M.O. 693

122

AIR MINISTRY
METEOROLOGICAL OFFICE

*Scientific Paper No. 9*

# An Experiment in the Verification of Forecast Charts

by C. E. WALLINGTON, M.Sc.

LONDON: HER MAJESTY'S STATIONERY OFFICE
FOUR SHILLINGS NET

M.O. 693

AIR MINISTRY
METEOROLOGICAL OFFICE

*Scientific Paper No. 9*

# An Experiment in the Verification of Forecast Charts

by C. E. WALLINGTON, M.Sc.

*LONDON*
HER MAJESTY'S STATIONERY OFFICE
1961

Decimal Index
551·509·2:551·509·59

## Contents

# An Experiment in the Verification of Forecast Charts

by C. E. Wallington, M.Sc.

## INTRODUCTION

A number of writers have proposed various indices as measures of the success or usefulness of forecast charts, but since the evaluation of these indices usually requires lengthy computation, experience of using them and assessing their value as verification indices is rather limited. As a step towards building up such experience at the Central Forecasting Office, Dunstable, some indices have been computed for a selection of forecast charts and for charts used in a recent four-month experiment in numerical forecasting.

## ITEMS COMPUTED

The forecast and actual data are in the form of 1000-millibar charts or 500–1000-millibar thicknesses and 500-millibar heights at $15 \times 10$ grid points in the "verification area" shown in Figure 1. Three sets of contour fields were available for each case: the forecast fields, the actual fields for the same time as the forecasts and the previous (24 hours earlier) fields. The contour heights at grid points of these three fields will be denoted by letters $F$, $A$ and $P$ respectively.

The computed items (not all of which are verification items) are:

1. Mean forecast height ($\bar{F}$)
2. Mean actual height ($\bar{A}$)
3. Root mean square error in the forecast heights
4. Correlation coefficient between forecast and actual heights ($r_H$)
5. Mean forecast 24-hour height changes ($\bar{F}—\bar{P}$)
6. Root mean square forecast 24-hour height changes
7. Mean actual 24-hour height changes ($\bar{A}—\bar{P}$)
8. Root mean square actual 24-hour height changes
9. Correlation coefficient between the forecast and actual 24-hour height changes ($r_C$)
10. The Bagrov scalar index of analogy (mean of exp $[-3(F-A)^2/4s^2]$ where $s$ =seasonal standard deviation of contour height at each grid point)

Using geostrophic winds averaged over each elementary square of grid points, the following items were also computed:

11. Mean "west" wind error, "west" being used in this report to denote components in the $x$ direction
12. Mean "south" wind error, "south" being used in this report to denote components in the $y$ direction
13. Root mean square west wind error
14. Root mean square south wind error

15. Root mean square vector wind error
16. Correlation coefficient between the forecast and actual westerly wind components
17. Correlation coefficient between the forecast and actual southerly wind components
18. Stretch vector correlation between forecast and actual winds
19. The Wasko[1]* westerly index (mean of the modulus of the mean westerly wind errors along each row of the $15 \times 10$ network)
20. The Wasko[1] southerly index (mean of the modulus of the mean southerly wind errors along each column of the $15 \times 10$ network)
21. The Teweles and Wobus[2] index (the sum of the contour gradient errors between each pair of the adjacent grid points divided by the sum of the greater of the forecast and actual gradients)
22. The Bagrov[3] vector index of analogy (mean of $\exp[-3\ (\text{vector wind error})^2/16S^2]$, where $S$=standard vector deviation of the winds at each grid point)

All items other than coefficients were obtained in units of metres or knots. Some of the items listed are not in common use and need further description.

Wasko[1] devised his index as a check on mean route winds of the type required in aviation forecasting. He studied the accuracy of forecasting headwind components along east–west routes, but there is no reason why east–west routes alone should be considered and in this experiment the index is computed for east–west and north–south components over the whole verification area.

In the Teweles and Wobus[2] index the denominator is intended to take into account the difficulty of forecasting strong contour gradients; when either the forecast or the actual contour gradients are large then the denominator is also large and the index (which is zero for perfect forecasting) is reduced. Teweles and Wobus found that in practice their index ranged from about 0·4 to 1·1.

The Bagrov[3] indices are such that perfect forecasting yields a value of 1 while a forecast of the climatological normals would be expected to yield a value of 0·632 and verification with charts selected at random would yield, on average, 0·5.

The computing programme enabled the twenty-two items listed to be computed for the 1000-millibar, 500–1000-millibar thickness and 500-millibar fields for whatever charts were available over the full verification area and over selected sub-areas which are shown in Figure 2. Winds could also be computed as geostrophic means over overlapping grid "squares" of size 1, 2, 3 and 4 grid lengths. Thus the computing programme itself provided a means of studying the effect of verification area and grid length on the verification indices.

## VARIATION OF INDICES WITH AREA

The effect on selected verification indices of reducing or changing the verification area was tested for eight of the 1000-millibar and 500-millibar forecast charts prepared by the Central Forecasting Office, Dunstable. The actual and forecast charts are shown in Figures 3 and 4, while the results are tabled below. All wind indices in these tables relate to contour gradients measured over one grid length.

In Tables I and II it is apparent that for both the 1000-millibar and 500-millibar levels the maximum differences between the $15 \times 10$ area and the sub-area values suggest that root

*The superscript figures refer to the bibliography on page 17.

TABLE I. *Root mean square height errors (item 3)—1000 millibars*

| C.F.O. 24-hour forecast for: | Areas indicated in Figure 2 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | metres | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 3(*a*)) | 69 | 63 | 55 | 48 | 85 | 16 | 72 | 57 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 3(*b*)) | 88 | 82 | 69 | 45 | 42 | 31 | 123 | 38 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 3(*c*)) | 57 | 60 | 59 | 48 | 54 | 38 | 77 | 49 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 3(*d*)) | 44 | 44 | 40 | 38 | 66 | 19 | 32 | 46 |
| 1200 G.M.T., 28 May 1957 (Figure 3(*e*)) | 29 | 24 | 23 | 25 | 31 | 33 | 19 | 29 |
| 1200 G.M.T., 2 July 1957 (Figure 3(*f*)) | 47 | 48 | 51 | 58 | 35 | 17 | 49 | 46 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 3(*g*) | 56 | 63 | 71 | 77 | 108 | 36 | 37 | 64 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 3(*h*)) | 35 | 33 | 33 | 31 | 33 | 22 | 26 | 35 |
| Maximum difference between adjacent columns | 7 | 13 | 24 | 31 | 72 | 92 | 85 | |

