$

where I,J = grid indices for the gridpoint closest to the location of the report.

GRADFAC= gradient factor.

INFOFAC= information density factor $0 \leq INFOFAC \leq 1$.

RADMAX, RADMIN = maximum and minimum range factors, respectively.

INFOFAC and GRADFAC will be described in detail in Sections II-B and II-C, respectively. The other two variables, RADMAX and RADMIN, are used to determine a range through which the radius size can vary on a given scan. Since INFOFAC varies between zero and one, this term will lie between RADMAX and RADMIN. The minimum range factor, RADMIN, is constant throughout the analysis program, but the maximum factor, RADMAX, decreases with each new cycle. The new RADMAX is the product of its value from the previous cycle and a fraction, RADFAC:

$$RADMAX_K = RADFAC * RADMAX_{K-1}$$

where the subscript K denotes the scan number. The values assigned to RADMAX, RADMIN, and RADFAC differ slightly from one atmospheric parameter to another. They can be found in the table of assembly variables, Table III-1.

TABLE III-1: ASSEMBLY VARIABLES FOR VARIOUS ANALYSIS TYPES

	Sea Surface Temperature	Surface Pressure	Temperature	Height	Wind
RADMAX	2.8	3.0	2.8	2.8	3.0
RADMIN	1.0	1.0	0.8	0.6	1.0
RADFAC	0.6	0.8	0.6	0.6	0.6
RAD	411.48	205.74	411.48	411.48	205.74
PER	0.20	0.20	0.20	0.20	0.20
A	1.50	1.25	1.25	1.25	-
B	1.00	1.00	1.00	1.00	-
FRACMAX	0.40	0.40	0.40	0.40	-
FRACMIN	0.05	0.10	0.10	0.10	-
PERLAPL	1.00	0.75	0.75	0.75	-
CRIT	1.0	0.5	1.0	1.0	1.0
REFAC	0.5	0.85	0.8	0.70	0.80
CONST	1.5	2.0	2.0	2.0	2.0
GROSFAC	0.9	0.7	0.8	0.9	1.0
IRAISE	2	2	2	2	1

In the surface pressure analysis, an initial RADMAX value of at least three is required for a good analysis of the large-scale anticyclones over the data-sparse oceans. On the other hand, a significant reduction in the maximum radius in later scans is necessary to define the centers of cyclones.

The effect of the variation of RADMAX can be seen in a comparison of the final assembled fields for two runs of the sea-level pressure analysis, shown in Figures III-3 and III-4. The assembly procedure was identical except that RADMAX was 2.5 in the first case and 3.5 in the second. The size of the high pressure centers over both the Atlantic and Pacific Oceans increases with the larger RADMAX, in better agreement with the observations. To use values of RADMAX greater than 3.5 was found to be counterproductive. Data information became so diffuse that the impact of the individual report was often lost.

The same figures demonstrate the need to reduce the maximum radius enough to accommodate the slightly smaller spatial scales of the cyclone centers. A cyclone located in the Pacific just north of 40° between 160° and 170°E is analyzed very differently in the analyses in Figures III-3 and III-4. During the final assembly scan, the maximum radius for the analysis in Figure III-3 was approximately seventy percent of the RADMAX in Figure III-4. Notice that the smaller radius permits the

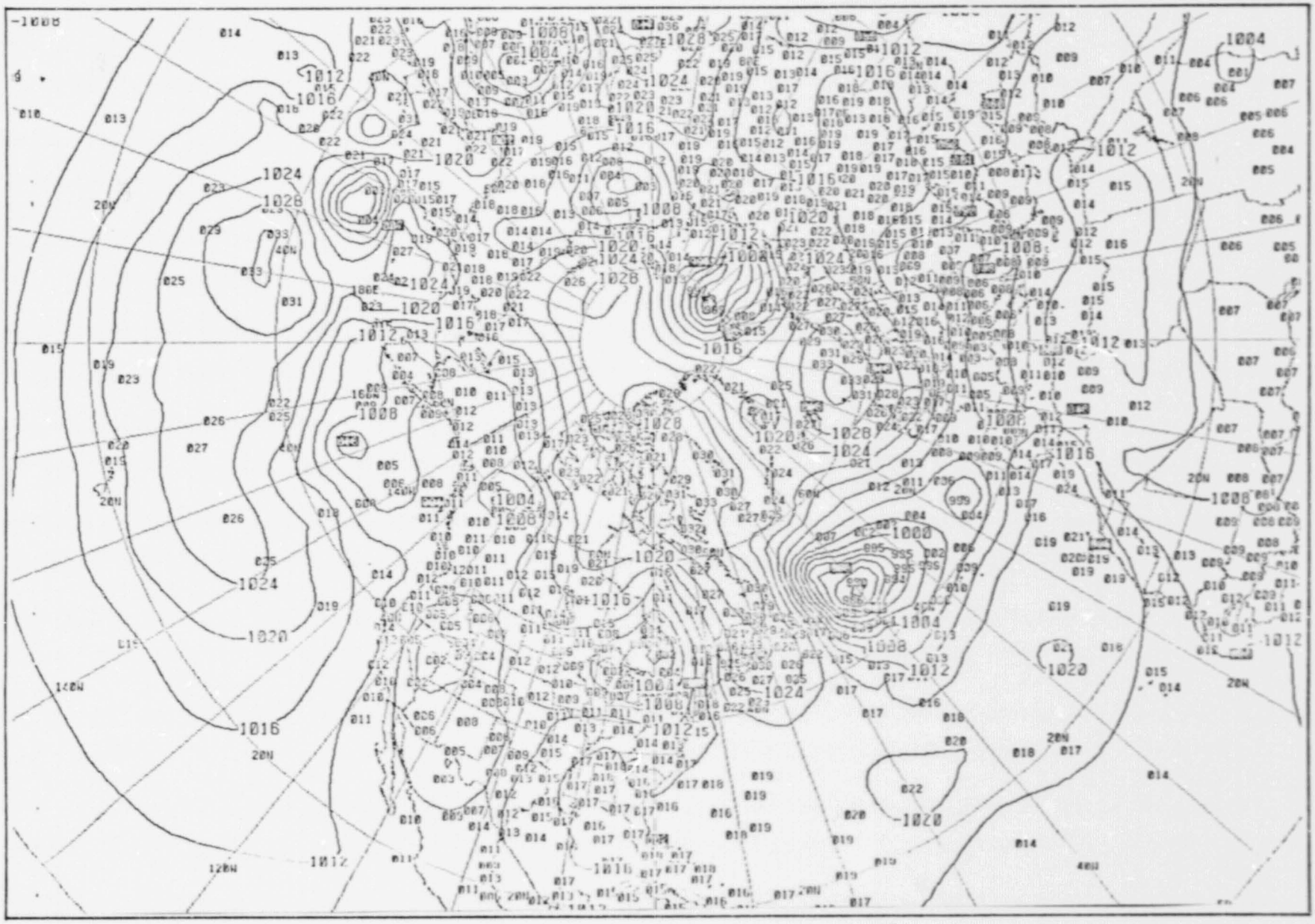


FIGURE III-3: SEA LEVEL PRESSURE ANALYSIS, FINAL ASSEMBLED FIELD, RADMAX=2.5

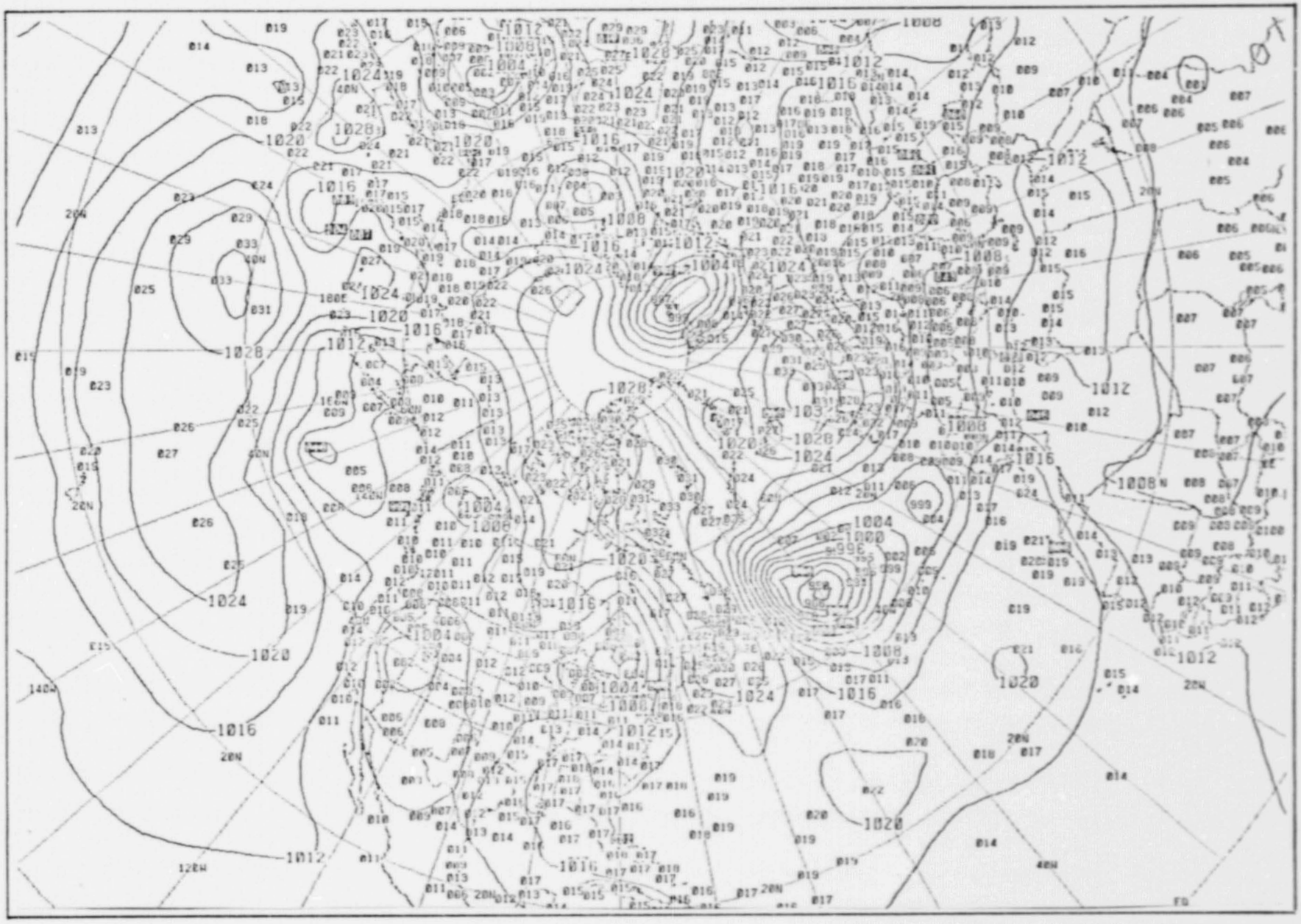


FIGURE III-4: SEA LEVEL PRESSURE ANALYSIS, FINAL ASSEMBLED FIELD, RADMAX=3.5

analyses to draw the 1004mb contour suggested by the data in the first case, whereas in the case of the larger radius, the data reports have been rejected, a result of the influence of the higher pressure reports surrounding them.

In conclusion, a definite range in the assembly radius is required to extract contributions from the data reports for the various synoptic spatial scales. A simple way to obtain this range is to make the radius reduction factor (RADFAC) small so that there is a significant change in the maximum radius values (RADMAX) between the first and last scans of the analysis program.

It may be noted in Table III-1 that RADFAC is greater (.8) in the surface pressure analysis than in the other analyses (.6); i.e., RADMAX does not decrease as much from one scan to another. This is due, in part, to the fact that RAD is included in the computation of the assembly radius of influence. For the surface pressure analysis, RAD is only half as great as in the other analyses, and if RADFAC is allowed to become too small, then the assembly radius introduces very small scales into the analysis which appear as noise relative to the large-scale systems.

B. Information Density (INFODEN)

The concept of the information density is based on the premise that where reports are closely spaced, the information from one report should not be spread beyond neighboring reports. Conversely, the reports in data sparse regions should be allowed to influence a relatively large area. Accordingly, the assembly radius for an observation is modified on the basis of the relative density of reported information in its vicinity.

The relative density at each point on the grid is determined in the subroutine, INFODEN. Here a report is assumed to influence the assembled field values at the gridpoints located within a certain radial distance of the report. Inside that circle, the contribution by a report to a gridpoint value is weighted by the distance between the report and the observation:

$$W_{I,J} = 1. - \left(\frac{D}{R}\right)$$

where W = weighted contribution by a report at a gridpoint

D = distance between the report and the gridpoint

R = the maximum radial extent of a report's influence

I,J = gridpoint indices

Once these contributions are calculated, they are added to the cumulative totals at those gridpoints. The same is done for each report. The sums at the gridpoints constitute the information densities.

The absolute density at a gridpoint is converted to a relative value INFOFAC, which lies between zero and one. The distribution of the densities is extremely skewed -- many gridpoints have little or no input from the current observations and a few have very high densities. For this reason, the absolute densities are normalized by a percentage of the maximum density. If this were not done, there would be very little variation in the PCT weighted on the basis of the information density (see Section II-A). The relative density INFOFAC cannot exceed one.

$$\text{INFOFAC}_{I,J} = \frac{\text{INFODEN}_{I,J}}{\text{DENMAX}} \qquad \text{and } 0 \leq \text{INFOFAC} \leq 1$$

where INFODEN = absolute information density at a gridpoint

INFOFAC = relative information density at a gridpoint

DENMAX = a percentage of the maximum absolute density for the grid.

I,J = grid location indices

DENMAX is defined as a value for which a given percentage of the total number of gridpoints has a higher density value. For example, PER (a program variable) is set to a percentage such as twenty percent. The gridpoints are divided into categories on the basis of their density values. DENMAX represents the category above which twenty percent of the gridpoints lie.

Recall that INFOFAC determines the magnitude of the term in the assembly radius which lies between RADMAX and RADMIN (see the beginning of Section III). When the density of information is high, i.e., when the reports are closely spaced, RADIUS will be small because the term approaches RADMIN. Similarly, as the data become less dense, INFOFAC is less and RADIUS approaches RADMAX, the maximum allowable value.

1. Tests of INFODEN

The first series of runs which included INFODEN used the product of the basic scan unit and the map factor, normalized by the standard meshlength to define the radius of the circular region within which the influence of a report could be felt. The radius of the INFODEN circle is expressed:

$$\text{INFODEN RADIUS}_M = \frac{IC * RAD * AMAP}{AMESH}$$

where RAD = basic scan unit

AMAP = map factor

AMESH = standard meshlength at the reference latitudes

M = data report index

IC = constant

IC was originally equal to one. However, in later runs, it was set to two or three since the assembly radius could be as large as three or four times RAD. It appears that with IC set to two, an INFODEN radius is defined which is representative of a report's significant influence region.

The percentage (PER) which would best serve as the cutoff maximum density limit, was examined also. Values for PER ranged from five to fifty percent. Currently, PER is set to twenty percent.

To find the maximum density limit, DENMAX, the absolute density values at the gridpoints are divided into intervals. Originally, we defined only ten intervals based on a linear scale. This proved to be too gross a division for determining a percentile in light of the extreme skewness of the distribution, so the number of intervals was increased to one hundred.

An example of the density distribution for the SST analysis based on a linear scale is shown in Figure III-5. Notice that most of the densities fall into the first dozen intervals. Given a PER of twenty percent, the DENMAX value falls into the fourteenth interval. In other words, twenty percent of the gridpoint locations have densities higher than the value corresponding to the fourteenth interval. If PER were greater than twenty percent, the selection of DENMAX would not be very precise because the overwhelming majority of the gridpoint values are concentrated in the lower categories. To overcome this problem, the interval scale was changed to a log 10 scale. The resulting distribution, Figure III-6, still shows a very large number of gridpoints in the first category, but this is to be expected. Many gridpoints are not influenced by any of the reports. Notice, however, that the rest of the gridpoints are spread more evenly through the remaining categories. In this case, with a PER of twenty percent, DENMAX is found in the fifty-fifth interval. A reasonable approximation of the density value representing even fifty percent of the reports could be made from this distribution.

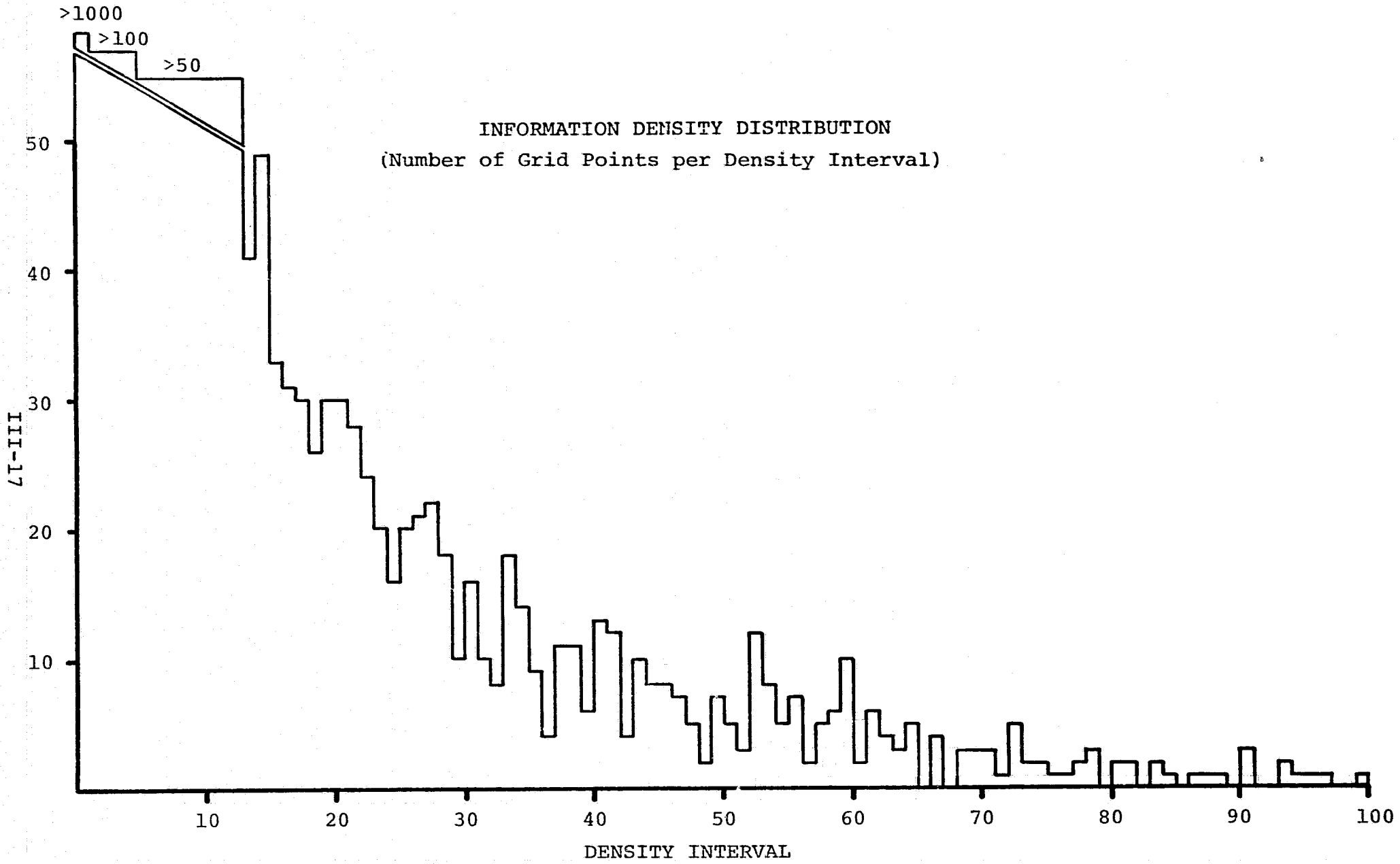


FIGURE III-5: HISTOGRAM OF INFORMATION DENSITY BASED ON LINEAR SCALE INTERVALS - SEA SURFACE TEMPERATURE ANALYSIS.