TABLE II. *Root mean square height errors (item 3)—500 millibars*

| C.F.O. 24-hour forecast for: | Areas indicated in Figure 2 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | metres | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 4(*a*)) | 89 | 92 | 95 | 100 | 60 | 40 | 129 | 77 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 4(*b*)) | 75 | 71 | 58 | 39 | 32 | 47 | 78 | 66 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 4(*c*)) | 39 | 40 | 41 | 41 | 43 | 38 | 26 | 41 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 4(*d*)) | 77 | 84 | 96 | 113 | 64 | 39 | 122 | 70 |
| 1200 G.M.T., 28 May 1957 (Figure 4(*e*)) | 40 | 34 | 26 | 20 | 12 | 25 | 38 | 25 |
| 1200 G.M.T., 2 July 1957 (Figure 4(*f*)) | 58 | 58 | 66 | 82 | 38 | 28 | 79 | 59 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 4(*g*)) | 81 | 87 | 102 | 113 | 63 | 23 | 124 | 80 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 4(*h*)) | 51 | 49 | 79 | 44 | 33 | 26 | 67 | 46 |
| Maximum difference between adjacent columns | 7 | 30 | 35 | 50 | 40 | 101 | 52 | |

mean square height errors over areas smaller than the 13 × 8 area vary considerably from one area to another. Since the maximum difference between the 15 × 10 and 13 × 8 values is only seven metres for both the 1000-millibar and 500-millibar levels, it seems likely that the area of 15 × 10 points is about the minimum size for using the root mean square height error as an index of the accuracy of the broad-scale features, the assumption being that the difference between a 1·5 × 10 and a slightly larger size would be less than seven metres. Unfortunately the computing programme used could not be easily modified to deal with a slightly larger area.

Although the results of only eight cases have been listed in these tables, they are sufficient to illustrate how unrepresentative a small area can be of the broader-scale accuracy of the forecast chart. The results also suggest that if the root mean square height error for the 15 × 10 area is low, say 40 metres or below, then it is likely to be low over most sub-areas, but if the full area value is high, say 50 metres or more, then there may be considerable variability of the root mean square height error from one sub-area to the next.

The distribution of the correlation coefficients listed in Tables III and IV also shows a tendency to approach a representative value as the verification area is increased to the 15 × 10 size. As with the root mean square contour height error the correlation coefficient applied to the 15 × 10 area is an index of the quality of the sub-areas only when the overall accuracy is very high. A correlation coefficient of about 0·98 or more indicates that the correlations will be high over most sub-areas but lower overall coefficients give no insight into the accuracy of any particular sub-area.

TABLE III. *Contour height correlation coefficients (item 4)—1000 millibars*

| C.F.O. 24-hour forecast for: | *Areas indicated in Figure 2* 15 × 10 | 13 × 8 | 11 × 6 | 9 × 4 | Britain 5 × 5 | France 5 × 5 | Atlantic 5 × 5 | Europe 10 × 10 |
|---|---|---|---|---|---|---|---|---|
| 1800 G.M.T., 11 Jan. 1957 (Figure 3(*a*)) | 0·75 | 0·74 | 0·73 | 0·73 | 0·81 | 0·95 | 0·79 | 0·78 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 3(*b*)) | 0·74 | 0·73 | 0·75 | 0·83 | 0·74 | 0·93 | —0·19 | 0·94 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 3(*c*)) | 0·88 | 0·88 | 0·86 | 0·84 | 0·94 | 0·69 | 0·27 | 0·94 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 3(*d*)) | 0·90 | 0·87 | 0·83 | 0·75 | 0·74 | 0·84 | 0·86 | 0·88 |
| 1200 G.M.T., 28 May 1957 (Figure 3(*e*)) | 0·88 | 0·90 | 0·88 | 0·81 | 0·91 | 0·89 | 0·91 | 0·94 |
| 1200 G.M.T., 2 July 1957 (Figure 3(*f*)) | 0·50 | 0·53 | 0·60 | 0·65 | 0·65 | 0·72 | 0·79 | 0·48 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 3(*g*)) | 0·84 | 0·85 | 0·85 | 0·82 | 0·81 | 0·42 | 0·24 | 0·85 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 3(*h*)) | 0·92 | 0·91 | 0·85 | 0·65 | 0·68 | 0·34 | 0·93 | 0·94 |
| Maximum difference between adjacent columns | | 0·07 | 0·1 | 0·2 | 0·25 | 0·49 | 1·12 | 1·13 |

TABLE IV. *Contour height correlation coefficients (item 4)—500 millibars*

| C.F.O. 24-hour forecast for: | 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | *Areas indicated in Figure 2* | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 4(*a*)) | 0·88 | 0·84 | 0·77 | 0·64 | 0·92 | 0·90 | 0·95 | 0·89 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 4(*b*)) | 0·92 | 0·91 | 0·92 | 0·94 | 0·94 | 0·91 | 0·34 | 0·93 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 4(*c*)) | 0·99 | 0·99 | 0·99 | 0·98 | 0·99 | 0·94 | 0·97 | 0·99 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 4(*d*)) | 0·96 | 0·96 | 0·94 | 0·90 | 0·94 | 0·85 | 0·97 | 0·95 |
| 1200 G.M.T., 28 May 1957 (Figure 4(*e*)) | 0·92 | 0·94 | 0·96 | 0·97 | 0·99 | 0·95 | 0·91 | 0·95 |
| 1200 G.M.T., 2 July 1957 (Figure 4(*f*)) | 0·89 | 0·90 | 0·91 | 0·90 | 0·91 | 0·96 | 0·96 | 0·89 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 4(*g*)) | 0·92 | 0·91 | 0·87 | 0·79 | 0·86 | 0·98 | 0·39 | 0·93 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 4(*h*)) | 0·98 | 0·97 | 0·96 | 0·94 | 0·97 | 0·94 | 0·98 | 0·97 |
| Maximum difference between adjacent columns | 0·04 | 0·07 | 0·13 | 0·28 | 0·12 | 0·59 | 0·59 | |

As indicated by Tables V and VI the 24-hour height change correlation coefficients also show considerable variations from one sub-area to another but the differences between the 15×10 and 13×8 areas are small.

TABLE V. *Correlation coefficients for 24-hour changes (item 5)—1000 millibars*

| C.F.O. 24-hour forecast for: | 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | *Areas indicated in Figure 2* | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 3(*a*)) | 0·54 | 0·47 | 0·42 | 0·38 | 0·65 | 0·72 | 0·58 | 0·63 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 3(*b*)) | 0·53 | 0·57 | 0·63 | 0·69 | 0·92 | 0·33 | 0·32 | 0·82 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 3(*c*)) | 0·79 | 0·80 | 0·80 | 0·83 | 0·85 | 0·15 | 0·70 | 0·87 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 3(*d*)) | 0·61 | 0·66 | 0·74 | 0·78 | 0·09 | 0·46 | 0·82 | 0·21 |
| 1200 G.M.T., 28 May 1957 (Figure 3(*e*)) | 0·61 | 0·66 | 0·61 | 0·36 | 0·75 | —0·26 | 0·90 | 0·66 |
| 1200 G.M.T., 2 July 1957 (Figure 3(*f*)) | 0·70 | 0·74 | 0·79 | 0·80 | 0·74 | 0·84 | 0·91 | 0·70 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 3(*g*)) | 0·47 | 0·54 | 0·64 | 0·66 | 0·91 | 0·29 | 0·30 | 0·49 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 3(*h*)) | 0·73 | 0·77 | 0·80 | 0·86 | 0·74 | 0·54 | 0·75 | 0·73 |
| Maximum difference between adjacent columns | 0·07 | 0·10 | 0·25 | 0·69 | 1·01 | 1·16 | 0·61 | |

TABLE VI. *Correlation coefficients for 24-hour height changes (item 5)—500 millibars*

| C.F.O. 24-hour forecast for: | Areas indicated in Figure 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Britain | France | Atlantic | Europe |
| | 15×10 | 13×8 | 11×6 | 9×4 | 5×5 | 5×5 | 5×5 | 10×10 |
| 1800 G.M.T., 11 Jan. 1957 (Figure 4(*a*)) | 0·60 | 0·58 | 0·58 | 0·59 | 0·91 | 0·92 | 0·09 | 081 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 4(*b*)) | 0·52 | 0·60 | 0·72 | 0·84 | 0·94 | 0·55 | 0·76 | 0·67 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 4(*c*)) | 0·84 | 0·85 | 0·86 | 0·88 | 0·83 | 0·45 | 0·73 | 0·90 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 4(*d*)) | 0·67 | 0·75 | 0·83 | 0·91 | 0·74 | 0·88 | 0·89 | 0·73 |
| 1200 G.M.T., 28 May 1957 (Figure 4(*e*)) | 0·46 | 0·41 | 0·33 | 0·61 | 0·77 | 0·52 | 0·47 | 0·28 |
| 1200 G.M.T., 2 July 1957 (Figure 4(*f*)) | 0·78 | 0·83 | 0·87 | 0·83 | 0·45 | 0·90 | 0·83 | 0·61 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 4(*g*)) | 0·63 | 0·60 | 0·47 | 0·29 | 0·73 | 0·78 | 0·58 | 0·65 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 4(*h*)) | 0·80 | 0·84 | 0·87 | 0·88 | 0·86 | 0·65 | 0·75 | 0·83 |
| Maximum difference between adjacent columns | 0·08 | 0·13 | 0·28 | 0·44 | 0·45 | 0·83 | 0·72 | |

Tables VII and VIII yield similar conclusions to those formed for the root mean square height error and correlation indices, namely that there is little to choose between at 15×10 and 13×8 verification areas but that smaller areas are representative of nothing but themselves.

TABLE VII. *Root mean square vector wind errors (item 15)—1000 millibars*

| C.F.O. 24-hour forecast for: | Areas indicated in Figure 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Britain | France | Atlantic | Europe |
| | 15×10 | 13×8 | 11×6 | 9×4 | 5×5 | 5×5 | 5×5 | 10×10 |
| | knots | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 3(*a*)) | 25 | 22 | 18 | 16 | 23 | 13 | 28 | 19 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 3(*b*)) | 30 | 32 | 35 | 33 | 21 | 21 | 48 | 21 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 3(*c*)) | 17 | 16 | 15 | 16 | 18 | 14 | 17 | 16 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 3(*d*)) | 21 | 22 | 21 | 20 | 32 | 12 | 18 | 22 |
| 1200 G.M.T., 28 May 1957 (Figure 3*e*)) | 13 | 10 | 9 | 8 | 8 | 19 | 8 | 13 |
| 1200 G.M.T., 2 July 1957 (Figure 3(*f*)) | 25 | 27 | 27 | 21 | 13 | 13 | 16 | 26 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 3(*g*)) | 21 | 21 | 21 | 22 | 27 | 20 | 20 | 22 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 3(*h*)) | 15 | 14 | 14 | 14 | 13 | 13 | 14 | 13 |
| Maximum difference between adjacent columns | 3 | 4 | 6 | 12 | 20 | 27 | 27 | |

TABLE VIII. *Root mean square vector wind errors (item 15)—500 millibars*

| C.F.O. 24-hour forecast for: | 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | *Areas indicated in Figure 2* (knots) | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 4(*a*)) | 37 | 36 | 36 | 36 | 24 | 38 | 38 | 36 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 4(*b*)) | 31 | 30 | 27 | 23 | 18 | 40 | 30 | 27 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 4(*c*)) | 14 | 13 | 13 | 13 | 12 | 13 | 13 | 13 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 4(*d*)) | 28 | 29 | 30 | 28 | 31 | 15 | 23 | 28 |
| 1200 G.M.T., 28 May 1957 (Figure 4(*e*)) | 16 | 15 | 12 | 11 | 6 | 14 | 18 | 15 |
| 1200 G.M.T., 2 July 1957 (Figure 4(*f*)) | 27 | 29 | 25 | 22 | 15 | 12 | 20 | 27 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 4(*g*)) | 33 | 33 | 32 | 30 | 31 | 15 | 37 | 30 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 4(*h*)) | 19 | 19 | 19 | 20 | 14 | 20 | 21 | 19 |
| Maximum difference between adjacent columns | 2 | 4 | 4 | 12 | 22 | 22 | 7 | |

Although the stretch vector correlations show very considerable variations in the sub-areas, the differences between the 15×10 and 13×8 areas are not large.

Study of the remaining verification indices computed for various areas showed that they all behaved in a similar way to that of the indices in Tables IX and X.

TABLE IX. *Stretch vector correlation (item 18)—1000 millibars*

| C.F.O. 24-hour forecast for: | 15×10 | 13×8 | 11×6 | 9×4 | Britain 5×5 | France 5×5 | Atlantic 5×5 | Europe 10×10 |
|---|---|---|---|---|---|---|---|---|
| | *Areas indicated in Figure 2* | | | | | | | |
| 1800 G.M.T., 11 Jan. 1957 (Figure 3(*a*)) | 0·61 | 0·65 | 0·70 | 0·69 | 0·48 | 0·88 | —0·41 | 0·77 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 3(*b*)) | 0·45 | 0·36 | 0·17 | 0·04 | 0·16 | 0·25 | —0·12 | 0·37 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 3(*c*)) | 0·70 | 0·71 | 0·83 | 0·90 | 0·56 | 0·73 | 0·79 | 0·75 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 3(*d*)) | 0·45 | 0·38 | 0·30 | 0·25 | —0·01 | 0·55 | 0·17 | 0·47 |
| 1200 G.M.T., 28 May 1957 (Figure 3(*e*)) | 0·75 | 0·81 | 0·83 | 0·85 | 0·54 | 0·00 | 0·54 | 0·63 |
| 1200 G.M.T., 2 July 1957 (Figure 3(*f*)) | 0·33 | 0·23 | 0·23 | 0·48 | 0·67 | 0·07 | 0·55 | 0·16 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 3(*g*)) | 0·58 | 0·53 | 0·46 | 0·35 | 0·62 | 0·72 | 0·25 | 0·65 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 3(*h*)) | 0·69 | 0·70 | 0·71 | 0·64 | 0·70 | 0·31 | 0·74 | 0·70 |
| Maximum difference between adjacent columns | 0·10 | 0·16 | 0·25 | 0·34 | 0·60 | 1·29 | 1·18 | |