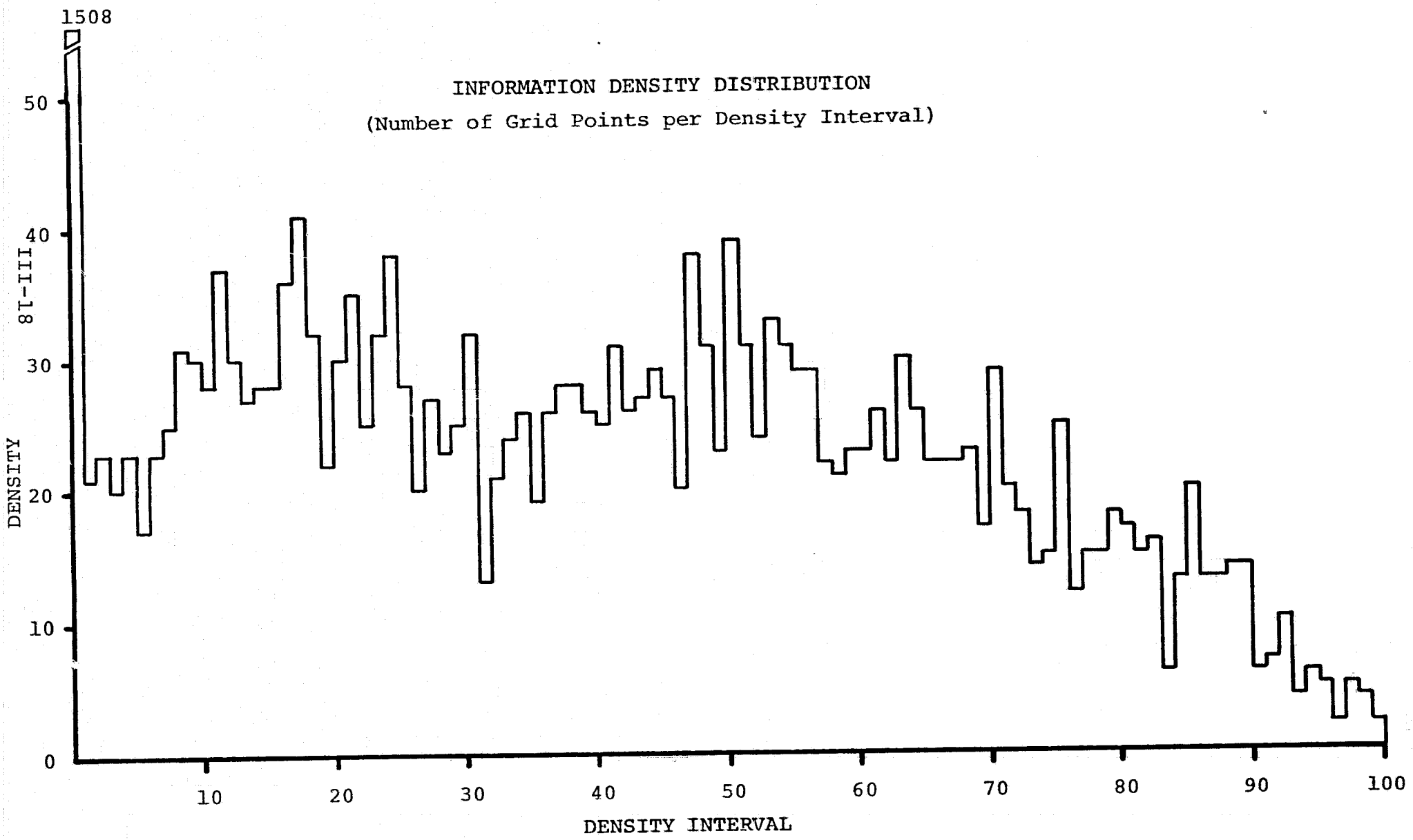


FIGURE III-6: HISTOGRAM OF INFORMATION DENSITY BASED ON LOG_{10} SCALE INTERVALS; SEA SURFACE TEMPERATURE ANALYSIS

One point needs to be made here regarding the categories. If PER is very small, such as five or ten percent, then a more accurate estimate of the representative DENMAX value is obtained using the linear scale. The reason is that the upper intervals on the log 10 scale encompass a much broader range of values than the same interval on a linear scale. The choice of scale, therefore, must be made with some consideration of the values of PER which will be used. Currently, the NASA-ODSI model employs the log 10 scale.

C. Gradient Factor

The assembly procedure can be further improved through the use of gradient information. Specifically, the radius of influence for each observation may be modified according to the guess-field gradients at the gridpoints closest to each report. In strong gradient regions, moreover, one wishes to limit the influence region of an observation to prevent the destruction of the existing gradients. The gradient factor, GRADFAC, is defined as:

$$GRADFAC_{I,J} = A - B * (GRAD_{I,J}/GRADMAX)$$

where GRAD = gradient with the maximum absolute value at the gridpoint closest to the report

GRADMAX = maximum gradient over the entire grid

A,B = constants, parameter-dependent

I,J = gridpoint indices

Recall that GRADFAC modifies the assembly radius. Its influence is best explained with an example. Let us assume that the constants, A and B, are set to 1.35 and 0.90, respectively. To find the assembly radius (RADIUS) for a report, the gradients in four directions from the gridpoint closest to that report are calculated. The gradient with the largest absolute value becomes GRAD in the expression above. GRAD is normalized by

the maximum gradient for the entire field, GRADMAX. If GRAD equals GRADMAX, their ratio is one and GRADFAC becomes 0.45. In other words, in an area of strong gradients, the assembly radius is significantly reduced from what it might otherwise have been. Of course, if the local gradients are weak, GRAD is close to zero, and GRADFAC approaches a value of 1.35. Here a report is allowed to influence a much greater region because there is less chance of interfering with the local gradients.

The constants in the GRADFAC expression vary somewhat from one analysis to another. They can be found in the Table of Assembly Variables, Table III-1.

D. Assembly Weights

The degree to which observational information is incorporated into an analysis depends on how much the reports are allowed to influence the field at the surrounding gridpoints. Weighted contributions from the nearby reports are combined to find the new assembled value at a gridpoint.

The formula for the new value is:

$$P_{I,J} = P_{I,J} + \frac{\Sigma[(W_K * DWT_K) * (W_K * DIF_K)]}{\Sigma (W_K * DIF_K)}$$

where P = the field value at gridpoint I,J

W_K = assembly weight for the Kth report

DWT_K = data weight for the Kth report (to be explained in the next section)

DIF_K = difference between the Kth report and the value interpolated from the field for its location

Clearly, the weights will largely determine the impact of the individual report.

The current NASA-ODSI model employs two criteria in the determination of the weights. The first is a distance dependence, i.e., the closer a report lies to a gridpoint, the greater is the weight assigned to its contribution. The second criterion is based on the magnitude of the curvature of the field in the

vicinity of the report. In regions of high curvature, there is a rapid spatial variation in the field. A weighting function which decreases very rapidly with increasing distance from a report is highly desirable under these conditions because it, in effect, localizes the influence of the report.

The first runs of the model used the Cressman weight function which is only distance dependent:

$$W_{I,J} = (R^2 - D^2)/(R^2 + D^2)$$

where W = weight of a report at gridpoint I,J

R = assembly radius of the Kth report

D = distance between the report location and the gridpoint I,J

Weights range from zero to one. Figure III-7 is a graph of the weights as a function of distance from the report. To increase the impact of the observations, a full weight of one was assigned to any gridpoint lying within a set distance of the report; e.g., within a distance of twenty percent of the report's assembly radius. Even with this technique, the analysis failed to depict the strength of the cyclones and anticyclones in the surface pressure analysis.

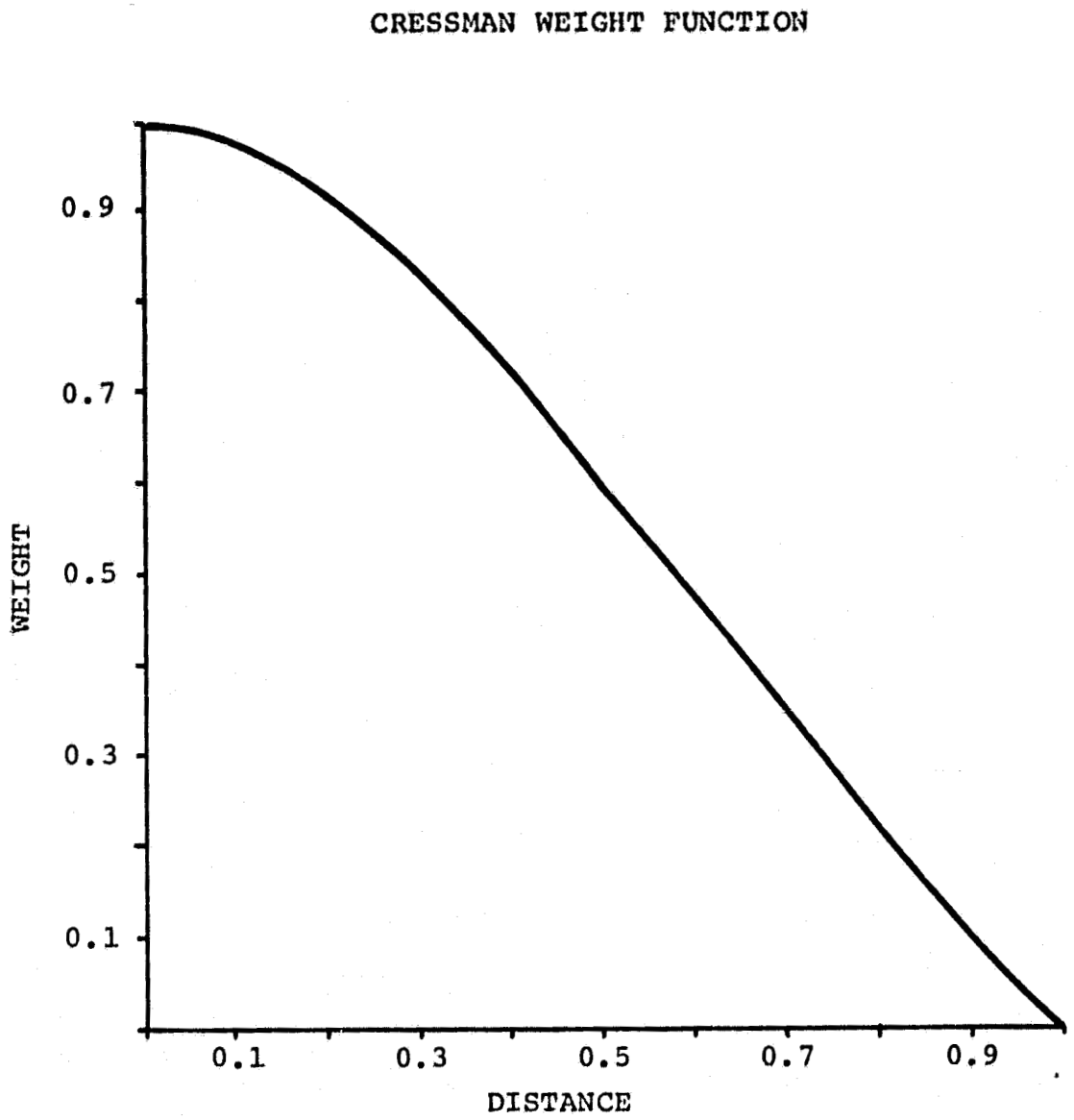


FIGURE III-7: GRAPH OF THE CRESSMAN WEIGHTS. WEIGHTS ARE A FUNCTION OF THE DISTANCE FROM A REPORT. THE DISTANCE HAS BEEN NORMALIZED BY THE RADIUS OF INFLUENCE.

The current formula allows weights to be a function of both the distance from the report and the curvature of the field. It is designed to ensure that data reports within a small distance of a gridpoint will receive a full weight of one. If this were not the case, it would be almost impossible for the analysis to exactly represent any report.

The weights are computed as:

$$W_{I,J} = (1. - RADPER)/(1. - FRAC)$$

where RADPER = D/R as defined above

FRAC = a function of the Laplacian curvatures
$0 \leq FRAC < 1$ and $0 \leq W \leq 1$.

When RADPER is less than FRAC, W is set to one.

The Laplacian value (WTLAPL), used in the determination of FRAC, is computed at the gridpoint closest to the report. WTLAPL is normalized by a percentage of the maximum Laplacian, WTLAPLM. The ratio of WTLAPL and WTLAPLM cannot exceed one so that FRAC lies between zero and one.

$$FRAC = 1. - (WTLAPL/WTLAPLM)$$

where WTLAPL = Laplacian at the gridpoint closest to a report

WTLAPLM = PERLAPL * WTLAPMX

PERLAPL = a fraction less than one

WTLAPMX = maximum Laplacian in the first-guess field

In the program, FRAC is limited by a maximum and minimum value. The maximum value, FRACMAX, represents the greatest extent within the circle of influence to which a weight of one might be assigned. (Recall that when RADPER is less than FRAC, the weight is set to one.) The minimum value, FRACMIN, defines the radius of a small circle within which the weight is always one.

The slope of this weight function can be seen in Figure III-8 for the case where FRACMAX and FRACMIN are 0.5 and 0.1. The function is actually a family of curves limited to the region between the maximum and minimum FRAC values.

Tests were run in which a range of values was assigned to the variables FRACMAX, FRACMIN and PERLAPL. Various combinations were examined for each of the different analyses; the values for these variables in Table III-1 appear to be the optimum values for the current forms of the analyses.

Several tests were made in which the Laplacian dependence was removed from the weighting function. These weights were solely distance dependent, but differed from the Cressman weights. Two formulae were examined:

$$W_{I,J} = 1. - \text{RADPER}$$

$$W_{I,J} = 1. - (\text{RADPER})^2$$

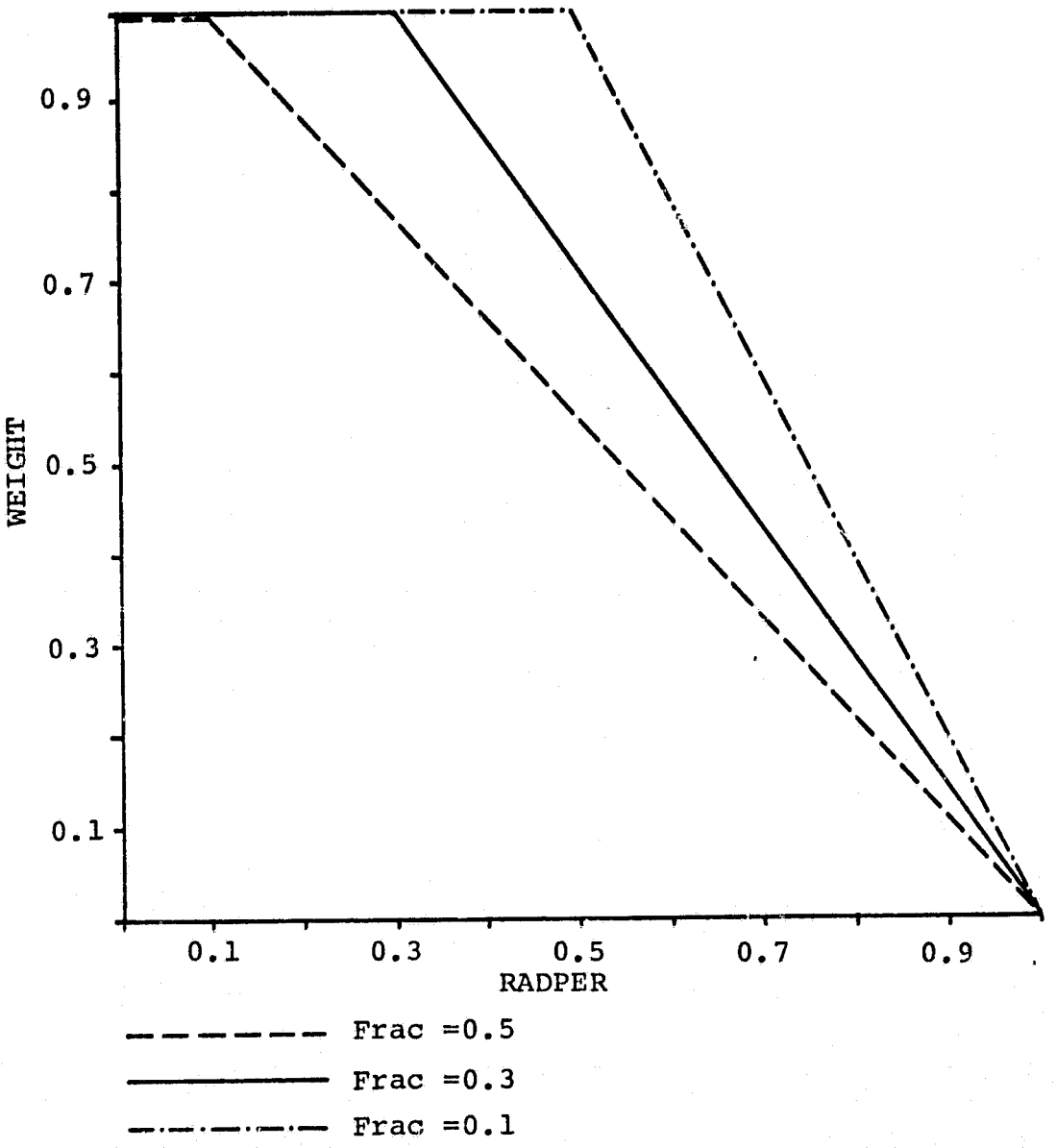


FIGURE III-8: DISTRIBUTION OF WEIGHTS AS A FUNCTION OF THE RADIAL DISTANCE FROM AN OBSERVATION. IN THIS CASE, FRACMAX=0.5 AND FRACMIN=0.1. WEIGHTS LIE BETWEEN THESE LIMITS.

All variables are defined as before. The lost sensitivity to the curvature of the field was evident in the assembled fields of these analyses, and the Laplacian dependence was reinstated. Compare Figures III-9 and III-10, the surface pressure fields after the third scan of the assembly process. The weights in the first analysis include the Laplacian dependence. For the analysis in Figure III-10, the Laplacian dependent weight function has been replaced by the second weight formula above. Several differences can be seen. In the second analysis, the central contour in the Pacific does not extend as far as in the first case, the cyclone in the Pacific is not as deep, nor does the 1000mb contour appear over the Indian subcontinent, a feature clearly suggested by the data.

It should be noted that the weight W, as described above, is in effect squared when used to calculate the assembled field values. Returning to the formula for the new assembled value,

$$P_{I,J} = P_{I,J} + \frac{\Sigma[(W_K * DWT_K) * (W * DIF)]}{\Sigma (W_K * DWT_K)}$$

notice that W appears twice in the numerator and once in the denominator. This formula was compared to another method of determining the assembled field values. The second formula,

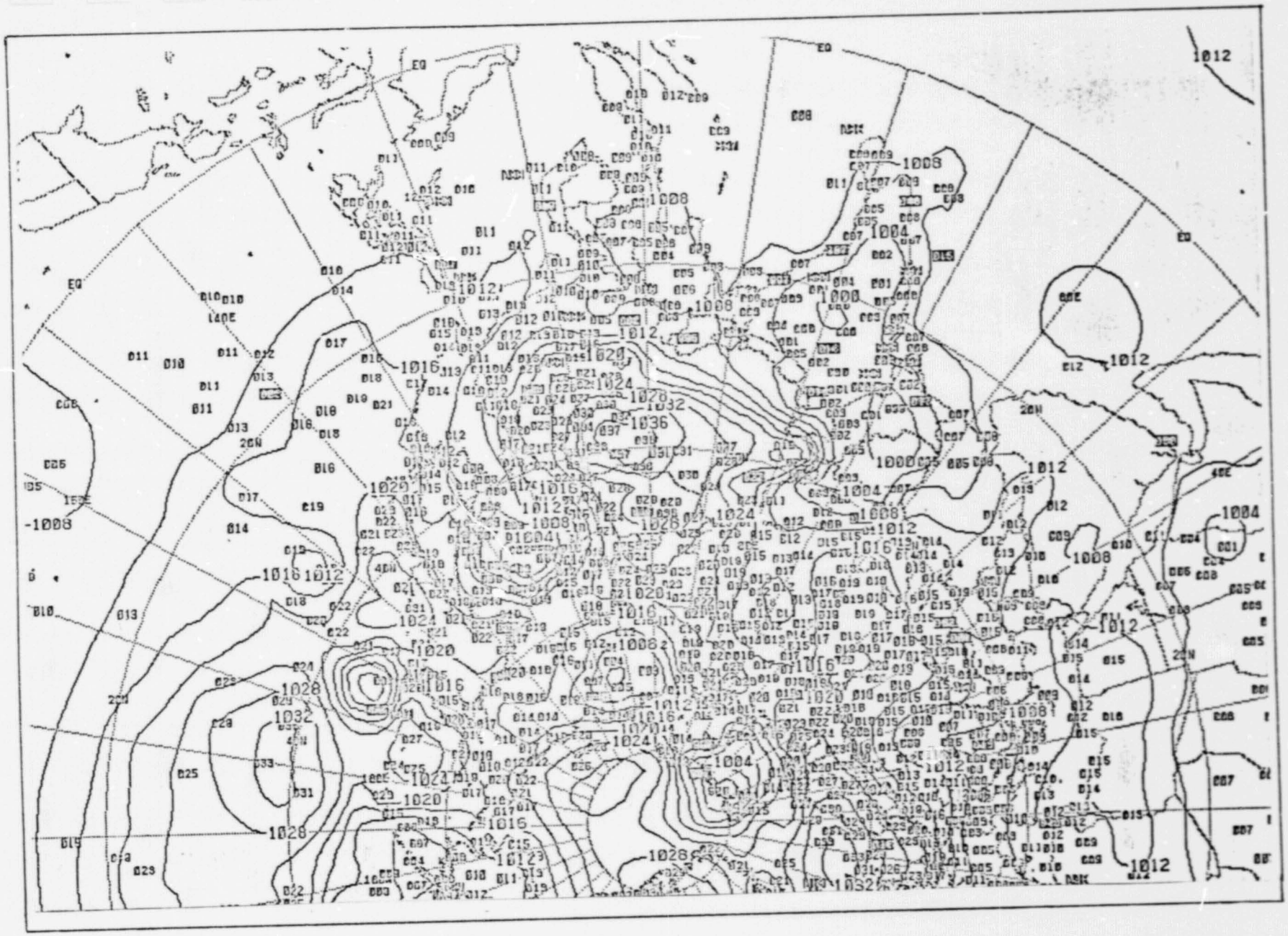


FIGURE III-9: SURFACE PRESSURE ANALYSIS, FINAL ASSEMBLED FIELD WITH LAPLACIAN DEPENDENT WEIGHTS.