TABLE X. *Stretch vector correlation (item 18)—500 millibars*

| C.F.O. 24-hour forecast for: | *Areas indicated in Figure 2* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Britain | France | Atlantic | Europe |
| | 15×10 | 13×8 | 11×6 | 9×4 | 5×5 | 5×5 | 5×5 | 10×10 |
| 1800 G.M.T., 11 Jan. 1957 (Figure 4(*a*)) | 0·56 | 0·62 | 0·64 | 0·59 | 0·95 | 0·09 | —0·26 | 0·62 |
| 0001 G.M.T., 6 Feb. 1957 (Figure 4(*b*)) | 0·92 | 0·91 | 0·92 | 0·94 | 0·94 | 0·91 | 0·34 | 0·93 |
| 1200 G.M.T., 17 Mar. 1957 (Figure 4(*c*)) | 0·84 | 0·82 | 0·72 | 0·70 | 0·85 | 0·95 | 0·62 | 0·88 |
| 1200 G.M.T., 18 Apr. 1957 (Figure 4(*d*)) | 0·63 | 0·58 | 0·53 | 0·50 | 0·42 | 0·83 | 0·00 | 0·69 |
| 1200 G.M.T., 28 May 1957 (Figure 4(*e*)) | 0·85 | 0·87 | 0·90 | 0·90 | 0·96 | 0·85 | 0·80 | 0·79 |
| 1200 G.M.T., 2 July 1957 (Figure 4(*f*)) | 0·71 | 0·69 | 0·78 | 0·86 | 0·47 | 0·75 | 0·55 | 0·40 |
| 1800 G.M.T., 25 Aug. 1957 (Figure 4(*g*)) | 0·38 | 0·30 | 0·27 | 0·36 | 0·77 | 0·63 | —0·36 | 0·64 |
| 0600 G.M.T., 16 Oct. 1957 (Figure 4(*h*)) | 0·78 | 0·78 | 0·79 | 0·73 | 0·68 | 0·30 | 0·57 | 0·76 |
| Maximum difference between adjacent columns | | 0·08 | 0·10 | 0·09 | 0·41 | 0·86 | 0·99 | 1·00 |

VARIATION OF GRID LENGTH

The computing programme enabled the wind indices to be computed over squares of side 1, 2, 3 and 4 grid lengths. The results of these computations for eight cases are listed in Table XI.

As was expected the effect of increasing the number of grid lengths used in measuring the geostrophic winds is to reduce the indicated errors. In the cases examined the approximate effect of quadrupling the grid lengths is to halve the root mean square wind errors and to increase (with one exception) the correlation coefficients by variable but significant amounts. The Teweles and Wobus, Wasko and Bagrov indices are also less stringent as verification indices when winds are measured over several grid lengths.

DISTRIBUTION OF VERIFICATION INDICES

During the first four months of 1959 a routine experiment in numerical forecasting was carried out at Dunstable. The data tapes for this experiment are in a form suitable for the computing programme which calculates the indices already listed. Forty-two of the forecasts (all 24-hour forecasts from 0001 G.M.T. on most of the Tuesdays, Wednesdays and Thursdays of the four-month period) together with the corresponding set of forecasts prepared by the Central Forecasting Office were used to study the ranges and distribution of the verification indices.

The same forecasts were also judged by six assessors who, studying the charts independently, awarded points on a scale of 0–10 for their subjective assessment of the quality of the forecast. The combined results of the numerical and Central Forecasting Office forecasts for the 1000-millibar level are shown in Table XII.

TABLE XI. *Wind verification indices, 15 × 10 area—1000-millibar level*

| C.F.O. 24-hour forecast for: | Grid length | r.m.s. west wind error (item 13) kt | r.m.s. south wind error (item 14) kt | r.m.s. vector wind error (item 15) kt | Corr. coeff. west winds (item 16) | Corr. coeff. south winds (item 17) | Stretch vector corr. coeff. (item 18) | Wasko west index (item 19) kt | Wasko south index (item 20) kt | Teweles and Wobus index (item 21) | Bagrov vector index (item 22) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1800 G.M.T., 11 Jan. 1957 | 4 | 11 | 10 | 14 | 0·84 | 0·75 | 0·76 | 8·6 | 7·8 | 0·51 | 0·91 |
| (Figure 3(*a*)) | 3 | 12 | 12 | 17 | 0·79 | 0·72 | 0·72 | 9·6 | 8·4 | 0·53 | 0·87 |
| | 2 | 14 | 15 | 21 | 0·72 | 0·67 | 0·67 | 10·0 | 9·4 | 0·55 | 0·83 |
| | 1 | 18 | 18 | 25 | 0·62 | 0·63 | 0·61 | 11·1 | 9·3 | 0·59 | 0·78 |
| 0001 G.M.T., 6 Feb. 1957 | 4 | 5 | 15 | 16 | 0·85 | 0·72 | 0·73 | 1·5 | 10·8 | 0·45 | 0·90 |
| (Figure 3(*b*)) | 3 | 9 | 18 | 20 | 0·73 | 0·62 | 0·62 | 2·6 | 12·4 | 0·52 | 0·85 |
| | 2 | 14 | 21 | 25 | 0·61 | 0·52 | 0·51 | 4·0 | 14·2 | 0·59 | 0·80 |
| | 1 | 19 | 24 | 30 | 0·54 | 0·42 | 0·44 | 4·7 | 14·2 | 0·65 | 0·72 |
| 1200 G.M.T., 17 Mar. 1957 | 4 | 7 | 5 | 8 | 0·82 | 0·89 | 0·86 | 5·2 | 3·0 | 0·40 | 0·93 |
| (Figure 3(*c*)) | 3 | 8 | 6 | 10 | 0·78 | 0·86 | 0·82 | 5·6 | 3·5 | 0·45 | 0·91 |
| | 2 | 11 | 8 | 13 | 0·71 | 0·84 | 0·77 | 7·1 | 3·8 | 0·50 | 0·86 |
| | 1 | 14 | 11 | 17 | 0·65 | 0·76 | 0·70 | 7·3 | 4·4 | 0·56 | 0 80 |
| 1200 G.M.T., 18 Apr. 1957 | 4 | 7 | 6 | 9 | 0·74 | 0·56 | 0·66 | 3·9 | 4·2 | 0·40 | 0·91 |
| (Figure 3(*d*)) | 3 | 8 | 9 | 12 | 0·76 | 0·43 | 0·62 | 3·9 | 5·8 | 0·48 | 0·89 |
| | 2 | 10 | 12 | 16 | 0·72 | 0·31 | 0·55 | 4·1 | 6·8 | 0·58 | 0·84 |
| | 1 | 14 | 16 | 21 | 0·61 | 0·23 | 0·45 | 4·3 | 7·5 | 0·66 | 0·77 |
| 1200 G.M.T., 28 May 1957 | 4 | 3 | 5 | 6 | 0·93 | 0·91 | 0·91 | 1·4 | 4·1 | 0·34 | 0·96 |
| (Figure 3(*e*)) | 3 | 5 | 6 | 8 | 0·89 | 0·89 | 0·88 | 2·6 | 4·3 | 0·39 | 0·94 |
| | 2 | 7 | 7 | 10 | 0·84 | 0·84 | 0·83 | 3·4 | 4·8 | 0·44 | 0·91 |
| | 1 | 9 | 9 | 13 | 0·76 | 0·75 | 0·75 | 4·1 | 5·3 | 0·50 | 0·88 |
| 1200 G.M.T., 2 July 1957 | 4 | 11 | 6 | 13 | —0·01 | 0·79 | 0·46 | 8·1 | 3·9 | 0·71 | 0·85 |
| (Figure 3(*f*)) | 3 | 14 | 9 | 17 | —0·04 | 0·68 | 0·41 | 9·1 | 5·7 | 0·73 | 0·77 |
| | 2 | 16 | 14 | 21 | 0·10 | 0·55 | 0·36 | 9·7 | 7·5 | 0·76 | 0·74 |
| | 1 | 18 | 17 | 25 | 0·15 | 0·45 | 0·33 | 9·6 | 8·6 | 0·77 | 0·70 |
| 1800 G.M.T., 25 Aug. 1957 | 4 | 6 | 11 | 12 | 0·71 | 0·51 | 0·58 | 4·7 | 9·1 | 0·50 | 0·83 |
| (Figure 3(*g*)) | 3 | 8 | 12 | 14 | 0·73 | 0·57 | 0·63 | 5·3 | 9·6 | 0·55 | 0·80 |
| | 2 | 10 | 14 | 17 | 0·71 | 0·57 | 0·63 | 6·1 | 9·7 | 0·59 | 0·74 |
| | 1 | 13 | 16 | 21 | 0·64 | 0·52 | 0·58 | 6·3 | 10·1 | 0·65 | 0·66 |
| 0600 G.M.T., 16 Oct. 1957 | 4 | 6 | 4 | 7 | 0·95 | 0·81 | 0·90 | 5·0 | 2·9 | 0·40 | 0·96 |
| (Figure 3(*h*)) | 3 | 7 | 6 | 9 | 0 91 | 0·71 | 0·84 | 4·9 | 3·7 | 0·46 | 0·95 |
| | 2 | 9 | 8 | 12 | 0·85 | 0·64 | 0·77 | 4·9 | 4·3 | 0·52 | 0·92 |
| | 1 | 11 | 10 | 15 | 0·74 | 0·62 | 0·69 | 4·5 | 4·0 | 0·57 | 0·88 |