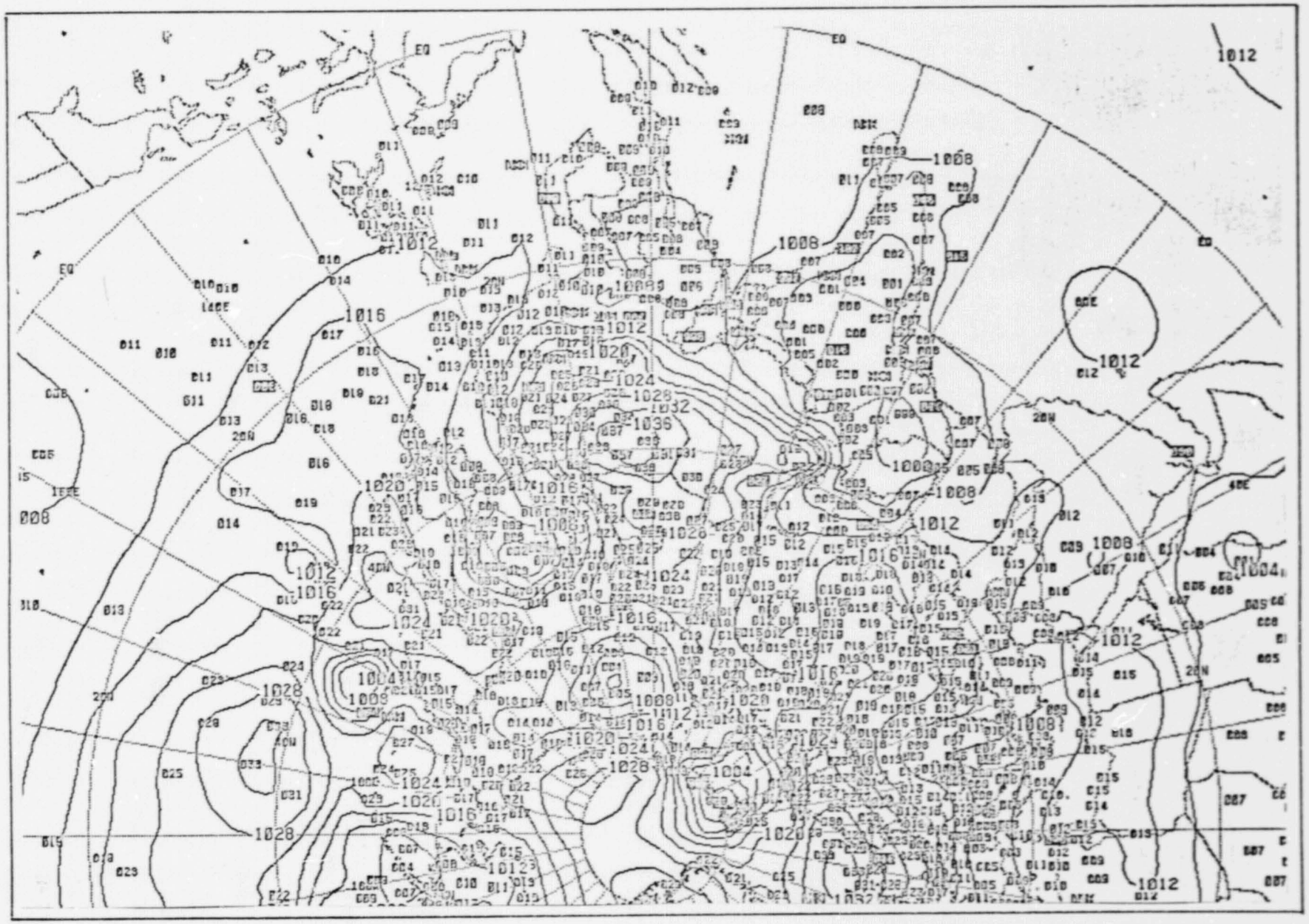


FIGURE III-10: SURFACE PRESSURE ANALYSIS, FINAL ASSEMBLED FIELD WITH NO LAPLACIAN DEPENDENCE.

$$P_{I,J} = P_{I,J} + \frac{\Sigma (W_K * DWT_K * DIF_K)}{\Sigma\ DWT_K}$$

does not work as well as evidenced by the change in RMS values, whether in the scalar or vector fields. For example, in the surface pressure analysis, the RMS value for the analysis was reduced by ten percent with the introduction of the W^2 formula. In the wind analysis, the same formula change resulted in an average reduction in the RMS value of twenty-two percent at each of the twelve levels.

Much of the preceding discussion of weights applies only to the scalar analyses. The weights in the wind analysis are derived with the Cressman formula only.

E. Data Weights and Reevaluation

The original method of data reevaluation was rather complex. It basically involved re-solving the PCT equations in the vicinity of an observation, omitting the report under reevaluation. Depending on the change in the analysis field, the observation's weight was restored to the original data weight or assigned some lesser value. This procedure was computationally very expensive, accounting for more than one-half the computer time for the analysis.

A new method of reevaluation has been installed in the scalar analysis programs. Initially, each observation is assigned a subjective data weight, DWT. At the end of each scan, the data weight is evaluated in the subroutine REVALWT according to a report's departure from the interpolated value for its location in the new analysis field. If there is a significant difference between the two, the data weight may be lowered. Should the analysis better agree with an observation whose weight was reduced in an earlier scan, the original data weight may be restored.

In the current sea surface temperature and sea-level pressure analyses, all reports are assigned the same initial data weight. Smaller data weights had been assigned to ship reports in the pressure analysis previously, but this was

found to be unnecessary. In the upper air temperature, height and wind analyses, a different basic data weight is assigned at each level. The role of the data weight in the reevaluation algorithm necessitates the separate specification. Only the wind analysis differentiates between types of reports; satellite-derived winds have half the weight of the conventional wind observations.

The reevaluation process has evolved to where the differences tolerated between the reports and the analysis fields are smaller with each new scan. It is felt that the new analyzed field should better accommodate the reports than the previous field. Observations which continue to show distinct departures from the field must be considered unrepresentative or unreliable data, and their importance reduced accordingly. Of course, in the first scan, all observations carry their full weight.

In the wind analysis program, the u and v wind components are processed separately. However, only one data weight is assigned to the two. Reevaluation of the wind reports is based upon the magnitude of the wind vector.

A report's weight is reevaluated if the parameter, REVAL, exceeds a constant critical value, CRIT. The parameter REVAL is expressed as:

$$REVAL_K = \frac{FP * REFAC * (DIF)^2}{ODWT_K}$$

where FP = the current scan number

REFAC = reevaluation factor, a constant specified for each type of analysis

DIF = the difference between a report and its interpolated value

ODWT = original data weight

K = subscript indicating the report number

Since the factor FP increases on succeeding scans, the differences tolerated between a report and the field at its location is less each time. The value of REFAC depends on the type of analysis. If REVAL is less than or equal to CRIT, the new data weight is set to the original value, ODWT. If REVAL is greater, the new data weight is determined as:

$$DWT_K = \frac{CONST * ODWT_K}{1. + REVAL}$$

where CONST = a constant

DWT = the new data weight assigned to a report

ODWT = the original data weight for the Kth report

The choice of constants for both CRIT and CONST is important. CRIT represents the upper limit of acceptable departures from the guess field. Beyond that a report's data weight is reduced. If the weights of too many reports are reduced, the analysis will fail to respond to the observational input. The value of CONST is also critical, for it determines the extent to which the data weight is changed.

The values of CRIT, CONST, and REFAC underwent many changes during the testing of the analyses as did the initial values of the data weights for each analysis. Adjustments were made so that only a few weights would be reduced and yet the unrepresentative reports would not unduly influence the analysis. The values for the first three factors appear in the table of assembly variables, Table III-1. The data weights for each analysis are given separately in Table III-2.

TABLE III-2: DATA WEIGHTS FOR VARIOUS ANALYSIS TYPES

Level	Sea Surface Temperature	Surface Pressure	Temperature	Height	Wind
SFC/1000MB	2.5	5.0	3.0	30.0	20.0
950	-	-	3.0	30.0	20.0
900	-	-	3.0	30.0	25.0
850	-	-	3.0	30.0	25.0
700	-	-	3.5	45.0	27.0
500	-	-	3.0	60.0	35.0
400	-	-	3.5	65.0	35.0
300	-	-	3.5	75.0	35.0
250	-	-	4.0	80.0	40.0
200	-	-	3.0	80.0	40.0
150	-	-	4.0	80.0	30.0
100	-	-	3.5	80.0	30.0

F. Reject Criteria

Whether an analysis begins with a poor or reasonable first-guess field, there will be some observations which deviate so greatly from the field that they must be considered bad reports. To incorporate them into the field would lead to a totally invalid analysis. Thus, a realistic objective criterion for rejecting such reports is essential.

Several procedures were tried in the NASA-ODSI analysis. The first involved the selection of a constant value for a gross reject criterion, GROS. It represented the critical amount by which a report could differ from its interpolated field value without being rejected. There is a problem, however, in using a constant value for the entire hemisphere. The magnitude of the variations in the fields is substantially less in the tropics than in the middle latitudes. It would seem, then, that the differences tolerated between the reports and the fields should also be less. A reject criterion that would vary with latitude was obviously needed.

A convenient way to incorporate the latitudinal dependence is to use the image plane factor, AMAP. The new reject criterion, GROSTOL, is defined as:

GROSTOL = GROS/(AMAP**IRAISE)

where GROSTOL = the latitudinally-dependent reject criterion

GROS = constant gross reject value

AMAP = map factor, a function of latitude

IRAISE = an exponent which is set to one or two, analysis-dependent

As one moves closer to the equator, the map factor (AMAP) increases. Correspondingly, the value of the gross tolerance, GROSTOL, becomes less.

The advantage of using GROSTOL versus GROS is illustrated in Figures III-11 and III-12. Both figures represent final analyses of the sea-level pressure.

In the first, the reject criterion was based on GROS; the second employed GROSTOL with IRAISE equal to 2. The areas surrounding the Indian subcontinent show a pattern of highs and lows which is unrealistic for the tropics (Figure III-11). The problem is eliminated with the use of GROSTOL (Figure III-12) which forces the rejection of many more reports in the area.

An additional change has been made in the procedure. Rather than allowing GROS to remain constant, its value is reduced for each new scan. It seems reasonable to make the reject criterion successively more stringent since with each cycle the analysis should be closer to fitting the observations.

FIGURE III-11: SURFACE PRESSURE ANALYSIS REJECT CRITERION BASED ON GROS.

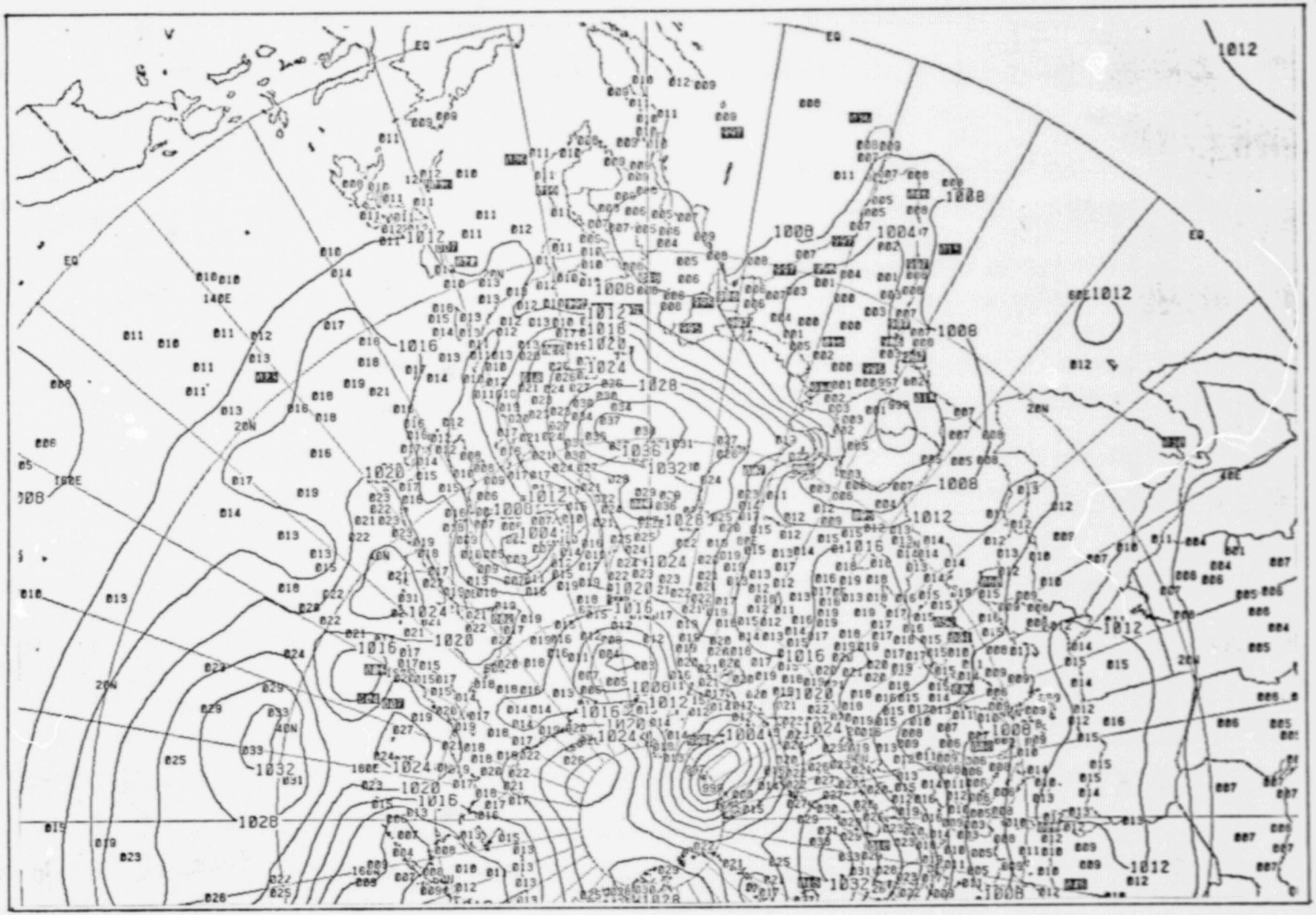


FIGURE III-12: SURFACE PRESSURE ANALYSIS LATITUDE DEPENDENT REJECT CRITERION.

The new value of GROS is calculated as a fraction of the old value:

$$GROS_M = GROSFAC * GROS_{M-1}$$

where GROSFAC = a constant less than one, parameter-dependent

M = a subscript denoting cycle number

The choice of the initial value of GROS can have a marked effect on the resulting analysis. As an example, compare two analyses of the sea surface temperature which were identical with the exception that for the first (Figure III-13) the initial value of GROS was seven, and for the second (Figure III-14) GROS was three. Not unexpectedly, many more reports are rejected with the smaller GROS value (20% versus 6%) but the field is much smoother. For this particular analysis, the smaller tolerance may be better.

In the upper air height and wind analyses, the initial values assigned to GROS are level-dependent. Larger GROS values are required in the upper levels because the range of field values at these levels is so much greater. In the temperature analysis, one initial value for GROS is assigned to all levels since the variation in the temperature field at any level is relatively small. As in the other analyses, GROSTOL is determined at each level as explained above, and the value of GROS at a given level decreases with each scan. Initial values of GROS appear in Table III-3; GROSFAC and IRAISE are shown with the assembly variables in Table III-1.

C-2

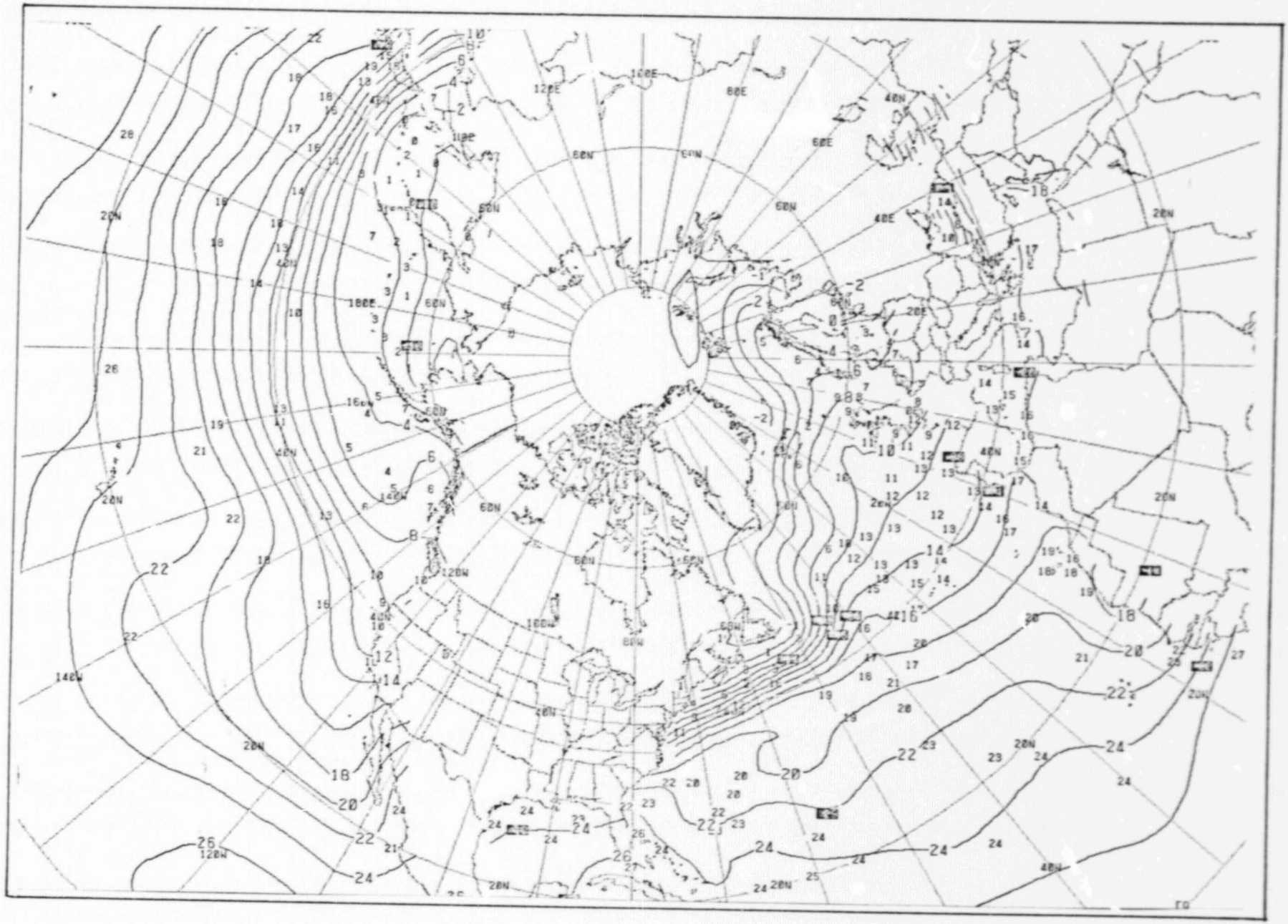


FIGURE III-13: SEA SURFACE TEMPERATURE ANALYSIS, GROS=7.

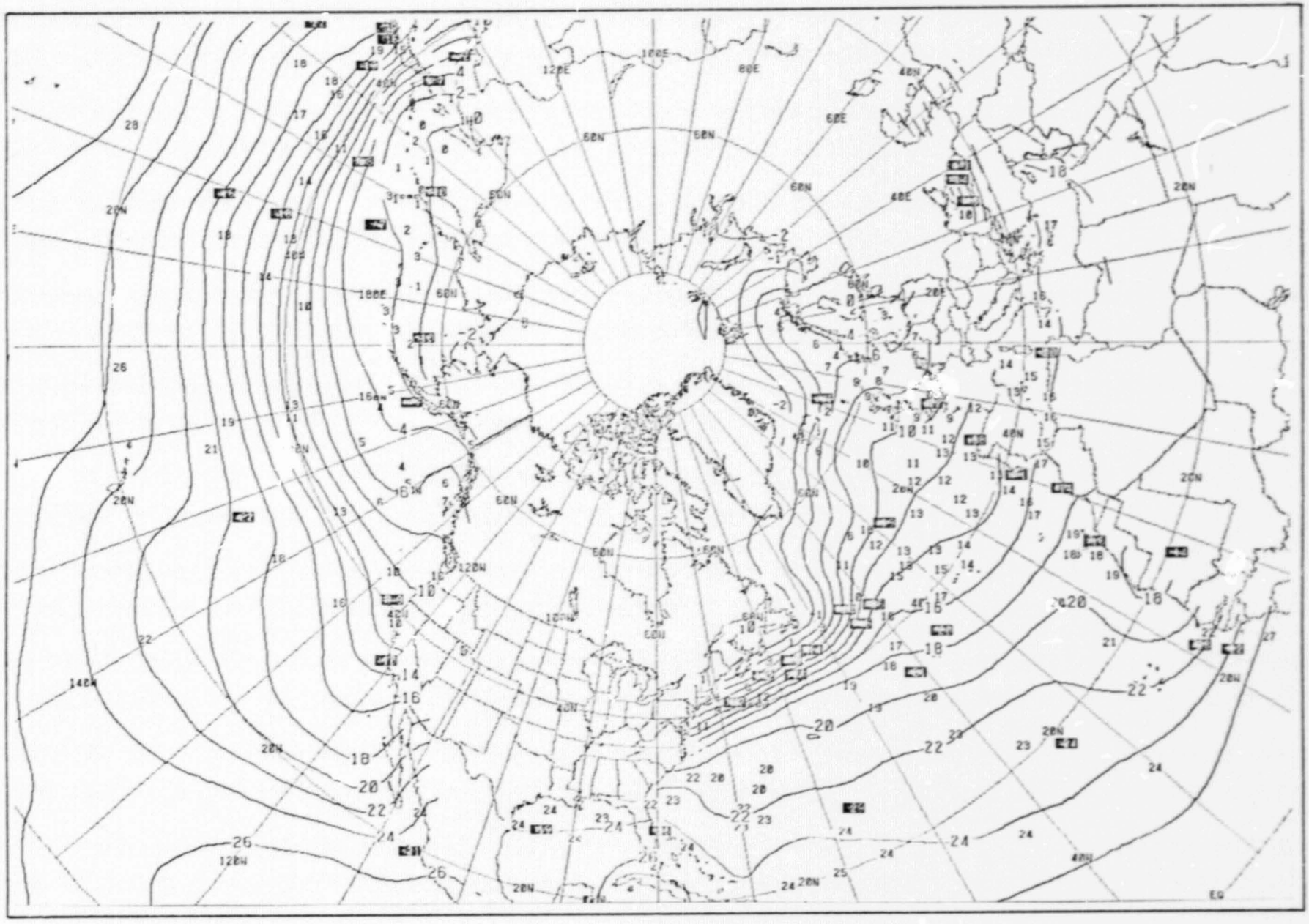


FIGURE III-14: SEA SURFACE TEMPERATURE ANALYSIS, GROS=3.