It is apparent from Table XII that the distributions of some indices such as the contour height correlation coefficient (item 4), the change correlation (item 9), the root mean square south wind error (item 14), the south wind correlation (item 17), the Wasko south wind index (item 20), the Teweles and Wobus index (item 21) and the Bagrov vector index (item 22) are leptokurtic by comparison with the distributions of the root mean square height error (item 3), the Bagrov scalar index (item 10), the root mean square vector error (item 15) and the Wasko west wind index (item 19).

TABLE XII. *Range and distribution of verification indices, 15 × 10 area—1000-millibar level*

| Verification item | *Range* Min. | Max. | *Range divided into ten intervals* 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 r.m.s. height error | 25 m | 126 m | 9 | 13 | 7 | 9 | 16 | 7 | 9 | 5 | 6 | 3 |
| 4 corr. coeff. for heights | 0·49 | 0·97 | 1 | 1 | 3 | 3 | 5 | 7 | 11 | 21 | 20 | 12 |
| 9 change corr. coeff. | —0·11 | 0·94 | 1 | 0 | 2 | 7 | 1 | 5 | 19 | 18 | 22 | 9 |
| 10 Bagrov scalar index | 0·40 | 0·96 | 2 | 5 | 6 | 4 | 10 | 9 | 15 | 19 | 11 | 3 |
| 13 r.m.s. error west wind | 9 kt | 30 kt | 11 | 18 | 13 | 10 | 7 | 5 | 5 | 6 | 5 | 4 |
| 14 r.m.s. error south wind | 8 kt | 40 kt | 10 | 19 | 22 | 12 | 10 | 6 | 3 | 1 | 0 | 1 |
| 15 r.m.s. vector wind error | 12 kt | 45 kt | 5 | 16 | 17 | 17 | 10 | 7 | 5 | 3 | 2 | 2 |
| 16 west wind corr. coeff. | 0·08 | 0·89 | 1 | 2 | 3 | 2 | 8 | 10 | 13 | 19 | 19 | 7 |
| 17 south wind corr. coeff. | —0·22 | 0·88 | 1 | 0 | 2 | 2 | 7 | 10 | 9 | 11 | 21 | 21 |
| 18 stretch vector corr. | 0·07 | 0·85 | 1 | 1 | 2 | 4 | 6 | 7 | 14 | 22 | 19 | 8 |
| 19 Wasko west index | 2 kt | 15 kt | 10 | 11 | 17 | 16 | 10 | 7 | 7 | 4 | 1 | 1 |
| 20 Wasko south index | 3 kt | 28 kt | 9 | 22 | 26 | 17 | 3 | 4 | 1 | 1 | 0 | 1 |
| 21 Teweles and Wobus index | 0·20 | 0·85 | 1 | 0 | 0 | 4 | 14 | 29 | 15 | 13 | 6 | 2 |
| 22 Bagrov vector index | 0·24 | 0·90 | 1 | 0 | 0 | 1 | 4 | 4 | 16 | 23 | 30 | 5 |
| 23 subjective assessment | 0 | 10 | 0 | 0 | 2 | 10 | 18 | 26 | 22 | 6 | 0 | 0 |

Winds are over one grid length.
42 numerical and 42 C.F.O. 24-hour forecasts were used in the period January to May 1959.

The dimodal character of the root mean square height error distribution is partly, but not entirely, due to the distribution being made up of both Central Forecasting Office forecasts and numerical forecasts in which height errors tend to err on the high side, but since an important aspect of verification will be to compare numerical techniques with subjective methods it is of interest to consider the combined distribution of indices.

### CORRELATION BETWEEN VERIFICATION INDICES

It is pertinent to inquire whether or not any one of the listed verification indices is effectively the same type of measure of forecasting success as any of the others. Also it is important to know how the computed indices compare with the subjective assessments. As a rough indication of the effective similarity between indices, appropriate correlation coefficients were calculated from the verification values computed for the 84 forecasts mentioned in the last section. These correlations are shown in Table XIII.