TABLE III-3: GROSS REJECT VALUES (GROS) FOR VARIOUS ANALYSIS TYPES

Level	Sea Surface Temperature	Surface Pressure	Temperature	Height	Wind
SFC/1000MB	7.0	15.0	13.0	50.0	15.0
950	-	-	13.0	50.0	15.0
900	-	-	13.0	50.0	20.0
850	-	-	13.0	60.0	21.0
700	-	-	13.0	75.0	23.0
500	-	-	13.0	100.0	26.0
400	-	-	13.0	100.0	28.0
300	-	-	13.0	120.0	35.0
250	-	-	13.0	120.0	37.0
200	-	-	13.0	120.0	40.0
150	-	-	13.0	120.0	35.0
100	-	-	13.0	120.0	30.0

IV. BOGUS REPORTS

In a situation where the number of observations is severely limited and the first guess is poor, the analysis will not adequately portray systems without additional input. This is especially critical in the case of a developing and/or fast moving system. Improvements in an analysis can be made with the use of bogus reports. These supplemental reports are introduced by a trained analyst whose decisions to "bogus" may be based on new information or simply on experience with certain types of weather situations.

In the program code, bogus reports are handled differently from the observational reports. First, the data weight assigned to a bogus report is higher and cannot be reduced in the reevaluation routine. Of course, the report cannot be rejected. Finally, the assembly radius for a bogus report is always the maximum allowed for a particular scan.

A situation involving a developing cyclone in the North Atlantic can be used to demonstrate the bogusing procedure. Figure IV-1 shows an analysis of the sea level pressure. Without additional information, the analysis shows a 996mb contour which fails to enclose the much lower 986mb observation. To ascertain the effect of bogusing, a series of runs was made

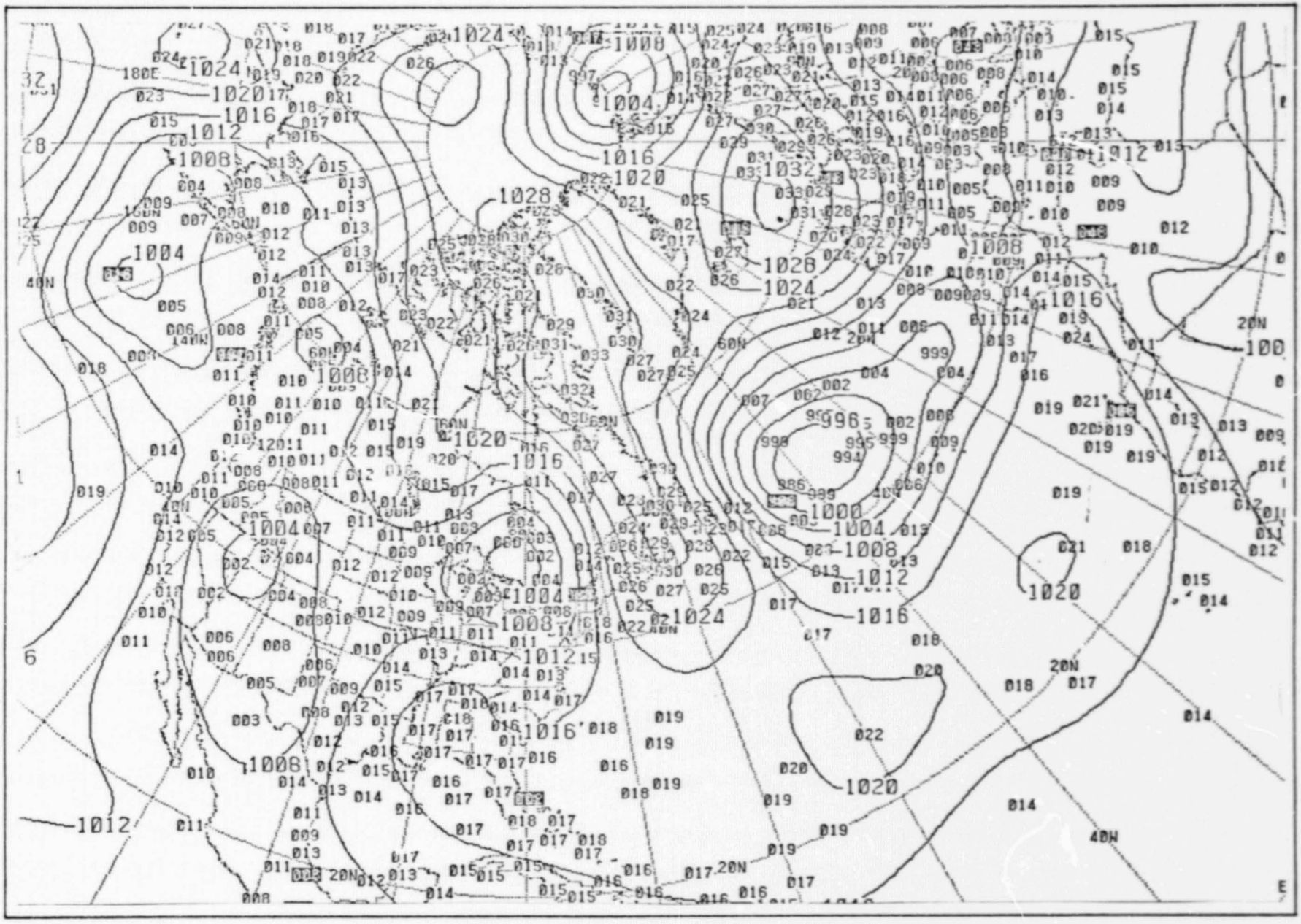


FIGURE IV-1: SEA LEVEL PRESSURE ANALYSIS, NO BOGUS REPORTS.

in which an increasing number of 980mb reports were introduced at locations near the cyclone center. The data weights for the bogus reports were three times the normal data weight. The impact of the bogus reports on the final pressure analysis is shown in Figures IV-2 through IV-5.

With the introduction of one bogus report, the North Atlantic cyclone deepens. Note that the area enclosed by the 996mb contour in Figure IV-2 is much greater than in Figure IV-1. A second bogus report further strengthens the cyclone as it now shows a 992mb contour (Figure IV-3). Note, too, that this results in the rejection of two more observational reports near the low center. With an additional bogus report, the low center shifts somewhat while encompassing a slightly larger region, as seen in Figure IV-4. Apparently, the shift of the center has allowed the 1004mb report to be retained. The last figure of the series shows the analysis using four bogus reports (Figure IV-5). In this figure, the cyclone has a 988mb contour.

In this particular situation, the bogusing effort is very much hampered by the implied strength of the gradients. It is difficult to analyze such strong gradients on the relatively coarse 63x63 grid. The influence of nearby reports can be a major factor, for although their influence may decrease on

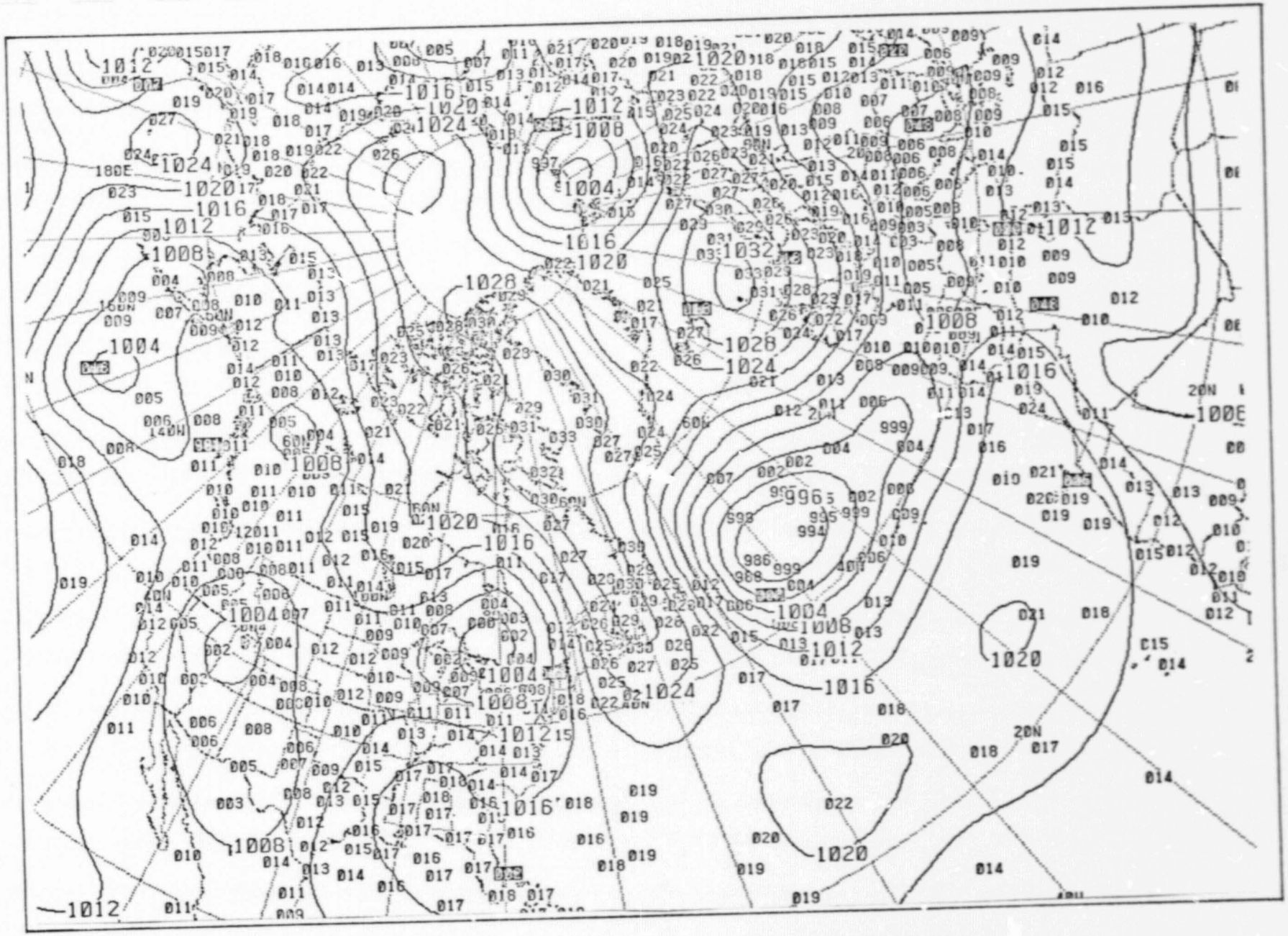


FIGURE IV-2: SEA LEVEL PRESSURE ANALYSIS, ONE BOGUS REPORT.

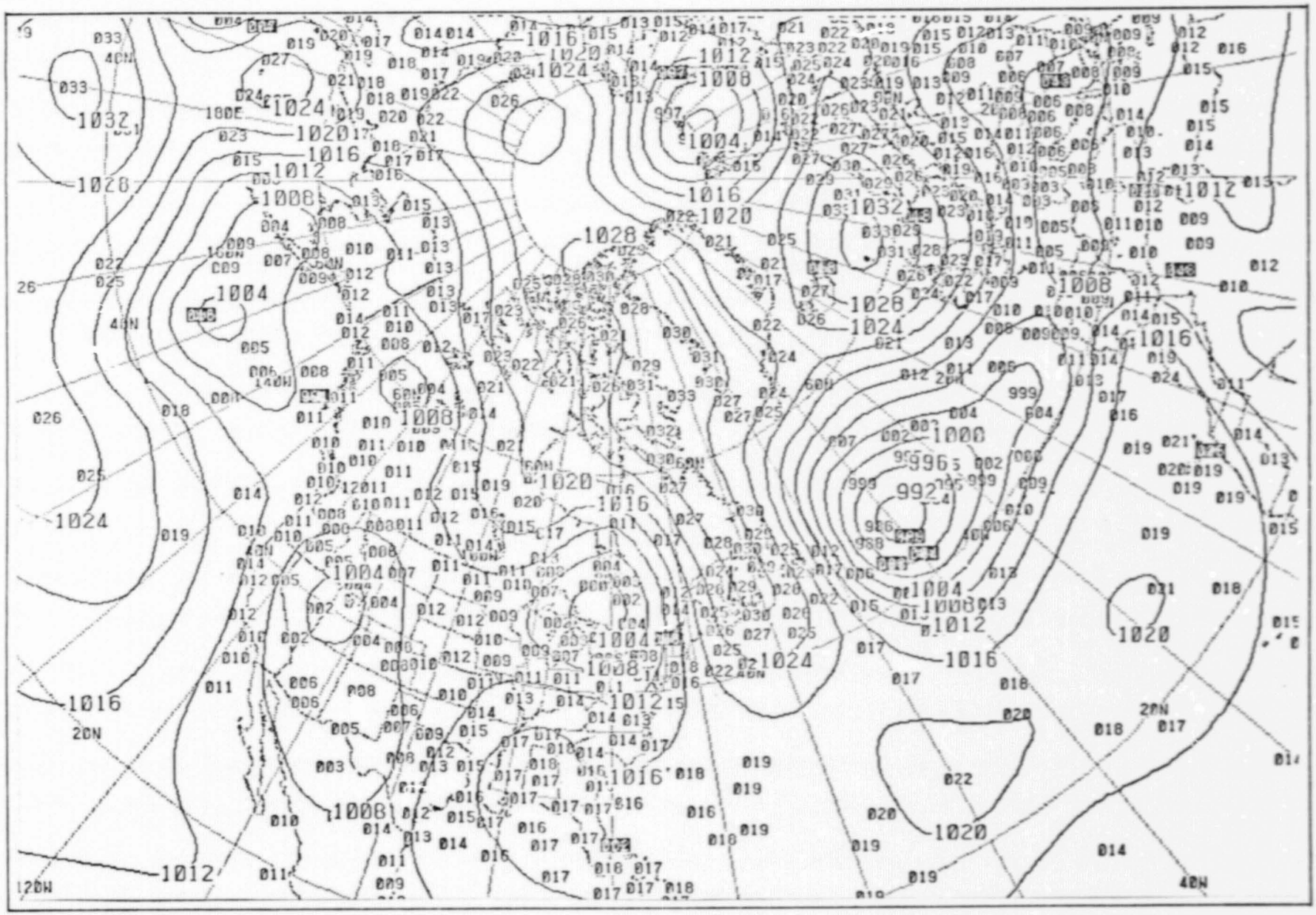


FIGURE IV-3: SEA LEVEL PRESSURE ANALYSIS, TWO BOGUS REPORTS.

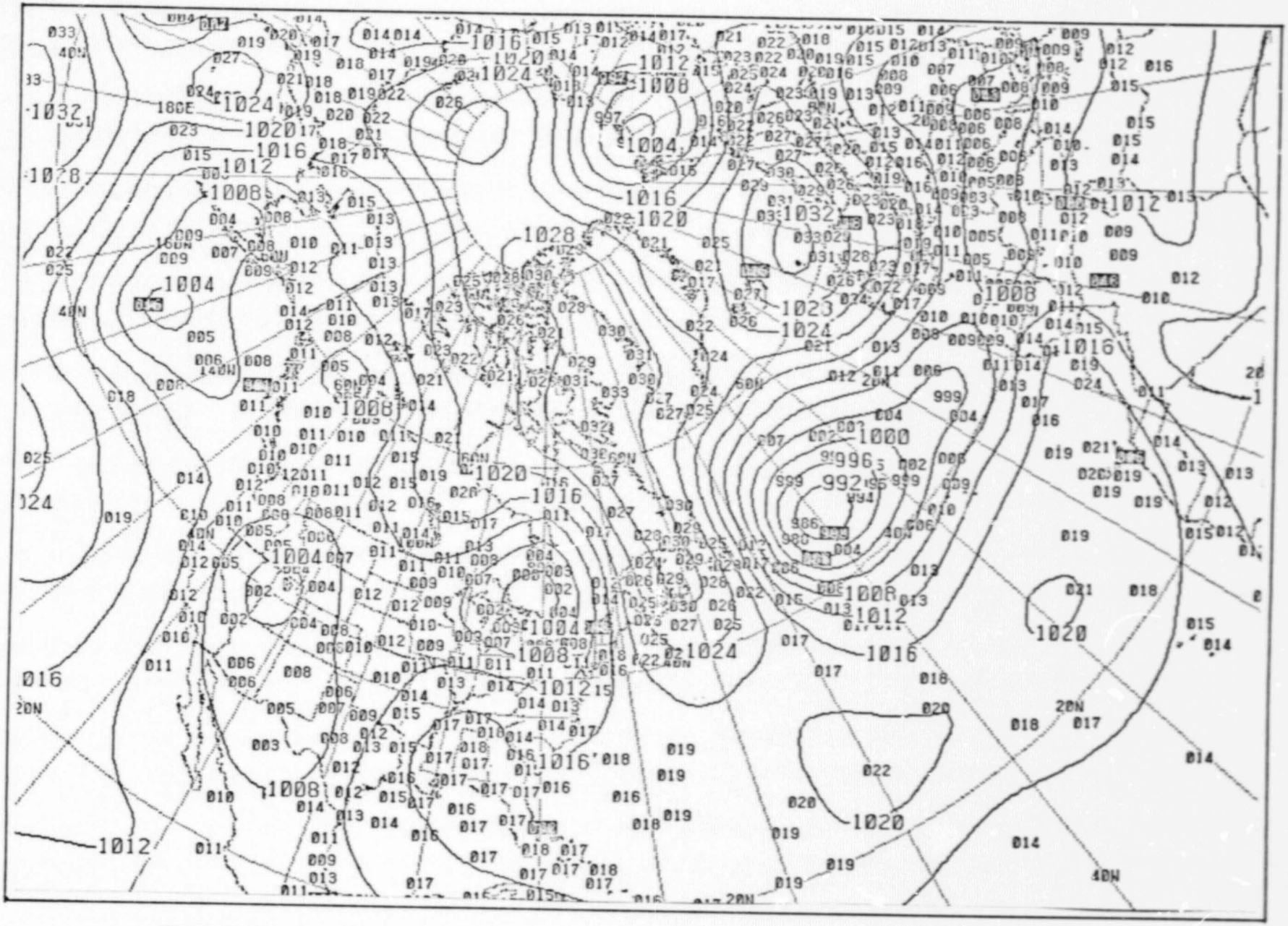


FIGURE IV-4: SEA LEVEL PRESSURE ANALYSIS, THREE BOGUS REPORTS.

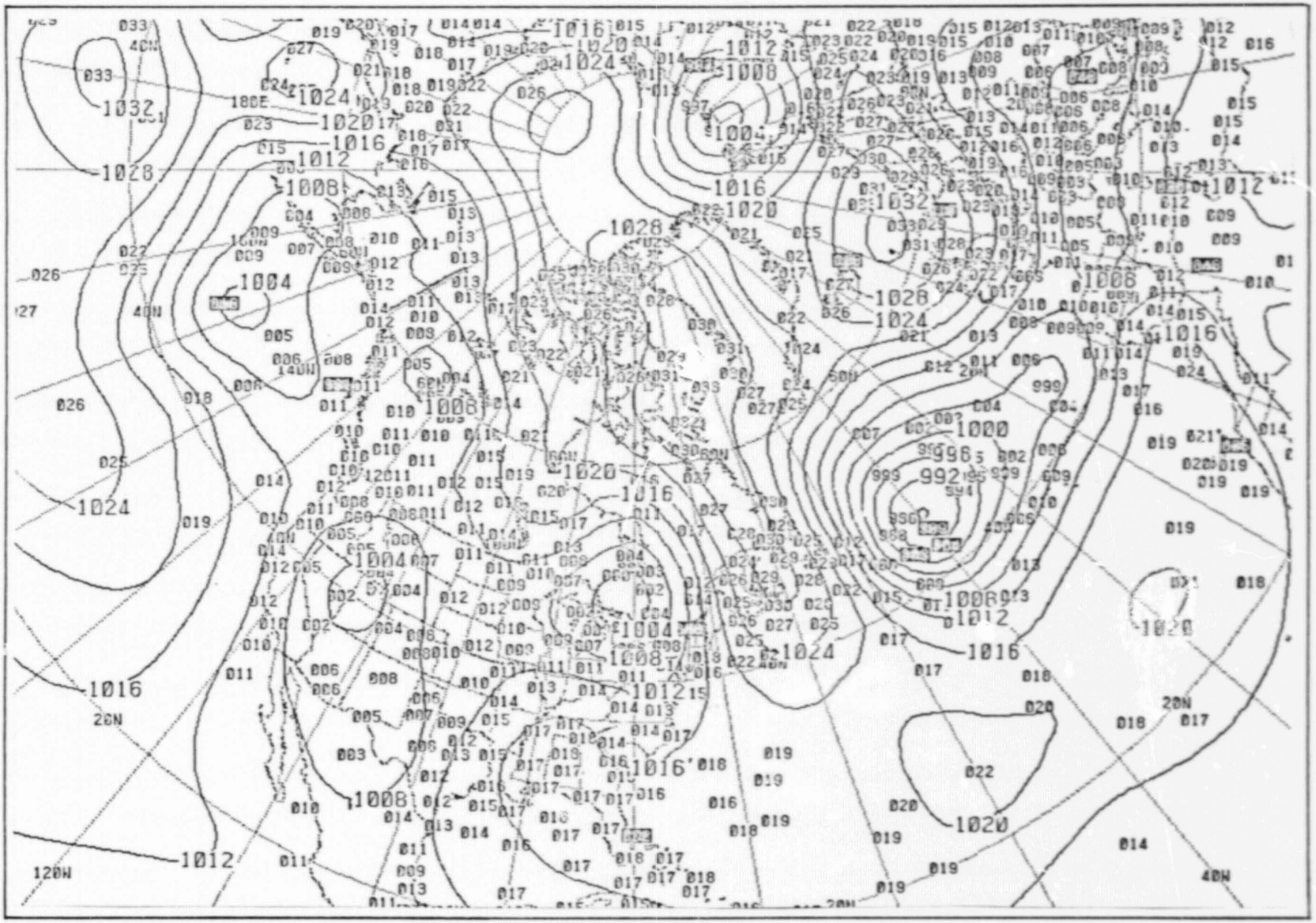


FIGURE IV-5: SEA LEVEL PRESSURE ANALYSIS, FOUR BOGUS REPORTS.

successive scans, they can have a significant impact on the initial scan. The PCT constraints which tend to bring the field back to the original (weaker, in this case) gradients are also a factor. Despite these difficulties, the bogusing technique can improve an analysis. The impact of the bogus reports can be controlled through:

1. The number of bogus reports.
2. The location of the reports in the field and with respect to other bogus reports.
3. The relative size of the data weight assigned to bogus data.

V. WIND INFORMATION IN HEIGHT ANALYSES

In an analysis, it is important to include as much of the available observational data as possible. This is especially true in an upper-air analysis where data are limited. In the original NASA-ODSI upper-air height analysis, gradient information derived from wind information was incorporated into the analysis via the PCT constraints. The geostrophic relationship was used to compute height gradients consistent with the reported winds. These height gradients replaced the x and y axis gradients computed from the first-guess field at the gridpoint nearest the observation. If the height gradient derived from the wind observations was greater than a specified limit, the guess-field gradient was retained. If more than one report influenced one gridpoint, the derived gradient nearest in magnitude to the gradient found in the first-guess field was substituted.

These procedures have been refined to better utilize the wind observations. First, the wind reject criteria has been modified to use a method which is physically based rather than one based upon an arbitrary constant gradient limit. The geostrophic wind speed derived from the first-guess height field is compared with the reported wind observation. The

wind observations are rejected if they differ significantly from the height derived winds. Reject limits which are level-dependent and consistent with those used in the wind analysis. Therefore, the wind observations which are used to influence the height analysis will be included in the wind analysis.

Originally, the derived height gradients were only incorporated into the gradients in the x and y directions. Tests indicated that this did not exert a strong enough influence on the PCT equations to effect much change. Recall that the PCT equations also include the diagonal gradients and the Laplacian. Accordingly, the appropriate diagonal gradients and Laplacian, derived from the winds, were substituted into the height field. This has greatly improved the sensitivity of the analysis. If more than one observation influences a particular gridpoint, the derived gradients and Laplacians from the various reports are averaged.

Some examples will show the advantage of including the wind observations. Figures V-1 through V-4 show the 76042212Z 900mb height analysis under different conditions. Figure V-1 shows the analysis with no wind observations; Figure V-2 shows the analysis with observations included only in the x and y direction height gradients; and two charts, Figures V-3 and V-4, with wind observations included in all PCT constraints. (One

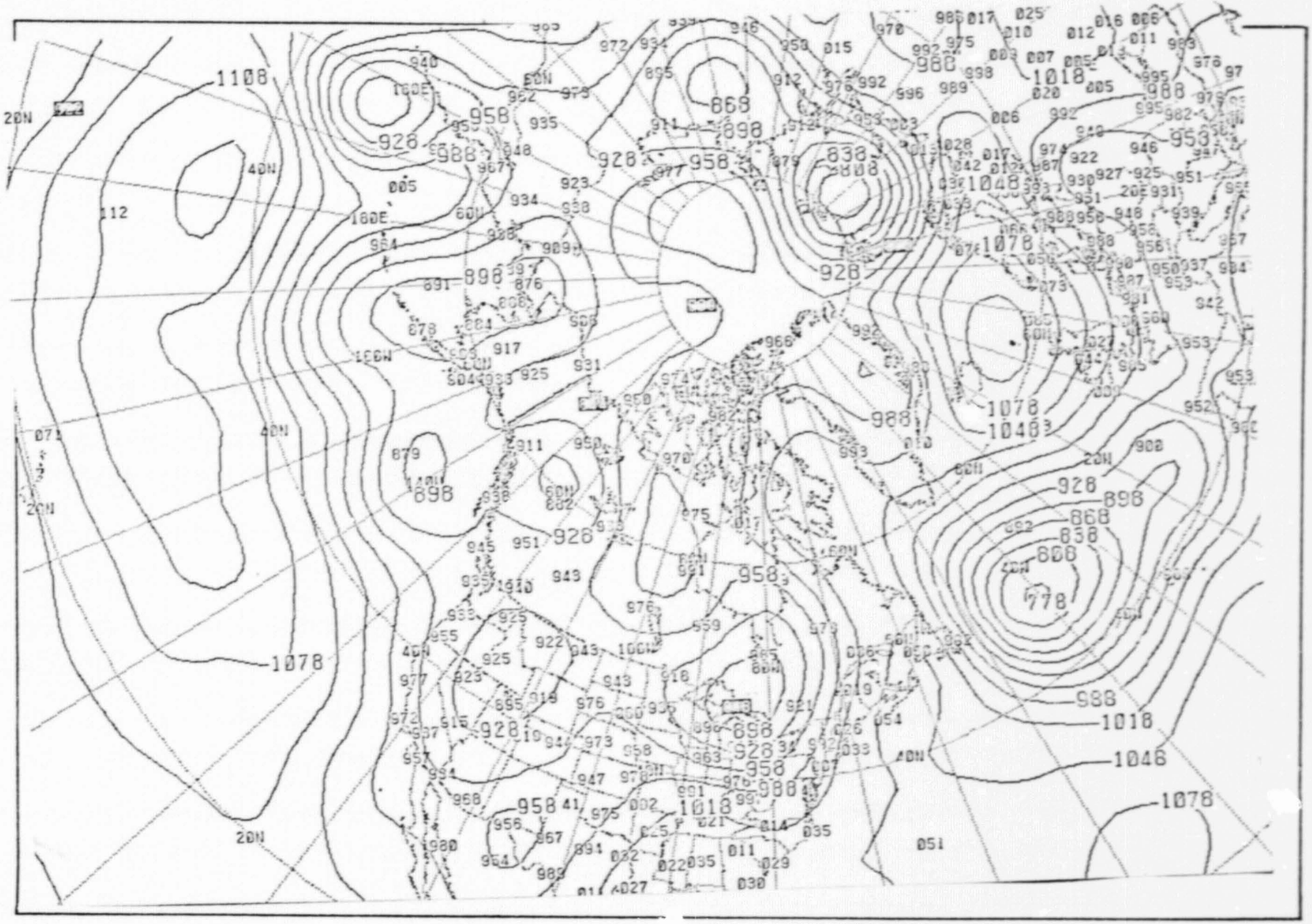


FIGURE V-1: 900MB HEIGHT ANALYSIS WITH NO WIND OBSERVATION GRADIENT INPUT.

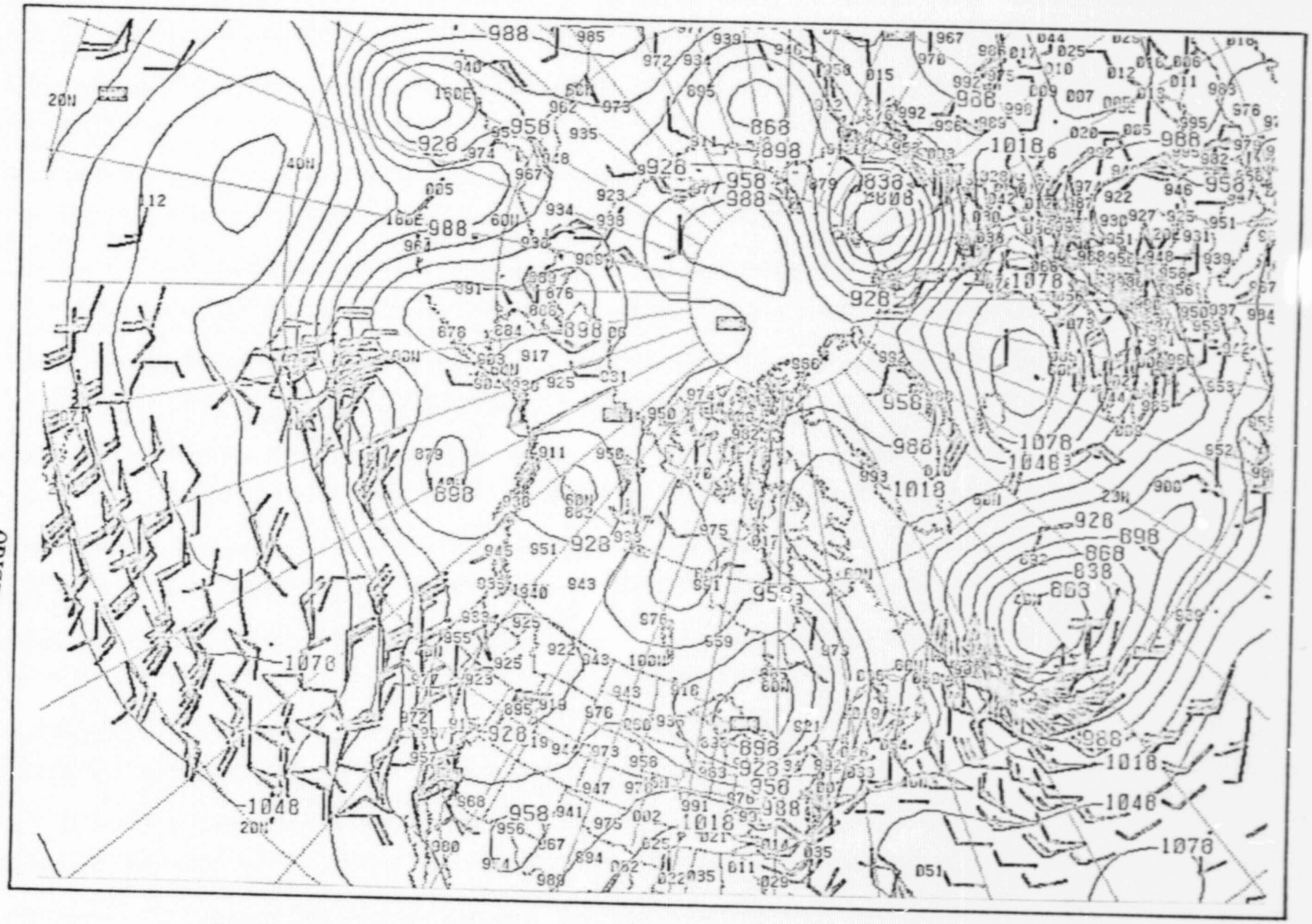


FIGURE V-2: 900MB HEIGHT ANALYSIS WITH X AND Y GRADIENT WIND OBSERVATION INPUT.

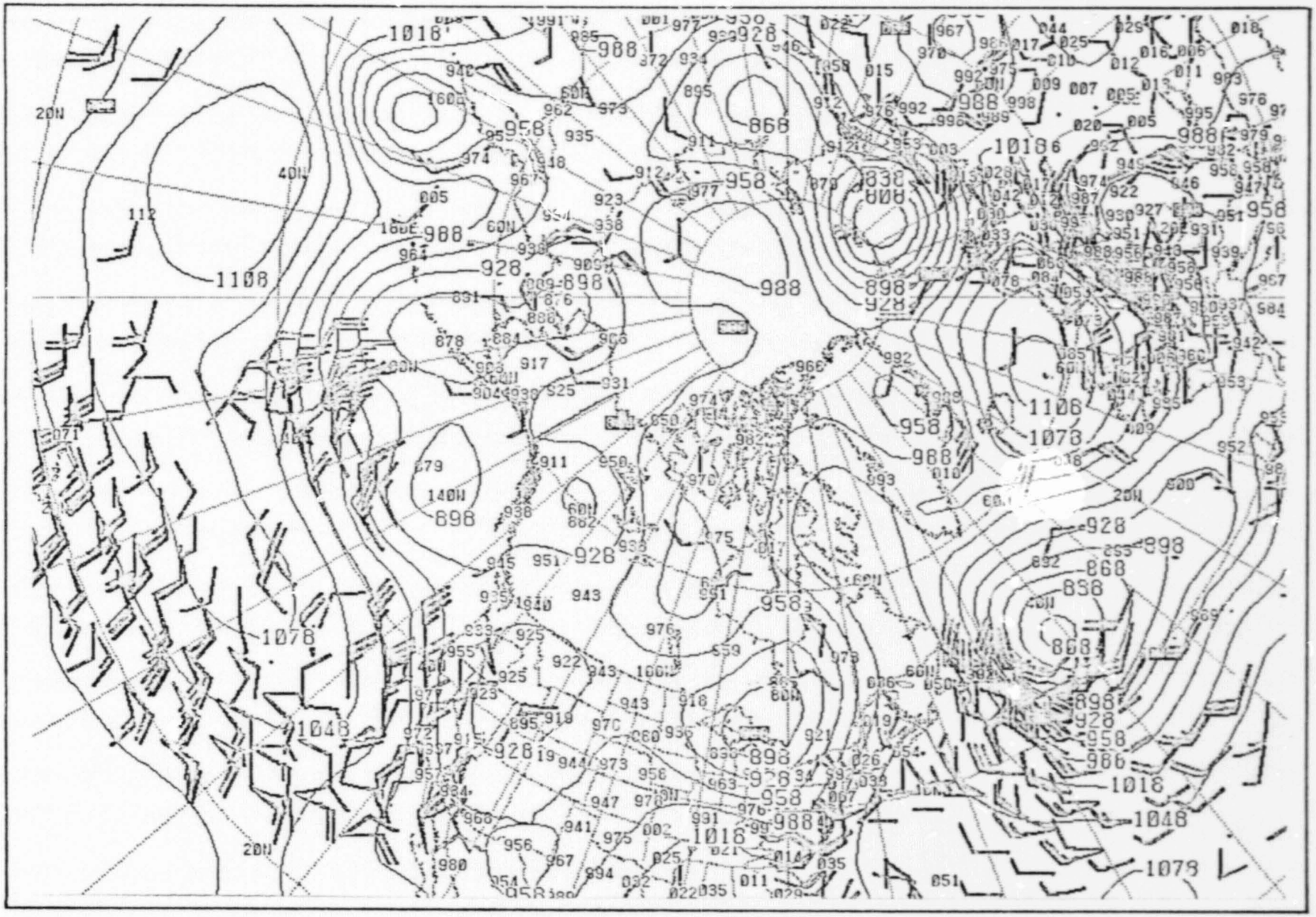


FIGURE V-3: 900MB HEIGHT ANALYSIS WITH ALL DIFFERENTIAL WIND OBSERVATION INPUT (WINDS PLOTTED).

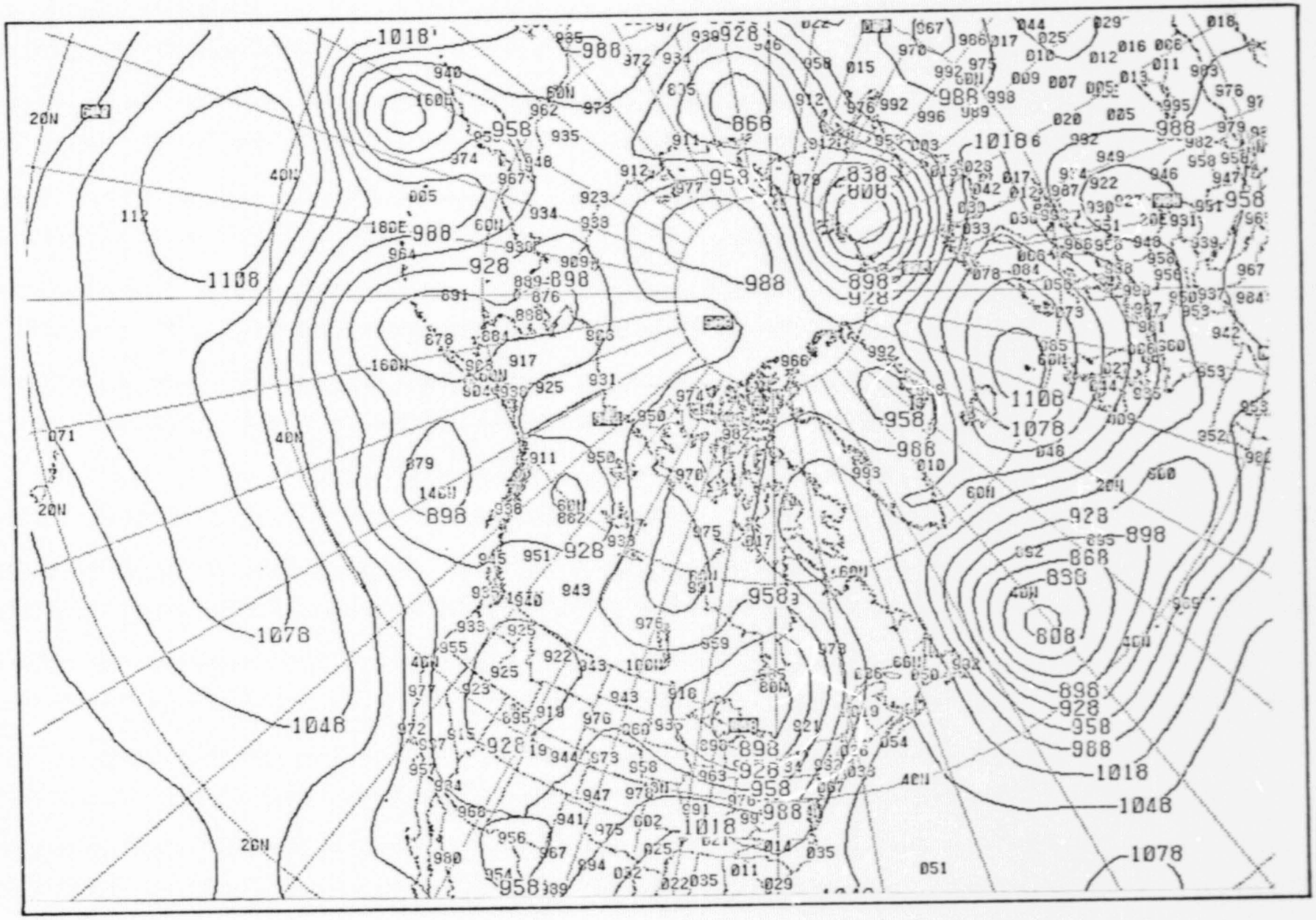


FIGURE V-4: 900MB HEIGHT ANALYSIS WITH ALL DIFFERENTIAL WIND OBSERVATION INPUT (WINDS NOT PLOTTED).

chart has the wind observations plotted; the other does not.) One should note in particular the eastward extension of the Pacific subtropical ridge. The PCT solution in which all gradients and the Laplacian were constrained to the wind observations (Figure V-4) shows a change in the orientation of the contours which better agrees with the available wind observations. The contour gradients and intensity of the ridge is considerably different from the fields with no wind observations or with contributions in only the x and y gradients. The cyclone off the east coast of the United States also exhibits a different intensity and gradients when all PCT constraints are altered by the wind observations.

Figures V-5 through V-8 show charts for similar conditions for the 250mb level. When wind derived gradients in all directions are included, the trough in the central North Atlantic shows a separate closed contour (see Figure V-7), and the trough orientation agrees much better with the available wind observations. When only x and y direction gradients are used, the wind effect is not sufficient to close off the low, as seen in Figure V-6.

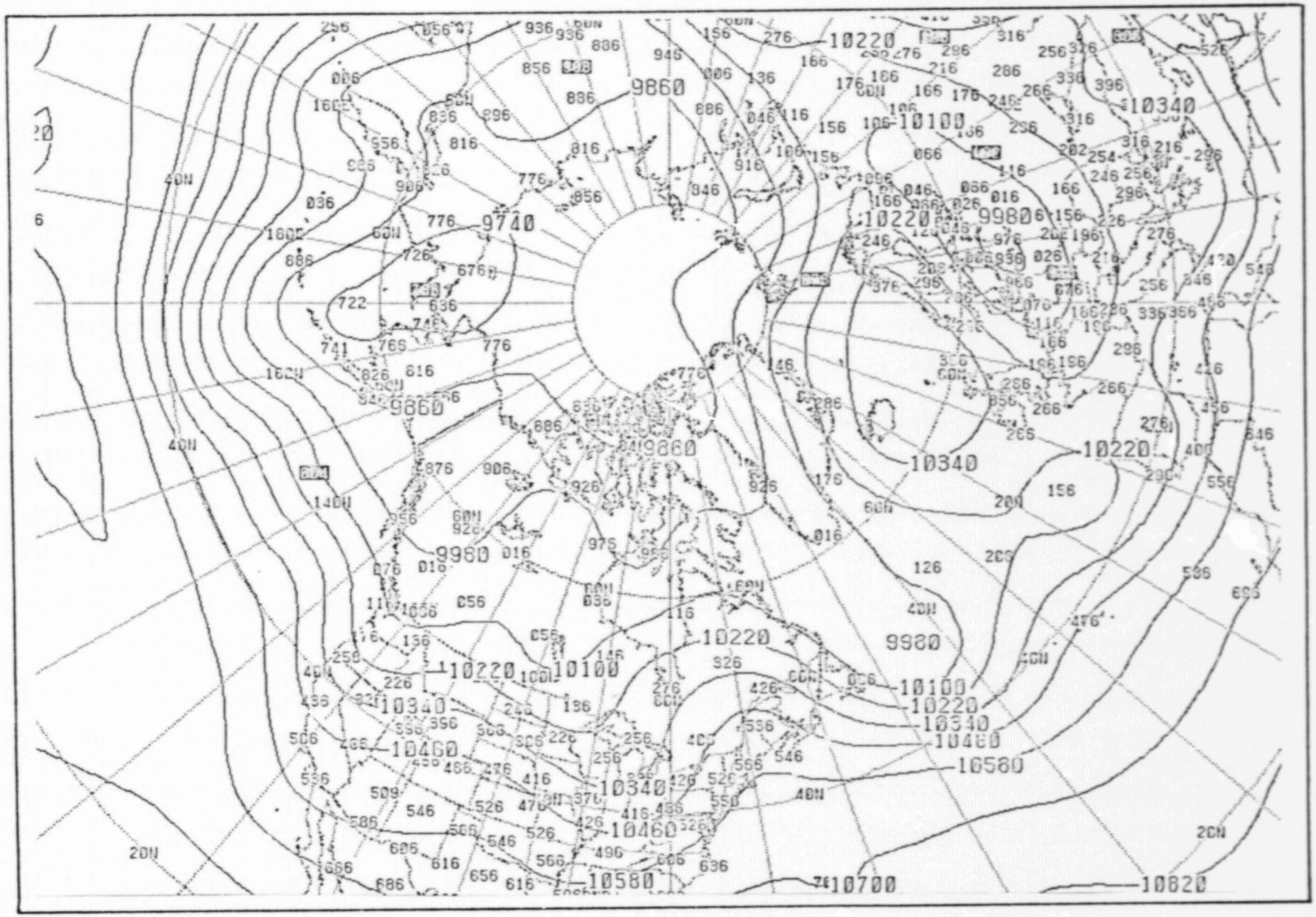


FIGURE V-5: 250MB HEIGHT ANALYSIS WITH NO WIND OBSERVATION GRADIENT INPUT.