The highest correlations between the calculated indices and the subjective assessment are 0·67 for the root mean square height error (item 3) and 0·60 for the Bagrov scalar index (item 10). Significant correlations are also found between the root mean square height error and all of the other indices except the wind correlations (items 16, 17 and 18). The root mean square vector wind error (item 15) is also moderately well correlated with most of the other indices, and it could be claimed that this index coupled with the root mean square height error virtually summarize the verification scores given by most of the other indices.

An index not covered by the root mean square height error or the root mean square vector wind error is the change correlation coefficient (item 9). This index forms a basically different type of test from the other verification indices.

TABLE XIII. *Correlations between verification indices, 15 × 10 area—1000-millibar level*

| | Verification item | Item number 3 | 4 | 9 | 10 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | r.m.s. height error | | —·53 | —·54 | —·84 | ·72 | ·66 | ·75 | —·37 | —·32 | —·42 | ·62 | ·56 | ·50 | —·57 | —·67 |
| 4 | corr. coeff. for heights | —·53 | | ·53 | ·45 | —·52 | —·55 | —·57 | ·44 | ·62 | ·65 | —·36 | —·52 | —·62 | ·56 | ·45 |
| 9 | change corr. coeff. | —·54 | ·53 | | ·35 | —·59 | —·49 | —·58 | ·49 | ·40 | ·55 | —·51 | —·37 | —·45 | ·46 | ·34 |
| 10 | Bagrov scalar index | —·84 | ·45 | ·35 | | —·47 | —·43 | —·52 | ·10 | ·12 | ·13 | —·41 | —·49 | —·26 | ·62 | ·60 |
| 13 | r.m.s. error west wind | ·72 | —·52 | —·59 | —·49 | | ·71 | ·89 | —·62 | —·36 | —·57 | ·71 | ·43 | ·53 | —·54 | —·43 |
| 14 | r.m.s. error south wind | ·66 | —·55 | —·49 | —·43 | ·71 | | ·94 | —·37 | —·45 | —·55 | ·35 | ·83 | ·48 | —·66 | —·37 |
| 15 | r.m.s. vector wind error | ·75 | —·57 | —·58 | —·52 | ·89 | ·94 | | —·51 | —·44 | —·60 | ·54 | ·71 | ·54 | —·66 | —·43 |
| 16 | west wind corr. coeff. | —·37 | ·44 | ·49 | ·10 | —·62 | —·37 | —·51 | | ·37 | ·78 | —·56 | —·19 | —·64 | ·25 | ·24 |
| 17 | south wind corr. coeff. | —·32 | ·62 | ·40 | ·12 | —·36 | —·45 | —·44 | ·37 | | ·84 | —·22 | —·46 | —·75 | ·34 | ·32 |
| 18 | stretch vector corr. | —·42 | ·65 | ·55 | ·13 | —·57 | —·55 | —·60 | ·78 | ·84 | | —·43 | —·47 | —·84 | ·38 | ·30 |
| 19 | Wasko west index | ·62 | —·36 | —·51 | —·41 | ·71 | ·35 | ·54 | —·56 | —·22 | —·43 | | ·17 | ·48 | —·39 | —·42 |
| 20 | Wasko south index | ·56 | —·52 | —·37 | —·49 | ·43 | ·83 | ·71 | —·19 | —·46 | —·47 | ·17 | | ·42 | —·65 | —·33 |
| 21 | Teweles and Wobus index | ·50 | —·62 | —·45 | —·26 | ·53 | ·48 | ·54 | —·64 | —·75 | —·84 | ·48 | ·42 | | —·39 | —·41 |
| 22 | Bagrov vector index | —·57 | —·56 | ·46 | ·62 | —·54 | —·66 | —·66 | ·25 | ·34 | ·38 | —·39 | —·65 | ·39 | | ·36 |
| 23 | subjective assessment | —·67 | ·45 | ·34 | ·60 | —·43 | —·37 | —·43 | ·24 | ·32 | ·30 | —·42 | —·33 | —·41 | ·36 | |

42 numerical and 42 C.F.O. 24-hour forecasts were used in the period January to May 1959.

The mainly low correlations between the calculated indices and the subjective assessment cannot be used to infer that the indices themselves are useless as guides to the quality of the forecast charts. The degree of subjectivity in assessing the quality of forecast charts by inspection may be illustrated by the correlation coefficients between the individual subjective assessments made by the team of six judges. These correlation coefficients are reproduced in Table XIV.

TABLE XIV. *Correlations between subjective assessments, 15 × 10 area—1000-millibar level*

| Assessor | A | B | C | D | E | F | Mean assessment |
|---|---|---|---|---|---|---|---|
| A | — | 0·33 | 0·44 | 0·32 | 0·49 | 0·39 | 0·65 |
| B | 0·33 | — | 0·43 | 0·44 | 0·42 | 0·41 | 0·64 |
| C | 0·44 | 0·43 | — | 0·57 | 0·58 | 0·34 | 0·78 |
| D | 0·32 | 0·44 | 0·57 | — | 0·41 | 0·45 | 0·72 |
| E | 0·49 | 0·42 | 0·58 | 0·41 | — | 0·47 | 0·81 |
| F | 0·39 | 0·41 | 0·34 | 0·45 | 0·47 | — | 0·70 |
| Mean assessment | 0·65 | 0·64 | 0·78 | 0·72 | 0·81 | 0·70 | — |

42 numerical and 42 C.F.O. 24-hour forecasts were used in the period January to May 1959.

These somewhat low correlations between one set of opinions and another suggest that the degree of subjectivity in individual assessment of a chart is perhaps higher than commonly supposed.

### EFFECT OF SYSTEMATIC DISTORTIONS

Inaccuracy of the forecast chart may arbitrarily but conveniently be considered to arise from combinations of:

(i) errors in local rates of change

(ii) displacement errors

(iii) over-emphasis of cyclonic or anticyclonic gradients.

It is seldom feasible to isolate these types of errors and assess them independently, but to form a tentative idea of their effect on the verification indices a few charts were prepared for use as artificial forecasts.

To provide an example of the effect of local rates of change, charts were constructed with values of contour heights ($F$) at the grid points given by:

$F=2A-P$ for $2\times$ the actual 24-hour rate of change

$F=\frac{1}{2}(3A-P)$ for $1\frac{1}{2}\times$ the actual 24-hour rate of change

$F=\frac{1}{2}(A+P)$ for $\frac{1}{2}\times$ the actual 24-hour rate of change

$F=2P-A$ for $-1\times$ the actual 24-hour rate of change

where $P$ and $A$ denote the actual contour heights at the beginning and end respectively of the 24-hour forecast period.

The indices obtained by making such artificial forecasts for the 1000-millibar and 500-millibar contour fields for 1800 G.M.T., 25 August 1957 are given in Tables XV and XVI, and a selection of the charts are included as Figures 5 and 6 of this paper.