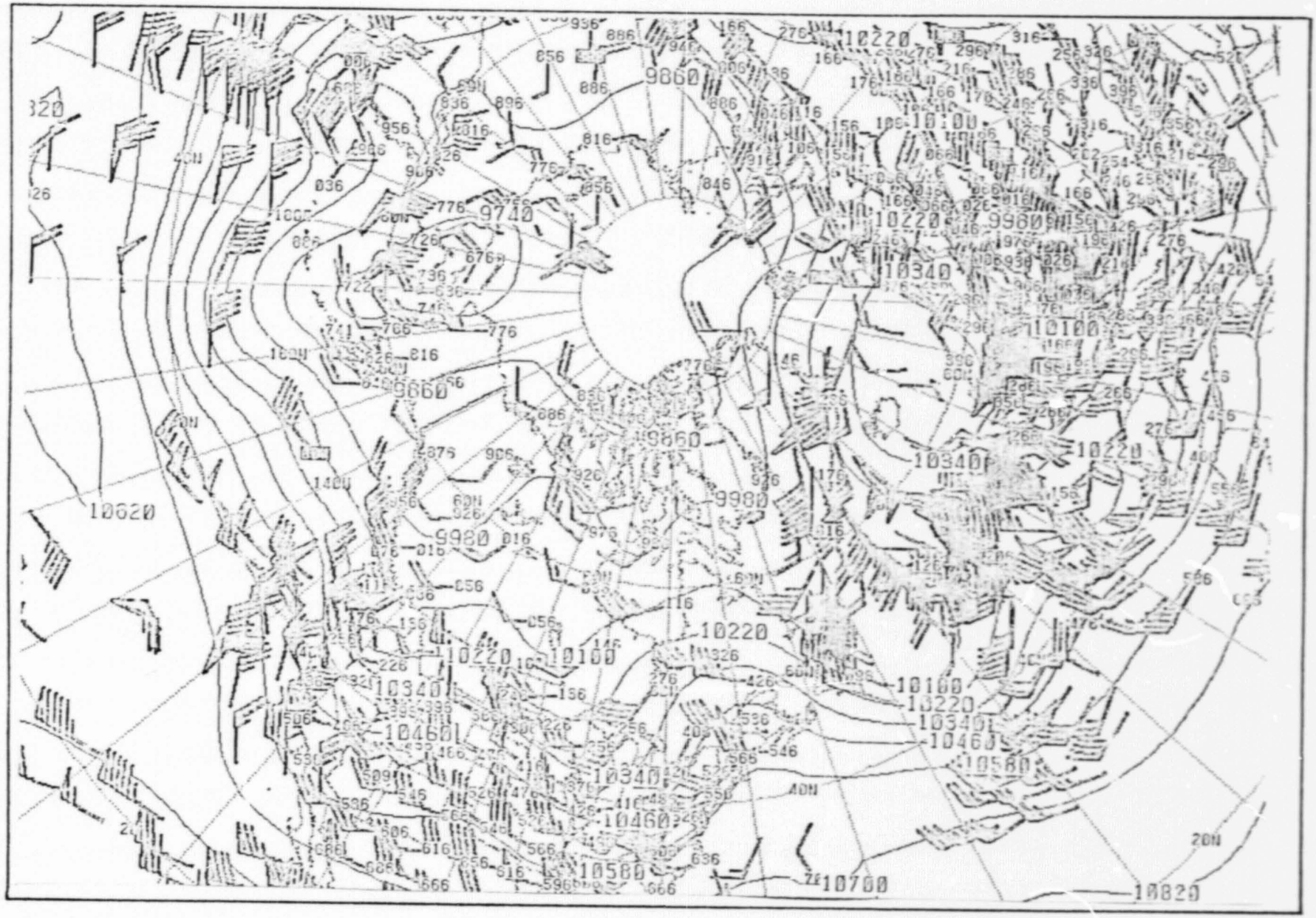


FIGURE V-6: 250MB HEIGHT ANALYSIS WITH X AND Y GRADIENT WIND OBSERVATION INPUT.

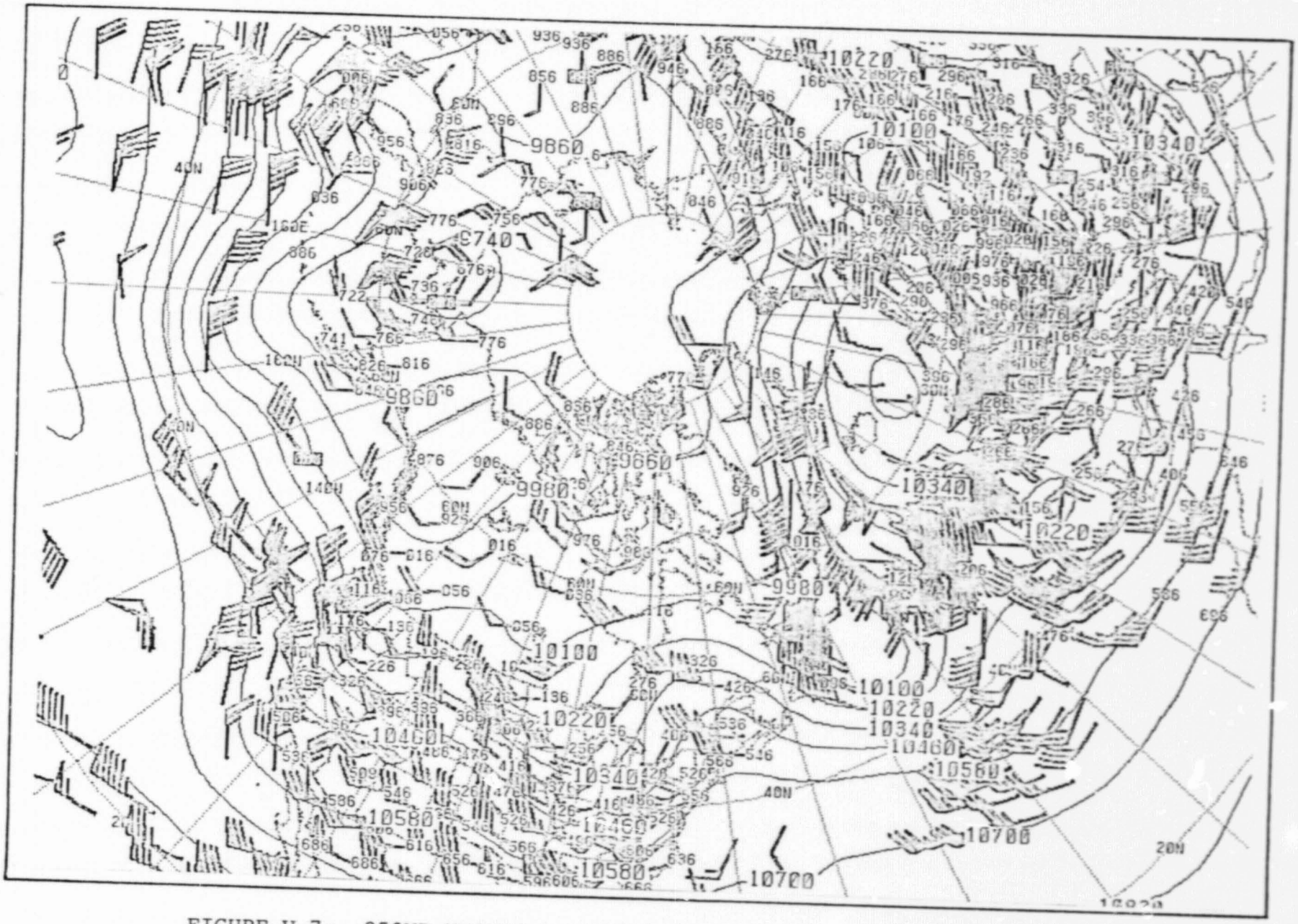


FIGURE V-7: 250MB HEIGHT ANALYSIS WITH ALL DIFFERENTIAL WIND OBSERVATION INPUT (WINDS PLOTTED).

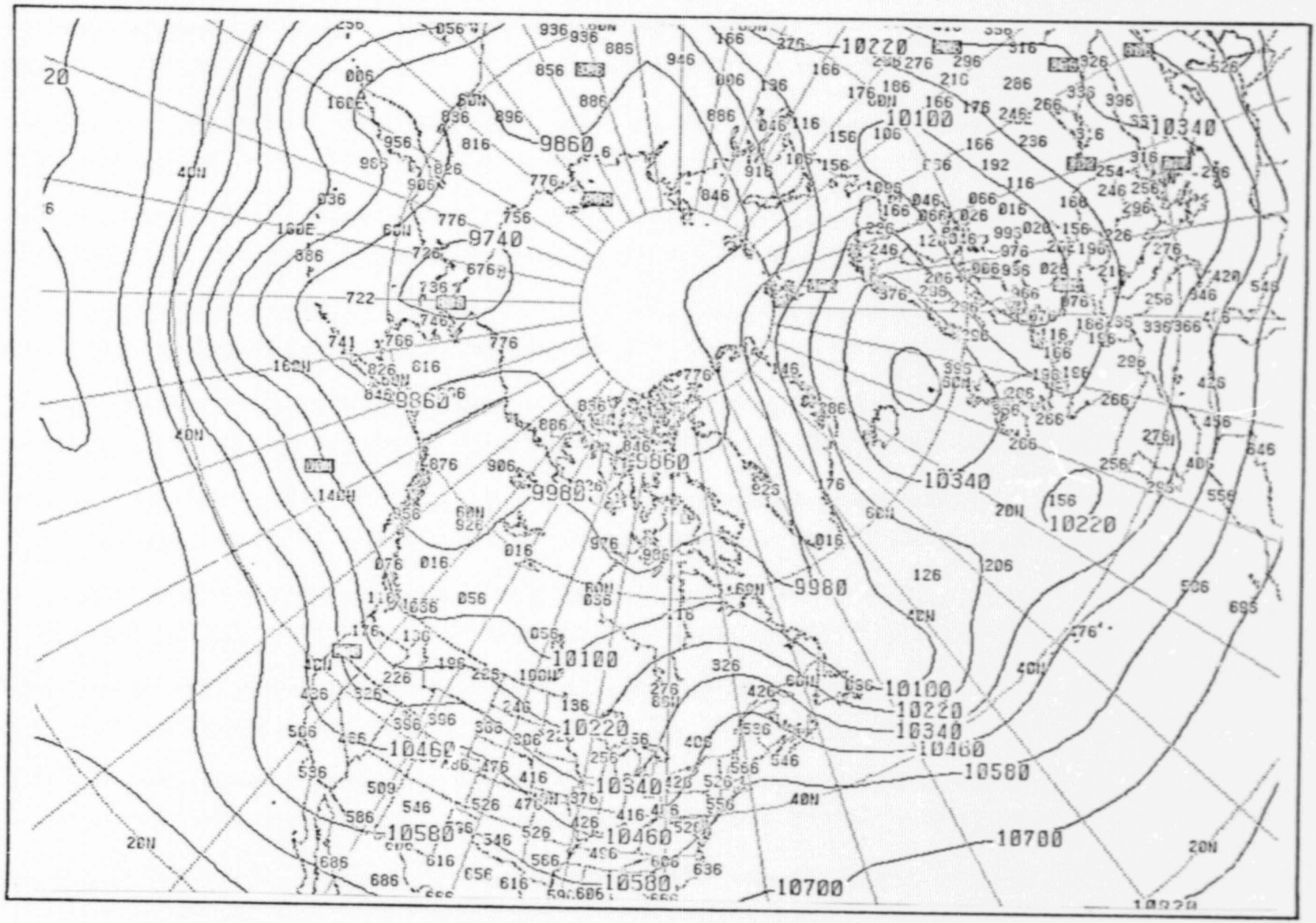


FIGURE V-8: 250MB HEIGHT ANALYSIS WITH ALL DIFFERENTIAL WIND OBSERVATION INPUT (WINDS NOT PLOTTED).

BIBLIOGRAPHY

"An Objective Analysis Based on the Variational Method", Y. Sasaki, Journal of the Meteorological Society of Japan, pp. 77-78, 1958.

"An Operational Objective Analysis System", G.P. Cressman, Monthly Weather Review, No. 10, pp. 367-374, 1959.

"Atmospheric Analysis and Prediction Model Development", ODSI Final Report for Econ, Inc. under Contract NASW-2558, August 1976 (in three volumes).

"Atmospheric Model Development in Support of SEASAT", ODSI Final Report for JPL under Contract No. 954668, June 1977 (in five volumes).

APPENDIX 1

APPENDIX I. SCALAR ANALYSIS USING THE PATTERN-CONSERVING TECHNIQUE

A. Introduction

In meteorological analysis, one piece of information that should not be ignored is the most recent past analysis, or, if available, a good forecast valid at the current analysis time. A man doing a hand analysis usually uses such information. In particular, the analyst needs an estimate of the positions of highs and lows (curvature) and of the areas of strong and weak gradients. In referring to the past analysis or forecast, the man usually is looking not so much for absolute magnitudes as for the shape of the field. When he draws his new analysis, he will first attempt to fit the shapes in the past field to the new data. If conflict occurs, the new data takes precedence unless the analyst suspects that the data is in error. In regions where data is routinely sparse, many conflicts need to be resolved because of the accumulation of errors. This results from the cycle of a poor analysis initializing a forecast which is consequently poor which is used as a deficient first guess for the next analysis/prediction cycle.

The best objective analysis scheme is probably one that follows the same rules that a person follows. If such objective techniques can be worked out, the machine may do the job in a more consistent manner than a man is able to do.

The basic goal of this analysis is to fit the following information to varying degrees: the new data; the most recent past analysis or forecast value (the first guess); the gradients of the first guess in eight directions from each grid point; and the Laplacian of the first guess. The degree of fit desired for each piece of information is specified by an array of weights. These variables and weights are named in Table I-1.

The desired fit is realized by minimizing the sum of the deviations of the various characteristics of the analysis from their counterparts in the first guess. The minimization is accomplished with an elementary application of the calculus of variations.

Information is spread through space by the gradient and Laplacian terms. In a surface analysis, there are sometimes natural obstacles (mountain ridges, coastlines, etc.) beyond which an analyst would not allow a new observation to influence the analysis. This kind of constraint can be simulated in the objective analysis by reducing the weights of the gradients and Laplacian along the demarcation zone.

The decision on the magnitudes of the various weights is less arbitrary if we view each weight as the inverse of the variance associated with the parameter it multiplies.

An analysis cycle consists of three basic steps:

1. Assemble the data at grid points.
2. Solve the minimization equation.
3. Re-evaluate the weight of each report.

In order to adequately evaluate the weight of each report, at least two cycles are required. It is desirable to include one additional cycle to allow initially suppressed data to enter the analysis with a high weight if supporting data some distance away causes the analysis to conform more closely to the report after the second cycle. The basic steps are detailed individually in the following sections.

B. Assembly

We shall refer to the guess field as $P_{i,j}$ with weight $A_{i,j}$. On the first cycle, it is the first guess, and $A_{i,j}$ has a low and probably uniform value. On subsequent cycles, $P_{i,j}$ is the result of the previous cycle, but $A_{i,j}$ keeps its original value.

The purpose of the assembly procedure is to incorporate the observational data into the first guess field $P_{i,j}$, taking into account the subjective specification of each report's reliability (DWT) and its distance from the grid point. Grid points within a circular influence region centered on each observation are affected by that observation. The size and shape of the influence function are determined by the data density and first guess field shape (i.e., gradient and Laplacian), respectively. An information density field is used to produce a factor (FACT) which varies the basic radius of the influence for each observation between a minimum and maximum limit. In areas of dense data concentration, the influence radius is set to the minimum value so as not to spread a data report's influence so far that it interferes with the already well-specified observed values. However, if the observation is isolated, its influence is spread to the maximum allowed.

The assembly radius from an observation which includes all gridpoints to be influenced by that observation is calculated as:

$$\text{RADIUS} = \frac{\text{FACT} * \text{GRADFAC} * \text{AMAP} * \text{RAD}}{\text{AMESH}}$$

where AMAP = map factor

AMESH = standard meshlength of the reference latitude

RAD = a multiple of AMESH

FACT = factor proportional to the information density

GRADFAC = gradient factor

FACT is computed as follows:

$$\text{FACT} = \text{RADMAX} - \text{INFOFAC}_{(I,J)} * (\text{RADMAX} - \text{RADMIN})$$

where RADMAX = maximum allowable factor

RADMIN = minimum allowable factor

INFOFAC(I,J) = value of information density factor nearest observation location

The maximum allowed radius (RADMAX) is decreased with each cycle in order to better define succeedingly smaller scales. The assembly radius (RADIUS) is further modified by GRADFAC, a factor which is inversely proportional to the gradient. Where gradients are large, RADIUS is reduced. Thus:

GRADFAC = A - B * (GRAD/GRADMAX)

where GRAD = maximum gradient at a gridpoint

GRADMAX = maximum gradient for the entire initial field.

A and B = constants

The basic influence function has a weight of one at its center (observation location), decreasing to zero at its maximum radius as determined by the information density. The fraction of the radius to which the weight value remains one (FRAC) varies between a minimum and maximum value determined by the curvature of the first-guess field. In systems such as cyclones, the curvature is large and an observation's full influence should not extend far from its location since it is not representative in the rapidly varying field. In anticyclones, the field varies less rapidly and it is acceptable to have the full weight of the observation included in the assembled fields at larger distances.

FRAC is computed as follows:

FRAC = 1.0 - (WTLAPL/WTLAPLM)

WTLAPL = |Laplacian(I,J)|

WTLAPLM = Percentage of the maximum Laplacian for the entire field.

WTLAPL $\leq$ WTLAPLM

FRACMIN < FRAC < FRACMAX

As the first step in the assembly procedure, the guess field is interpolated at the observation location and the difference between the observation and the guess field determined (DIF). If DIF is greater than the gross tolerance for the parameter being analyzed, it is excluded from the assembly process.

Next, the value of the influence function appropriate to the distance of the grid point from the observation is computed (w), where:

$$w = 1.0 \text{ if } \frac{\text{distance}}{\text{RADIUS}} < \text{FRAC}$$

$$\text{otherwise } w = \frac{1.0 - \frac{\text{distance}}{\text{RADIUS}}}{1.0 - \text{FRAC}}$$

For each gridpoint affected by the observation, a cumulative sum of the product (W*DWT)*W*DIF is computed at the appropriate I,J. Also, a field of the product W*DWT is accumulated. Once all observations have been processed, the assembled value is obtained by dividing the two fields at all grid points:

$$P_{I,J} = P_{I,J} + \frac{\sum_{K=1}^{NOBS} [W*DWT(K)]*W*DIF}{\sum_{K=1}^{NOBS} W*DWT(K)} \qquad \text{for } \begin{matrix} I=1,M \\ J=1,N \end{matrix}$$

P = assembled field value
NOBS = number of observations
DWT = data weight assigned to an observation

C. Minimizing the Deviations

TABLE I-1: PCT SCALAR CONSTRAINTS

Constraint	Weight
$P_{i,j}$ = Variable being analyzed (assembled value)	$A_{i,j}$
$\mu_{i,j}$ = y axis gradient = $P_{i,j+1} - P_{i,j}$ (computed from non-assembled value of first guess)	$B_{i,j}$
$\nu_{i,j}$ = x axis gradient = $P_{i+1,j} - P_{i,j}$ (computed from non-assembled value of first guess)	$C_{i,j}$
$\alpha_{i,j}$ = x-1,y+1 gradient = $P_{i-i,j+1} - P_{i,j}$ (computed from non-assembled value of first guess)	$E_{i,j}$
$\beta_{i,j}$ = x+1,y+1 gradient = $P_{i+1,j+1} - P_{i,j}$ (computed from non-assembled value of first guess)	$F_{i,j}$
$L_{i,j}$ = Laplacian = $P_{i+1,j} + P_{i-1,j} + P_{i,j+1} + P_{i,j-1} - 4P_{i,j}$ (computed from non-assembled value of first guess)	$D_{i,j}$

The first guess shapes μ, ν, α, β and L and their respective weights B, C, E, F and D have a constant value during the entire analysis. Within limits specified by the weights, we require the final analysis to have similar values of the constraints as the first guess field.