To provide examples of displacement effects a few artificial forecast charts were constructed by simply moving the actual contour fields one or two grid lengths towards the south and/or east. The results for a 1000-millibar and a 500-millibar field are listed in Tables XVII and XVIII, while Figure 7 shows an example of the effect of an easterly displacement at the 500-millibar level.

TABLE XV. *Effect of artificial rates of change, 15 × 10 area, 1800 G.M.T, 25 August 1957—1000-millibar level*

| Verification item | Rate of change over 24 hours<br>2 × actual rate<br>$F=2A-P$ | $1\frac{1}{2}$ × actual rate<br>$F=\frac{1}{2}(3A-P)$<br>(Figs. 5(*a*) & (*b*)) | $\frac{1}{2}$ × actual rate<br>$F=\frac{1}{2}(A+P)$<br>(Figs. 5(*a*) & (*c*)) | Actual rate reversed<br>$F=2P-A$ |
|---|---|---|---|---|
| 3 r.m.s. height error (m) | 62 | 31 | 31 | 124 |
| 4 corr. coeff. for heights | 0·82 | 0·94 | 0·95 | 0·59 |
| 9 change corr. coeff. | 1·0 | 1·0 | 1·0 | —1·0 |
| 10 Bagrov scalar index | 0·66 | 0·86 | 0·86 | 0·47 |
| 13 r.m.s. error west wind (kt) | 13 | 6 | 6 | 26 |
| 14 r.m.s. error south wind (kt) | 20 | 10 | 10 | 39 |
| 15 r.m.s. error vector wind (kt) | 24 | 12 | 12 | 47 |
| 16 west wind corr. coeff. | 0·80 | 0·93 | 0·90 | 0·35 |
| 17 south wind corr. coeff. | 0·87 | 0·94 | 0·75 | —0·29 |
| 18 stretch vector corr. | 0·83 | 0·93 | 0·83 | —0·01 |
| 19 Wasko west index (kt) | 4·1 | 2·1 | 2·0 | 8·1 |
| 20 Wasko south index (kt) | 12·1 | 6·0 | 6·0 | 24·1 |
| 21 Teweles and Wobus index | 0·57 | 0·37 | 0·42 | 0·91 |
| 22 Bagrov vector index | 0·66 | 0·88 | 0·88 | 0·36 |

Winds are over one grid length.

TABLE XVI. *Effect of artificial rates of change, 15 × 10 area, 1800 G.M.T., 25 August 1957—500-millibar level*

| Verification item | Rate of change over 24 hours<br>2 × actual rate<br>$F=2A-P$<br>(Figs. 6(*a*) & (*b*)) | $1\frac{1}{2}$ × actual rate<br>$F=\frac{1}{2}(3A-P)$ | $\frac{1}{2}$ × actual rate<br>$F=\frac{1}{2}(A+P)$<br>(Figs. 6(*a*) & (*c*)) | Actual rate reversed<br>$F=2P-A$ |
|---|---|---|---|---|
| 3 r.m.s. height error (m) | 84 | 42 | 42 | 168 |
| 4 corr. coeff. for heights | 0·90 | 0·97 | 0·97 | 0·81 |
| 9 change corr. coeff. | 1·0 | 1·0 | 1·0 | —1·0 |
| 10 Bagrov scalar index | 0·70 | 0·88 | 0·88 | 0·66 |
| 13 r.m.s. error west wind (kt) | 17 | 9 | 9 | 35 |
| 14 r.m.s. error south wind (kt) | 23 | 11 | 11 | 46 |
| 15 r.m.s. error vector wind (kt) | 29 | 15 | 14 | 58 |
| 16 west wind corr. coeff. | 0·78 | 0·93 | 0·94 | 0·56 |
| 17 south wind corr. coeff. | 0·91 | 0·96 | 0·90 | —0·05 |
| 18 Stretch vector corr. | 0·86 | 0·95 | 0·91 | 0·23 |
| 19 Wasko west index (kt) | 8·5 | 4·2 | 4·2 | 17·1 |
| 20 Wasko south index (kt) | 16·3 | 8·2 | 8·2 | 32·7 |
| 21 Teweles and Wobus index | 0·44 | 0·26 | 0·28 | 0·75 |
| 22 Bagrov vector index | 0·79 | 0·92 | 0·92 | 0·50 |

Winds are over one grid length.

TABLE XVII. *Effect of artificial displacements, 15 × 10 area, 1800 G.M.T., 11 January 1957—1000-millibar level*

| Verification item | Displacement: 2 grid lengths west | 1 grid length west | 1 grid length east | 2 grid lengths east |
|---|---|---|---|---|
| 3 r.m.s. height error (m) | 61 | 34 | 37 | 71 |
| 4 corr. coeff. for heights | 0·80 | 0·94 | 0·95 | 0·85 |
| 9 change corr. coeff. | 0·62 | 0·89 | 0·91 | 0·73 |
| 10 Bagrov scalar index | 0·83 | 0·90 | 0·92 | 0·80 |
| 13 r.m.s. error west wind (kt) | 14 | 8 | 9 | 18 |
| 14 r.m.s. error south wind (kt) | 19 | 13 | 13 | 20 |
| 15 r.m.s. error vector wind (kt) | 23 | 15 | 16 | 26 |
| 16 west wind corr. coeff. | 0·80 | 0·93 | 0·92 | 0·76 |
| 17 south wind corr. coeff. | 0·54 | 0·79 | 0·81 | 0·54 |
| 18 stretch vector corr. | 0·69 | 0·87 | 0·87 | 0·67 |
| 19 Wasko west index (kt) | 5·0 | 2·8 | 3·4 | 7·3 |
| 20 Wasko south index (kt) | 12·4 | 7·7 | 7·1 | 11·7 |
| 21 Teweles and Wobus index | 0·56 | 0·36 | 0·36 | 0·56 |
| 22 Bagrov vector index | 0·79 | 0·89 | 0·89 | 0·77 |

Winds are over one grid length.

TABLE XVIII. *Effect of artificial displacements, 15 × 10 area, 1200 G.M.T., 18 April 1957—500-millibar level*

| Verification item | Displacement: 2 grid lengths east | 1 grid length east | 1 grid length west | 1 grid length north and west | 1 grid length north |
|---|---|---|---|---|---|
| 3 r.m.s. height error (m) | 66 | 34 | 34 | 88 | 73 |
| 4 corr. coeff. for heights | 0·96 | 0·99 | 0·99 | 0·97 | 0·98 |
| 9 change corr. coeff. | 0·68 | 0·84 | 0·69 | 0·73 | 0·72 |
| 10 Bagrov scalar index | 0·86 | 0·96 | 0·92 | 0·76 | 0·82 |
| 13 r.m.s. error west wind (kt) | 16 | 9 | 9 | 22 | 20 |
| 14 r.m.s. error south wind (kt) | 16 | 10 | 9 | 16 | 9 |
| 15 r.m.s. error vector wind (kt) | 23 | 13 | 13 | 27 | 22 |
| 16 west wind corr. coeff. | 0·82 | 0·95 | 0·95 | 0·69 | 0·73 |
| 17 south wind corr. coeff. | 0·53 | 0·82 | 0·84 | 0·65 | 0·89 |
| 18 stretch vector corr. | 0·73 | 0·91 | 0·92 | 0·68 | 0·78 |
| 19 Wasko west index (kt) | 4·2 | 2·2 | 2·4 | 13·3 | 12·8 |
| 20 Wasko south index (kt) | 7·2 | 4·5 | 3·8 | 7·9 | 4·0 |
| 21 Teweles and Wobus index | 0·41 | 0·26 | 0·25 | 0·45 | 0·37 |
| 22 Bagrov vector index | 0·92 | 0·97 | 0·93 | 0·90 | 0·93 |

Winds are over one grid length.