To effect this matching, we shall minimize the following integral:

$$I \equiv \iint [\qquad [I.1]$$

(a) $A_{i,j} (P^*_{i,j} - P_{i,j})^2 +$

(b) $B_{i,j} (P^*_{i,j+1} - P^*_{i,j} - \mu_{i,j})^2 +$

(c) $B_{i,j-1} (P^*_{i,j} - P^*_{i,j-1} - \mu_{i,j-1})^2 +$

(d) $C_{i,j} (P^*_{i+1,j} - P^*_{i,j} - \nu_{i,j})^2 +$

(e) $C_{i-1,j} (P^*_{i,j} - P^*_{i-1,j} - \nu_{i-1,j})^2 +$

(f) $E_{i,j} (P^*_{i-1,j+1} - P^*_{i,j} - \alpha_{i,j})^2 +$

(g) $E_{i+1,j-1} (P^*_{i,j} - P^*_{i+1,j-1} - \alpha_{i+1,j-1})^2 +$

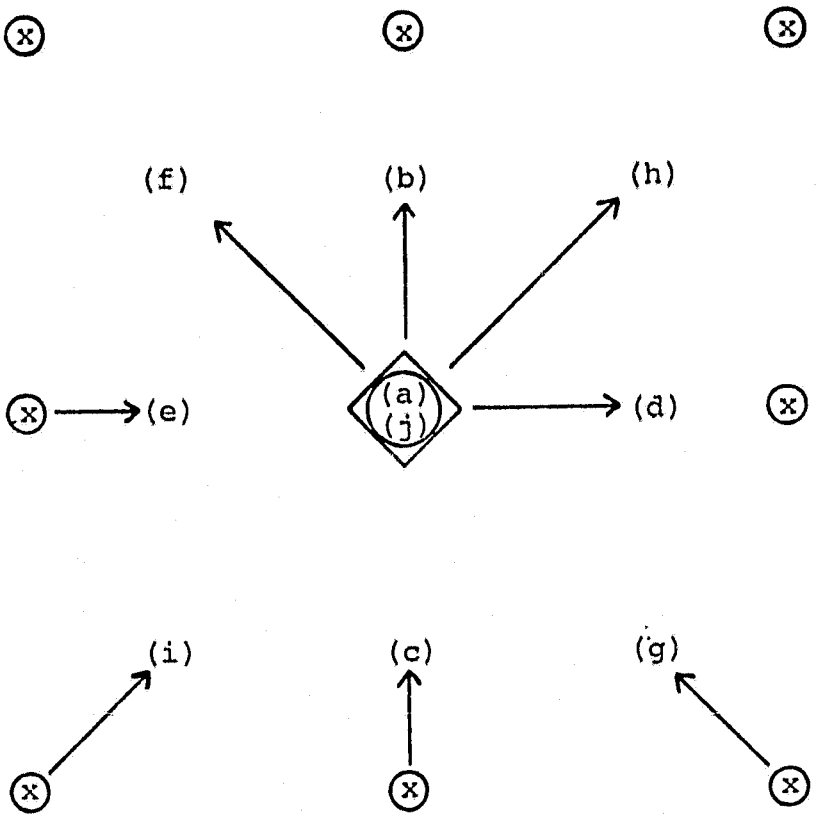
(h) $F_{i,j} (P^*_{i+1,j+1} - P^*_{i,j} - \beta_{i,j})^2 +$

(i) $F_{i-1,j-1} (P^*_{i,j} - P^*_{i-1,j-1} - \beta_{i-1,j-1})^2 +$

(j) $D_{i,j} (P^*_{i+1,j} + P^*_{i-1,j} + P^*_{i,j+1} + P^*_{i,j-1} - 4P^*_{i,j} - L_{i,j})^2 +$

$]dxdy$

In the above, the starred quantities are the analysis values we are seeking. Each term is a departure from the desired matching of differential properties. Extra terms have been added to account for the effect of a changing $P^*_{i,j}$ on the differential properties computed at surrounding points. Their effect is to more closely couple neighboring grid points. See Figure I-1 for a depiction of the minimization stencil as it relates to the terms of equation [I.1]. To minimize the integral, we simply take the first variation with respect to $P^*_{i,j}$, and set it to zero (see equation [I.2]). The solution of the resulting equation will be the $P^*_{i,j}$ that will cause the integral to be minimized. The fact that each term is squared ensures a minimum as opposed to a maximum value.



LEGEND: ( ) = constraint from Equation [I.1]

= difference

= gradient

= laplacian at grid point

= grid points

FIGURE I-1: SCALAR MINIMIZATION STENCIL

$$\frac{\delta I}{\delta P^*} = \iint [\ 2A_{i,j}\ (P^*_{i,j} - P_{i,j}) \qquad \text{[I.2]}$$

$$-2\ B_{i,j}\ (P^*_{i,j+1} - P^*_{i,j} - \mu_{i,j})$$

$$+\ 2\ B_{i,j-1}\ (P^*_{i,j} - P^*_{i,j-1} - \mu_{i,j-1})$$

$$-\ 2\ C_{i,j}\ (P^*_{i+1,j} - P^*_{i,j} - \nu_{i,j})$$

$$+\ 2\ C_{i-1,j}\ (P^*_{i,j} - P^*_{i-1,j} - \nu_{i-1,j})$$

$$-\ 2\ E_{i,j}\ (P^*_{i-1,j+1} - P^*_{i,j} - \alpha_{i,j})$$

$$+\ 2\ E_{i+1,j-1}\ (P^*_{i,j} - P^*_{i+1,j-1} - \alpha_{i+1,j-1})$$

$$-\ 2\ F_{i,j}\ (P^*_{i+1,j+1} - P^*_{i,j} - \beta_{i,j})$$

$$+\ 2\ F_{i-1,j-1}\ (P^*_{i,j} - P^*_{i-1,j-1} - \beta_{i-1,j-1})$$

$$-\ 8\ D_{i,j}\ (P^*_{i+1,j} + P^*_{i-1,j} + P^*_{i,j+1} + P^*_{i,j-1} - 4P^*_{i,j} - L_{i,j})$$

$$]\ dxdy \overset{\text{set}}{=} 0$$

The terms in $\frac{\delta I}{\delta P^*}$ can be grouped into three categories:

1. Those involving $P^*_{i,j}$.
2. Those involving P* at surrounding points.
3. Those not involving P*.

$$S_{i,j} P^*_{i,j} \left\{ \begin{array}{l} [\ A_{i,j} + B_{i,j} + B_{i,j-1} + C_{i,j} + C_{i-1,j} + E_{i,j} \\ \quad + E_{i+1,j-1} + F_{i,j} + F_{i-1,j-1} + 16\ D_{i,j}\ ]\ P^*_{i,j} \end{array} \right.$$

$$-H_{i,j} \begin{cases} - B_{i,j}\, P^*_{i,j+1} - B_{i,j-1}\, P^*_{i,j-1} - C_{i,j}\, P^*_{i+1,j} - C_{i-1,j}\, P^*_{i-1,j} \\ - E_{i,j}\, P^*_{i-1,j+1} - E_{i+1,j-1}\, P^*_{i+1,j-1} - F_{i,j}\, P^*_{i+1,j+1} \\ - F_{i-1,j-1}\, P^*_{i-1,j-1} - 4\, D_{i,j}\, P^*_{i+1,j} - 4\, D_{i,j}\, P^*_{i-1,j} \\ - 4\, D_{i,j}\, P^*_{i,j+1} - 4\, D_{i,j}\, P^*_{i,j-1} \end{cases}$$

$$-G_{i,j} \begin{cases} - A_{i,j}\, P_{i,j} + B_{i,j}\, \mu_{i,j} - B_{i,j-1}\, \mu_{i,j-1} + C_{i,j}\, \nu_{i,j} \\ - C_{i-1,j}\, \nu_{i-1,j} + E_{i,j}\, \alpha_{i,j} - E_{i+1,j-1}\, \alpha_{i+1,j-1} \\ + F_{i,j}\, \beta_{i,j} - F_{i-1,j-1}\, \beta_{i-1,j-1} + 4\, D_{i,j}\, L_{i,j} \end{cases}$$

Note that all terms in S and G except $A_{i,j}$ in $S_{i,j}$ and $-A_{i,j}\, P_{i,j}$ in $G_{i,j}$ involve first-guess pattern information which is consistent during the analysis.

The minimization may be written as

$$S_{i,j}\, P^*_{i,j} - (G_{i,j} + H_{i,j}) = 0 \qquad [I.3]$$

In $H_{i,j}$, let us group together the coefficients of P* at each point.

$$
\begin{aligned}
-H_{i,j} = &+ (-C_{i-1,j} - 4\,D_{i,j})\,P^{*}_{i-1,j} \\
&+ (-C_{i,j} - 4\,D_{i,j})\,P^{*}_{i+1,j} \\
&+ (-E_{i,j})\,P^{*}_{i-1,j+1} \\
&+ (-B_{i,j} - 4\,D_{i,j})\,P^{*}_{i,j+1} \\
&+ (-F_{i,j})\,P^{*}_{i+1,j+1} \\
&+ (-F_{i-1,j-1})\,P^{*}_{i-1,j-1} \\
&+ (-B_{i,j-1} - 4\,D_{i,j})\,P^{*}_{i,j-1} \\
&+ (-E_{i+1,j-1})\,P^{*}_{i+1,j-1}
\end{aligned}
$$

Define:

$$
\begin{aligned}
X_{i,j} &\equiv C_{i,j} \\
Y_{i,j} &\equiv B_{i,j} \\
Z_{i,j} &\equiv -F_{i,j} \\
R_{i,j} &\equiv -E_{i+1,j}
\end{aligned}
$$

Note that X, Y, Z and R have a constant value during the analysis.

Then:

$$
\begin{aligned}
-H_{i,j} = &\; X_{i-1,j}\,P^{*}_{i-1,j} - X_{i,j}\,P^{*}_{i+1,j} \\
&+ R_{i-1,j}\,P^{*}_{i-1,j+1} - Y_{i,j}\,P^{*}_{i,j+1} \\
&+ Z_{i,j}\,P^{*}_{i+1,j+1} + Z_{i-1,j-1}\,P^{*}_{i-1,j-1} \\
&- Y_{i,j-1}\,P^{*}_{i,j-1} + R_{i,j-1}\,P^{*}_{i+1,j-1} \\
&- 4\,D_{i,j}(P^{*}_{i-1,j}+P^{*}_{i+1,j}+P^{*}_{i,j+1}+P^{*}_{i,j-1})
\end{aligned}
\qquad [I.4]
$$

The minimization equation [I.3] is solved by simultaneous over-relaxation. The matrices $S_{i,j}$ and $G_{i,j}$ may be computed initially except for the $A_{i,j}$ term and will not change throughout the analysis. Matrix $H_{i,j}$ must be recomputed for every iteration of the relaxation.

The relaxation proceeds as follows: At Point (i,j) the terms of the minimization equation are evaluated using the assembled P field for P*. In general, the equation is not satisfied and a residual is defined as

$$S_{i,j}\ P^{*\tau}_{i,j} - (G_{i,j} + H_{i,j}) \equiv R \qquad [I.5]$$

The superscript τ is an iteration counter. The value of $P^{*}_{i,j}$ is to be altered so that on the next iteration, the residual will be zero, provided $H_{i,j}$ does not change. Of course, $H_{i,j}$ will change, but if the equation is fairly well behaved, repetition of the procedure should lead to convergence on the correct solution.

$$S_{i,j}\ P^{*\tau+1}_{i,j} - (G_{i,j} + H_{i,j}) = 0 \qquad [I.6]$$

Subtracting [I.6] from [I.5],

$$S_{i,j}\ (P^{*\tau}_{i,j} - P^{*\tau+1}_{i,j}) = R \qquad [I.7]$$

and

$$P^{*\,T+1}_{i,j} = P^{*\,T}_{i,j} - \frac{R}{S_{i,j}}$$

Convergence can be hastened by increasing the correction term in [I.7] by a factor ALFA. The factor by which it is increased is called the over-relaxation coefficient. Equation [I.7] becomes:

$$P^{*\,T+1}_{i,j} = P^{*\,T}_{i,j} - \text{ALFA}\,\frac{R}{S_{i,j}} \qquad [I.8]$$

One iteration consists of making the correction [I.8] at every grid point. Testing has shown the convergence can be speeded up and unwanted solution noise decreased if the grid points are processed in a circular manner. Therefore, the field is scanned in a counter-clockwise circular sweep starting at the center and working toward the boundaries. Iterations are repeated until the maximum residual is less than a specified convergence criterion. The resulting P* field is the solution of equation [I.3].

D. Reevaluating the Data Weights

At the end of each cycle, the weight of each report is reevaluated. An observation will have its weight reduced if the report differs from the analysis value on the current scan by more than a subjectively determined amount. REVAL is the reevaluation parameter and CRIT, the critical value at which a report's weight is reduced. REVAL is calculated as:

$$\frac{FP * REFAC * (DIF)^2}{ODWT}$$

where FP = scan number

REFAC = constant, reevaluation

ODWT = original data weight factor

If REVAL is less than CRIT, the observation retains its original weight, even though it may have been reduced another scan. If REVAL is greater than CRIT, then:

$$DWT = \frac{CONST * ODWT}{1 + REVAL}$$

where CONST = constant

ODWT = original data weight

DWT = new data weight

Notice that on any cycle, every data point may have its original weight restored, even if it had been reduced previously. In this way, a report that causes a large change in the analysis may have full effect if it is supported by data nearby.

E. Program Description - 63x63 Grid Version

1. INFODEN

The logic in the assembling process requires a field which quantifies the density of the available observations. The subroutine INFODEN produces such a field and related statistics. The density is computed in a manner similar to that used in the assembly process without the variation introduced by FRAC and FACT. An influence function with a value of one at the center, reducing to zero at its periphery is superimposed at each observation location. All grid points within the circle accumulate a contribution equal to the appropriate influence value multiplied by the observations subjective weight. The resulting field is written to the random access file TAPE9, with the record name IDEN as described in the PCT description. A 100-word array is computed giving the distribution of the information density using a $\log_{10}$ increment between the minimum and maximum values in the field. Finally, the density value corresponding to having PPER percent of the grid points with a density less than PPER is computed and defined as DENLIM for use in the assembly process.

2. PCT

The calling arguments are described in detail in the comments of the program. All the arrays are variably dimensioned, using the dimensions M and N provided in the calling arguments. The date-time group is provided through common block /DTG/, a title for plotting, a contour starting point and a contour interval are in common block /INFO/ and sense switch variables are passed in /ISW/.

A random-access file TAPE9 is used for temporary storage and must be declared on the PROGRAM card of the calling program. The writing of some arrays on the random-access file for later retrieval allows their use as work arrays. First, the data list, the I and J data location lists, the initial data weight list and the initial weight field of the first-guess A and laplacian field D are all written on the random-access file so that these arrays can be used in subroutine BKGRND. Since the same array is frequently used to hold two different fields, two names, separated by a space, make up the names of these arrays. Of course, the space is ignored by Fortran. The two names are interpreted as one, but this convention helps in reading the listing.

After BKGRND computes matrices S and G (see page I-12 and I-13), they, along with Y and X, are written on TAPE9 and the data-related arrays are read back in.

DO loop 100 is the main loop controlling the number of cycles to be made through the program. A cycle consists of assembly (Section I-B), solving equation [I.3] (Section I-C), and reevaluating the data weights (Section I-D). First, the subroutine ASSMBL is called to include the influence of the available observations. The latest solution of P, the field being analyzed, is available for use as the first-guess field for the assembly.

After calling ASSMBL on the last analysis cycle only, subroutine PLTDAT (see Appendix) writes the data list on the plot file. The data list is rewritten on TAPE9 because in ASSMBL, gross errors were flagged by setting the last bit of the data word. The last bit of all good data is cleared. Also, the data weight for the current cycle is written to TAPE9 as CURWT for later use by the subroutine REVALWT. Matrices A, G and S (see page I-12 and I-13) are read from TAPE9 and subroutine UPDATE adds A to S and A*P to G. Weight field D and arrays X and Y are read from TAPE9, and subroutine BLEND solves equation [I.3], resulting in a new analysis field.

After restoring the arrays DATA, AIS and AJG, and the array A with the current data weight CURWT, REVALWT is called. The subroutine reevaluates the data weight and computes RMS values. If the number of cycles completed is less than NOPAS, the program continues through another entire cycle after reducing the scan radius and gross reject limit. If the analysis is complete, the analyzed field is passed through a variable number of passes of a filter. If appropriate, the tropical latitudes are smoothed by calling the subroutine SMTHP (see Appendix). Depending upon external sense switch settings, the analysis field is written on the plot file by subroutine PLOT (see Appendix); PRT (see Appendix) makes a printer map of the field, and the field may be written to disk using the Fleet Numerical Weather Central (FNWC) random access routine ZRANDIO (see Appendix).

2. BKGRND

Matrices S, G, X, Z and R (see pages I-12, I-13 and I-14) are computed and returned. There is no problem in interpreting the code with the aid of the comments. A maximum value for the laplacian and absolute gradients within the first-guess field are computed and stored in common.

3. ASSMBL

First, the arrays SUM and DENOM are set to zero, and DENSITY initialized with the information density field. Then, the data list is scanned and the guess field interpolated to each report location. If the interpolated value differs from the report by more than GROS, the report is rejected by setting the last bit of the report word. Otherwise, this bit is cleared. The value of GROS varies with latitude and decreases with each cycle. If the report is flagged as being a bogus report, it is not tested for possible rejection since bogus observations are entered to correct areas of incorrect analysis. The constants FRAC and FACT are computed as described in Section I-B using the current first-guess field. A bogus observation has FRAC set to FRACMAX, and FACT set to the smaller of RADMAX and $1\frac{1}{2}$ the computed FACT. All grid points within the subgrid computed around the observation location have accumulated in sum the product W*DWT*W*DIF (see I-B for notation) and in DENOM the product W*DWT. After all data have been scanned, SUM is divided by DENOM to obtain the contribution to the first-guess field producing the assembled field (see Section I-B).

Finally, statistics describing accepted and rejected distributions are computed and displayed.

4. UPDATE

The current assembled weight field A is added to matrix S (see page I-12 and the BKGRND listing) and A*P is added to matrix G, where P is the current guess field.

5. BLEND

The minimization equation [I.3] is solved by simultaneous over-relaxation. The method is described in detail on pages I-15 and I-16. No further discussion is warranted.

6. REVALWT

The weight of each datum is reevaluated as discussed in Section I-D. First, a distribution of weights is computed and printed before reevaluation. No reevaluation of bogus observations are allowed. Once the difference between the observation and the resulting field is determined, λ^2 is computed. If λ^2 is greater than one, the observation weight is decreased according to the equation in Section I-D. Once all observations have been processed, the distribution of weights is recomputed and displayed as before.

APPENDIX 2

APPENDIX II: VECTOR WIND ANALYSIS USING THE PATTERN CONSERVING TECHNIQUE

A. Introduction

The pattern-conserving technique described in Section I is used to analyze a scalar variable. In this section, we will concentrate on those aspects which are peculiar to the wind problem.

The most essential feature of the pattern-conserving technique is that, while fitting new data, it tends to retain certain differential properties of the first-guess field. For scalar analysis, we were only concerned with gradients and the Laplacian. The wind, being a vector, complicates the problem slightly. Some of the properties we would like to conserve; e.g., vorticity and divergence, involve both scalar components. We must analyze both components simultaneously.

The differential properties that we choose to conserve are the gradients of each wind component in eight directions from each grid point, the vorticity and the divergence. The same method is used here as in the scalar analysis, the main difference being that two minimization equations rather than one must be solved simultaneously.

The equations are considerably simplified by using the staggered grid illustrated by Figure II-1 and defining the divergence, vorticity and gradients as in Table II-1 and Figure II-2. This arrangement causes certain matrices to be tridiagonal.

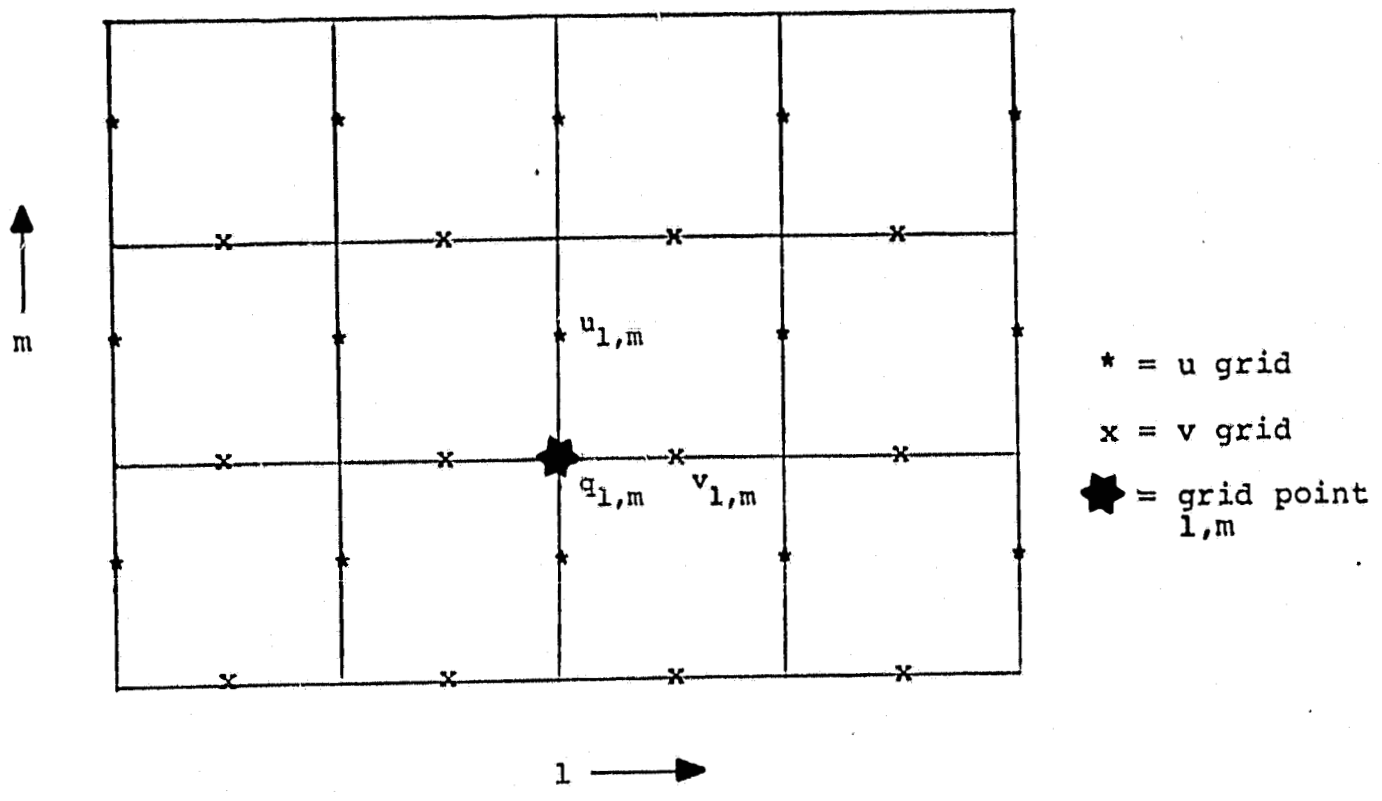


FIGURE II-1: STAGGERED u,v GRID

B. Assembly

The assembly of data to grid points is done in the same way as in the scalar analysis. The wind components are moved to grid points within an influence function, and a weighted average is computed at each grid point for each component. Since the grid is staggered (Figure II-1), the components of an observation may be moved to grid points with different weightings.

For the wind analysis, we need a field SUMU and SUMV initialized to zero for accumulating the contribution to the assembled field of each component. The denominator is carried as two packed floating point values in the array DENOM. The influence function radius is varied as in the scalar analysis based upon the information density to compute the factor FACT. Unlike the scalar influence function, the radius of weight equals one is not variable and decreases from one at the center to zero at the maximum radius.

C. Minimizing the Deviations

In the scalar analysis, we wanted to conserve the gradients and the Laplacian. With the regular grid, the finite difference expressions for the gradients and Laplacian did not provide good horizontal coupling among the grid points, so we included in the integral to be minimized the gradients and Laplacian at surrounding grid points. In the case of the wind analysis, the more complex differential properties and the staggered grid extend the influence of the computations at a grid point further than in the scalar analysis, and it is not necessary to add the contributions at the surrounding points.

TABLE II-1: PCT VECTOR CONSTRAINTS

Constraint	Weight
$u_{l,m}$ = Variable being analyzed (assembled value)	$A_{l,m}$
$v_{l,m}$ = Variable being analyzed (assembled value)	$\hat{A}_{l,m}$
$d_{l,m}$ = divergence = $\partial u/\partial x + \partial v/\partial y$	
$= u_{l+1,m} - u_{l,m} + v_{l,m+1} - v_{l,m}$	$D_{l,m}$
(Computed from non-assembled value of first guess.)	
$q_{l,m}$ = vorticity = $\partial v/\partial x - \partial u/\partial y$	$Q_{l,m}$
$= v_{l,m} - v_{l-1,m} - u_{l,m} + u_{l,m-1}$	
(Computed from non-assembled value of first guess.)	
$e_{l,m}$ = x-1,y+1 u gradient = $u_{l-1,m+1} - u_{l,m}$	$E_{l,m}$
(Computed from non-assembled value of first guess.)	
$\hat{e}_{l,m}$ = x-1 y+1 v gradient = $v_{l-1,m+1} - v_{l,m}$	$\hat{E}_{l,m}$
(Computed from non-assembled value of first guess.)	
$f_{l,m}$ = y axis u gradient = $u_{l,m+1} - u_{l,m}$	$F_{l,m}$
(Computed from non-assembled value of first guess.)	
$\hat{f}_{l,m}$ = y axis v gradient = $v_{l,m+1} - v_{l,m}$	$\hat{F}_{l,m}$
(Computed from non-assembled value of first guess.)	
$g_{l,m}$ = x+1,y+1 u gradient = $u_{l+1,m+1} - u_{l,m}$	$G_{l,m}$
(Computed from non-assembled value of first guess.)	
$\hat{g}_{l,m}$ = x+1,y+1 v gradient = $v_{l+1,m+1} - v_{l,m}$	$\hat{G}_{l,m}$
(Computed from non-assembled value of first guess.)	
$h_{l,m}$ = x axis u gradient = $u_{l+1,m} - u_{l,m}$	$H_{l,m}$
(Computed from non-assembled value of first guess.)	
$\hat{h}_{l,m}$ = x axis v gradient = $v_{l+1,m} - u_{l,m}$	$\hat{H}_{l,m}$
(Computed from non-assembled value of first guess.)	

We shall minimize the following integral:

$$I \equiv \iint [$$

(a) $$A_{l,m}(u^*_{l,m} - u_{l,m})^2 \qquad \text{[II.1]}$$

(b) $$+ \hat{A}_{l,m}\,(v^*_{l,m} - v_{l,m})^2$$

(c) $$+ D_{l,m}(u^*_{l+1,m} - u^*_{l,m} + v^*_{l,m+1} - v^*_{l,m} - d_{l,m})^2$$

(d) $$+ Q_{l,m}(v^*_{l,m} - v^*_{l-1,m} - u^*_{l,m} + u^*_{l,m-1} - q_{l,m})^2$$

(e) $$+ E_{l,m}(u^*_{l-1,m+1} - u^*_{l,m} - e_{l,m})^2$$

(f) $$+ \hat{E}_{l,m}(v^*_{l-1,m+1} - v^*_{l,m} - \hat{e}_{l,m})^2$$

(g) $$+ F_{l,m}(u^*_{l,m+1} - u^*_{l,m} - f_{l,m})^2$$

(h) $$+ \hat{F}_{l,m}(v^*_{l,m+1} - v^*_{l,m} - \hat{f}_{l,m})^2$$

(i) $$+ G_{l,m}(u^*_{l+1,m+1} - u^*_{l,m} - g_{l,m})^2$$

(j) $$+ \hat{G}_{l,m}(v^*_{l+1,m+1} - v^*_{l,m} - \hat{g}_{l,m})^2$$

(k) $$+ H_{l,m}(u^*_{l+1,m} - u^*_{l,m} - h_{l,m})^2$$

(l) $$+ \hat{H}_{l,m}(v^*_{l+1,m} - v^*_{l,m} - \hat{h}_{l,m})^2$$

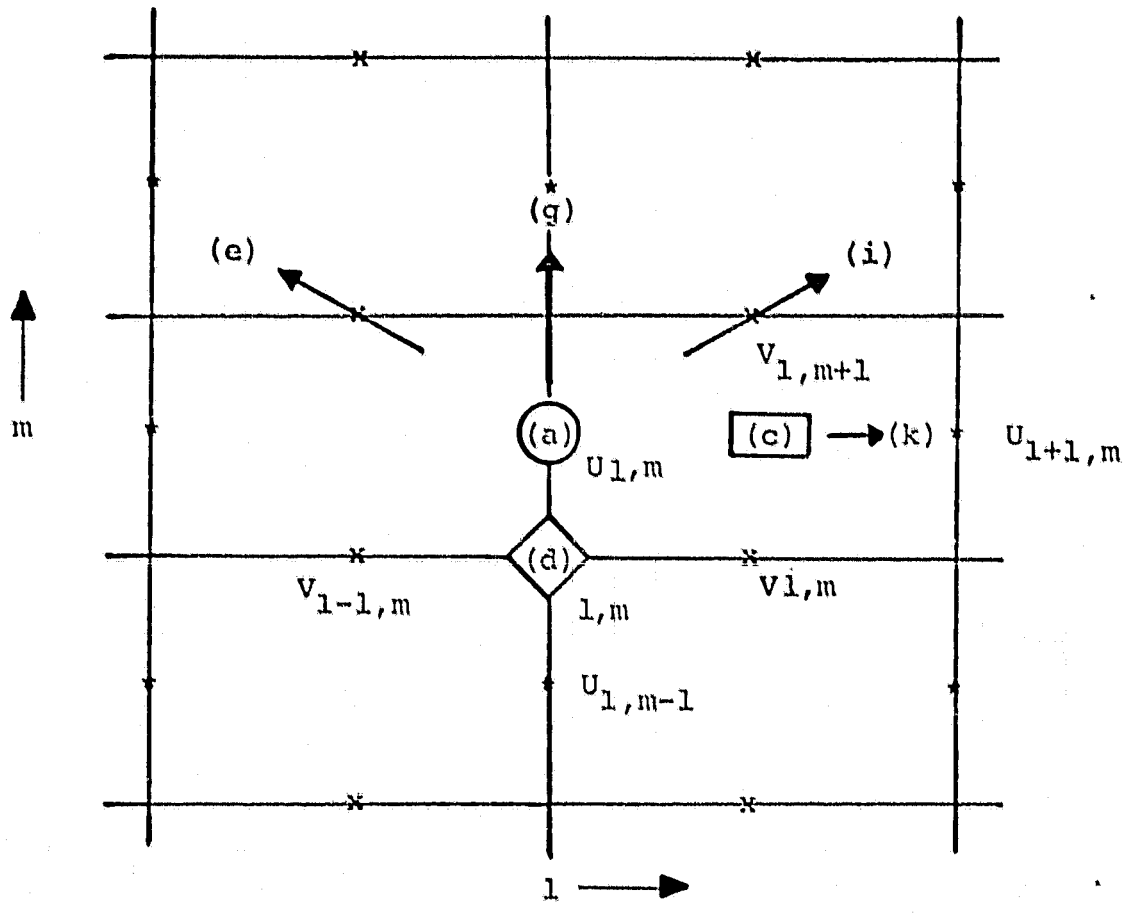
$$] \, dxdy$$

The superscript (*) indicates the values we seek. The differential properties of the first-guess field are defined in Table II-1, and a depiction of the u component minimization stencil as it relates to the u terms of equation [II.1] is given in Figure II-2.

To minimize the integral we take the first variation with respect to $u^*_{l,m}$ and with respect to $v^*_{l,m}$, yielding the following two equations:

$$\frac{\delta I}{\delta u^*_{l,m}} = \iint [\, A_{l,m}(u^*_{l,m} - u_{l,m}) \qquad \text{[II.2]}$$
$$- D_{l,m}\,(u^*_{l+1,m} - u^*_{l,m} + v^*_{l,m+1} - v^*_{l,m} - d_{l,m})$$
$$- Q_{l,m}\,(v^*_{l,m} - v^*_{l-1,m} - u^*_{l,m} + u^*_{l,m-1} - q_{l,m})$$
$$- E_{l,m}\,(u^*_{l-1,m+1} - u^*_{l,m} - e_{l,m}) - F_{l,m}\,(u^*_{l,m+1} - u^*_{l,m} - f_{l,m})$$
$$- G_{l,m}\,(u^*_{l+1,m+1} - u^*_{l,m} - g_{l,m}) - H_{l,m}\,(u^*_{l+1,m} - u^*_{l,m} - h_{l,m})]\,dxdy \overset{\text{set}}{=} 0$$

$$\frac{\delta I}{\delta v^*_{l,m}} = \iint [\, \hat{A}_{l,m}\,(v^*_{l,m} - v_{l,m}) \qquad \text{[II.3]}$$
$$- D_{l,m}\,(u^*_{l+1,m} - u^*_{l,m} + v^*_{l,m+1} - v^*_{l,m} - d_{l,m})$$
$$+ Q_{l,m}\,(v^*_{l,m} - v^*_{l-1,m} - u^*_{l,m} + u^*_{l,m-1} - q_{l,m})$$
$$- \hat{E}_{l,m}\,(v^*_{l-1,m+1} - v^*_{l,m} - \hat{e}_{l,m}) - \hat{F}_{l,m}\,(v^*_{l,m+1} - v^*_{l,m} - \hat{f}_{l,m})$$
$$- \hat{G}_{l,m}\,(v^*_{l+1,m+1} - v^*_{l,m} - \hat{g}_{l,m}) - \hat{H}_{l,m}\,(v^*_{l+1,m} - v^*_{l,m} - \hat{h}_{l,m})]\,dxdy \overset{\text{set}}{=} 0$$



( ) = constraint from equation [II.1]

* = U grid points

x = V grid points

○ = difference

▭ = divergence

◇ = vorticity

→ = gradient

FIGURE II-2: U COMPONENT MINIMIZATION STENCIL

In equation [II.2] group terms involving 1) $u^*_{l,m}$; 2) u* at surrounding points; 3) v* and 4) everything else.

$$\iint [\overbrace{(A_{l,m} + D_{l,m} + Q_{l,m} + E_{l,m} + F_{l,m} + G_{l,m} + H_{l,m})}^{S_{l,m}} u^*_{l,m} \qquad [II.4]$$

$$X_{l,m}\begin{cases} + (- D_{l,m} - H_{l,m})\, u^*_{l+1,m} + (-Q_{l,m})\, u^*_{l,m-1} \\ + (-E_{l,m})\, u^*_{l-1,m+1} + (-F_{l,m})\, u^*_{l,m+1} + (-G_{l,m})\, u^*_{l+1,m+1} \end{cases}$$

$$Y_{l,m}\ \{+ (D_{l,m} - Q_{l,m})\, v^*_{l,m} + (-D_{l,m})\, v^*_{l,m+1} + Q_{l,m}\, v^*_{l-1,m}$$

$$Z_{l,m}\begin{cases} - A_{l,m}\, u_{l,m} + D_{l,m}\, d_{l,m} + Q_{l,m}\, q_{l,m} + E_{l,m}\, e_{l,m} + F_{l,m}\, f_{l,m} \\ + G_{l,m}\, g_{l,m} + H_{l,m}\, h_{l,m}] \, dxdy = 0 \end{cases}$$

Group [II.3] similarly:

$$\iint [\overbrace{(\hat{A}_{l,m} + D_{l,m} + Q_{l,m} + \hat{E}_{l,m} + \hat{F}_{l,m} + \hat{G}_{l,m} + \hat{H}_{l,m})}^{\hat{S}_{l,m}} V^*_{l,m} \qquad [II.5]$$

$$\hat{X}_{l,m}\begin{cases} + (-\hat{D}_{l,m} - \hat{F}_{l,m})\, V^*_{l,m+1} + (- Q_{l,m}) V^*_{l-1,m} + (-\hat{E}_{l,m})\, V^*_{l-1,m+1} \\ + (-\hat{G}_{l,m}) V^*_{l+1,m+1} + (-\hat{H}_{l,m})\, V^*_{l+1,m} \end{cases}$$

$$\hat{Y}_{l,m}\ \{+ (D_{l,m} - Q_{l,m})\, u^*_{l,m} - D_{l,m}\, u^*_{l+1,m} + Q_{l,m}\, u^*_{l,m-1}$$

$$\hat{Z}_{l,m}\begin{cases} - \hat{A}_{l,m}\, v_{l,m} + D_{l,m}\, d_{l,m} - Q_{l,m}\, q_{l,m} + \hat{E}_{l,m} \hat{e}_{l,m} + \hat{F}_{l,m}\, \hat{f}_{l,m} \\ + \hat{G}_{l,m}\, \hat{g}_{l,m} + \hat{H}_{l,m}\, \hat{h}_{l,m}] \, dxdy = 0 \end{cases}$$

Note that all terms in S and Z except $A_{l,m}$ in $S_{l,m}$ and $- A_{l,m}u_{l,m}$ in $Z_{l,n}$ involve first-guess information which is constant during the analysis. Similar conditions hold for $\hat{S}$ and $\hat{Z}$.

Equations [II.4] and [II.5] can be written in matrix form:

$$\underline{\underline{S}}_{l,m}\,\underline{u}^* + \underline{\underline{X}}_{l,m} + \underline{\underline{Y}}_{l,m} + \underline{\underline{Z}}_{l,m} = 0 \qquad \text{[II.6]}$$

$$\hat{\underline{\underline{S}}}_{l,m}\,\underline{v}^* + \hat{\underline{\underline{X}}}_{l,m} + \hat{\underline{\underline{Y}}}_{l,m} + \hat{\underline{\underline{Z}}}_{l,m} = 0 \qquad \text{[II.7]}$$

These equations must be solved simultaneously. The method of solution used is Liebmann over-relaxation. Using a first-guess for u* and v*, equation [II.6] is, in general, not satisfied. A residual is defined by:

$$\underline{\underline{S}}_{l,m}\,\underline{u}^{*\tau} + \underline{\underline{X}}_{l,m} + \underline{\underline{Y}}_{l,m} + \underline{\underline{Z}}_{l,m} = R \qquad \text{[II.8]}$$

The superscript τ is the iteration counter. We wish to find a next guess at $\underline{u}^*$ such that the residual is zero, if the values at surrounding points do not change.

$$\underline{\underline{S}}_{l,m}\underline{u}^{*\tau+1} + \underline{\underline{X}}_{l,m} + \underline{\underline{Y}}_{l,m} + \underline{\underline{Z}}_{l,m} = 0 \qquad \text{[II.9]}$$

Subtracting [II.9] from [II.8],

$$\underline{\underline{S}}_{l,m}\,(\underline{u}^{*\tau} - \underline{u}^{*\tau+1}) = R$$

and $$\underline{u}^{*\tau+1} = \underline{u}^{*\tau} - \frac{R}{\underline{\underline{S}}_{l,m}} \,. \qquad \text{[II.10]}$$

Convergence is more rapid if the correction in [II.10] is exaggerated by the inclusion of ALFA factor.

$$\underline{u}^{*\tau+1} = \underline{u}^{*\tau} - \text{ALFA} \frac{R}{\underline{S}_{1,m}} . \qquad \text{[II.11]}$$

At a particular grid point, u* is corrected by equation [II.11] and v* is then corrected in an analogous way. In computing R from equation [II.8] or from the analogous equation in v*, the latest estimate of both u* and v* at surrounding points is used. Some of them have been changed on the current iteration and some have not. As in the scalar analysis, the field is scanned in a counter-clockwise circular sweep starting at the center and working toward the boundaries.

During each iteration through the grid, the maximum residual is checked. When it becomes less than a prescribed convergence criterion, the equations are considered solved.

D. Reevaluating the Data Weights

The validity of wind reports is judged according to the vector difference between the reported wind and the analyzed wind. The analyzed wind is obtained by interpolation from the analysis fields. The root-mean-square difference is computed and averaged for all the observations that were accepted on the current scan as a diagnostic output. If the report differs in vector magnitude from the analysis by more than the expected difference, its weight is reevaluated. The expected difference is defined as the square root of the class variance assigned to the report initially, which is the inverse of the original data weight. Define:

$$\lambda^2 = \left| \frac{W_n - W_a}{\frac{1}{A_o}} \right|$$

where W_n is the nth report, W_a is the interpolated analyzed wind, and A_o is the original report weight.

If λ^2 is greater than 1, which implies actual error is greater than expected error, the report weight is computed as:

$$A_n = \frac{2A_o}{1 + \lambda^2}$$

If λ^2 is less than 1, the report is assigned the weight A_o even if its weight was previously reduced.

E. Program Description - 63 x 63 Grid Version

1. INFODEN

As in the scalar analysis, a field is required which quantifies the density of the available observations. An influence function with a value of one at the center, reducing to zero at its periphery is superimposed at each observation location. All grid points within the circle accumulate a contribution equal to the appropriate influence value multiplied by the observations subjective weight. Once all observations have been processed, the field is written to the random file TAPE9 as IDEN. A $\log_{10}$ histogram of the resulting grid point values is computed and printed. Finally, the density value corresponding to having PPER percent of the grid points with a density less than PPER is calculated and defined as DENLIM for use by the assembly process.

2. PCTWND

All the arrays have variable dimensions. Common block /DTG/ provides the date-time group, /INFO/ the ident information and /ISW/ the switch settings. Random-access file TAPE9 must be declared on the PROGRAM card of the calling program and is used for temporary storage space within PCTWND. The data location lists, the data weight list and the initial weight of the first-guess field are

written on TAPE9. These arrays are used in the call to subroutine BKGRND, which computes matrices S, ZU and ZV. In the notation of Section II-C, a (^) referred to the v wind component. Since S is the same as $\hat{S}$, no distinction needs to be made. ZU is used in the program for Z, and ZV for $\hat{Z}$.

The convention of assigning two names to an array and separating them by a space is used here as it was in PCT. Array AI S holds the I coordinate data location list initially, but returns matrix S from BKGRND. The matrices S, ZU and ZV are written on TAPE9 and their arrays are refilled with their original fields. Weight fields E, F, G, and H are stored on TAPE9 so they can be used as work arrays later. DO loop 100 is the main loop controlling the number of cycles to be made through the program. Subroutine ASSMBL adds the influence of the observations into the first-guess field as described in Section II-B. The data list which includes reject bits set in the call to ASSMBL are written to TAPE9 for later access. After calling ASSMBL on the last analysis cycle only, subroutine PLTWIND (see Appendix) writes the data list on the plot file.

Next, the matrices S, A, ZU and ZV are read from TAPE9. Subroutine UPDATE adds A to S, -A*U to ZU and -A*V to ZV. Arrays E, F, G and H are refilled from TAPE9. Subroutine BLEND solves Equations [II.6] and [II.7], resulting in the analysis fields U and V. After restoring the current data weights into array A and DWT ZV, the data itself and its I,J location arrays are restored. Now the data weights are reevaluated by REVALWT as described in Section II-D, and a vector root-mean-square difference between the observations accepted on the current cycle and the resulting analysis of the cycle is computed. The reevaluated weights are written to TAPE9 as CURWT. If the number of cycles completed is less than NOPAS, the program continues through another entire cycle after reducing the scan radius and gross rejection limit. If the analysis is complete and the sense switch settings are appropriately set, the analysis fields and the associated divergence are printed by PRT, the U and V analysis fields are written on the plot file by subroutine PLOT, and the fields are written to the disk using the FNWC random access routine ZRANDIO.

2. BKGRND

Matrices S, ZU and ZV are computed as indicated on page II-9. The comments adequately describe each step.

3. ASSMBL

Arrays SUMU and SUMV are used to accumulate the contribution of the observations to the first-guess field. They are initialized to zero. Since the grid is staggered (see Figure II-1), the J coordinate of the U component report and the I coordinate of the V component report are decreased by .5. Then the guess U and V are interpolated to the adjusted report location. The assembly equation is the same as used in the scalar analysis but is computed for each component. If a gross error has not occurred, the product of the influence function value times the data weight times the difference of the observation and the first guess are added to SUMU and SUMV for each grid point influenced by each observation. The data weight is also accumulated in DENOM and DENOMV and packed into DENOM. It should be noted that the distance from the data location to the staggered U(I,J) and V(I,J) will be different resulting in a different influence function value for the same data report.

Gross errors are rejected by setting the last bit of the DATA word. For good data, this bit is cleared. Bogus reports cannot be rejected.

Finally, the weighted average of U and V are computed for each grid point along with accept and reject statistics.

4. UPDATE

Matrix S applies to both the U and V equations, but the terms A*U and A*V have been left out. UPDATE adds them in and two matrices result. These two are packed into array S. Also, A*U is subtracted from ZU and A*V from ZV.

5. BLEND

The two minimization equations, [II.6] and [II.7], are solved as described in detail on page II-10. No further discussion is needed.

6. REVALWT

The explanation in Section II-D is followed closely and the comments in the listing suffice to explain the code.