Artificial over-emphasis of cyclonic or anticyclonic developments was achieved by taking the forecast heights $F$ to be given by the arbitrary but convenient formula

$$F = A - \alpha\,\frac{H-L}{16}\left[1 - 4\frac{(H+L-2A)^2}{(H-L)^2}\right],$$

where $A$ is the actual contour height (at each point), $H$ the maximum of $A$ over the $15\times10$ area, $L$ the minimum of $A$ over the $15\times10$ area and $\alpha$ is an arbitrary constant.

When α is between 0 and 1 the effect of using the formula is to raise the contour heights in the central regions of the high- and low-pressure systems while leaving the mean contour height over the whole 15×10 area approximately unchanged. Thus the contour gradients near the high cells are tightened while those towards the centre of the depressions become weaker. Conversely values of α between 0 and —1 effect a lowering of the pressure and tightening of the pressure gradients in the low-pressure systems.

Some examples of the effect of this type of modification to the charts are shown in Tables XIX and XX and Figures 8 and 9.

TABLE XIX. *Effect of over-emphasis of cyclonic or anticyclonic gradients, 15×10 area, 0600 G.M.T., 4 November 1957—1000-millibar level*

| Verification item | $\alpha=\pm\frac{1}{4}$ | $\alpha=\pm\frac{1}{2}$ (Figure 8(*a*)) | $\alpha=\pm 1$ (Figure 8(*b*)) |
|---|---|---|---|
| 3 r.m.s. height error (m) | 7 | 14 | 30 |
| 4 corr. coeff. for heights | 1·0 | 0·99 | 0·97 |
| 9 change corr. coeff. | 0·99 | 0·97 | 0·86 |
| 10 Bagrov scalar index | 1·0 | 0·99 | 0·95 |
| 13 r.m.s. error west wind (kt) | 2 | 4 | 9 |
| 14 r.m.s. error south wind (kt) | 3 | 6 | 12 |
| 15 r.m.s. error vector wind (kt) | 4 | 7 | 15 |
| 16 west wind corr. coeff. | 1·0 | 0·98 | 0·95 |
| 17 south wind corr. coeff. | 1·0 | 0·98 | 0·92 |
| 18 stretch vector corr. | 0·99 | 0·98 | 0·93 |
| 19 Wasko west index (kt) | 1·3 | 2·5 | 5·7 |
| 20 Wasko south index (kt) | 1·9 | 3·9 | 7·9 |
| 21 Teweles and Wobus index | 0·09 | 0·17 | 0·32 |
| 22 Bagrov vector index | 0·99 | 0·97 | 0·90 |

Winds are over one grid length.

TABLE XX. *Effect of over-emphasis of cyclonic or anticyclonic gradients, 15×10 area, 1200 G.M.T., 28 May 1957—500-millibar level*

| Verification item | $\alpha=\pm\frac{1}{4}$ | $\alpha=\pm\frac{1}{2}$ | $\alpha=\pm 1$ (Figure 9) |
|---|---|---|---|
| 3 r.m.s. height error (m) | 6 | 13 | 25 |
| 4 corr. coeff. for heights | 1·0 | 0·99 | 0·98 |
| 9 change corr. coeff. | 0·99 | 0·95 | 0·82 |
| 10 Bagrov scalar index | 1·0 | 0·99 | 0·97 |
| 13 r.m.s. error west wind (kt) | 2 | 4 | 8 |
| 14 r.m.s. error south wind (kt) | 2 | 3 | 6 |
| 15 r.m.s. error vector wind (kt) | 2 | 5 | 10 |
| 16 west wind corr. coeff. | 0·99 | 0·98 | 0·93 |
| 17 south wind corr. coeff. | 1·0 | 0·98 | 0·94 |
| 18 stretch vector corr. | 0·99 | 0·98 | 0·93 |
| 19 Wasko west index (kt) | 0·6 | 1·3 | 2·6 |
| 20 Wasko south index (kt) | 0·7 | 1·3 | 2·6 |
| 21 Teweles and Wobus index | 0·09 | 0·18 | 0·30 |
| 22 Bagrov vector index | 1·0 | 1·0 | 0·99 |

Winds are over one grid length.

## COMPARISON OF INDICES AS TESTS OF FORECASTING SYSTEMS

During a further experiment in numerical forecasting, verification indices were computed for 50 pairs of 24-hour forecasts during the period 30 June to 12 November 1959. Each pair of forecasts comprised a numerical prediction and a Central Forecasting Office forecast. A comparison between the indices for the two forecasting systems is shown in Table XXI.

TABLE XXI. *Verification indices for numerical and C.F.O. 24-hour forecasts, 15 × 10 area—1000-millibar level*

| | Verification item | Numerical forecasts | C.F.O. forecasts |
|---|---|---|---|
| * 3 | r.m.s. height error (m) | 57 | 42 |
| † 4 | corr. coeff. for heights | 0·85 | 0·86 |
| † 9 | change corr. coeff. | 0·61 | 0·70 |
| †10 | Bagrov scalar index | 0·74 | 0·80 |
| *13 | r.m.s. error west wind (kt) | 12 | 11 |
| *14 | r.m.s. error south wind (kt) | 13 | 12 |
| *15 | r.m.s. error vector wind (kt) | 23 | 18 |
| †16 | west winds corr. coeff. | 0·54 | 0·60 |
| †17 | south wind corr. coeff. | 0·57 | 0·58 |
| †18 | stretch vector corr. | 0·58 | 0·60 |
| †19 | Wasko west index (kt) | 7 | 5 |
| †20 | Wasko south index (kt) | 7 | 6 |
| †21 | Teweles and Wobus index | 0·64 | 0·61 |
| †22 | Bagrov vector index | 0·74 | 0·81 |

* r.m.s. of indices computed for the 50 forecasts.
† Mean of indices computed for the 50 forecasts.
Winds are over one grid length.

On the basis of index comparisons the Central Forecasting Office forecasts are seen to be of a slightly better overall standard than that of the numerical forecasts, but the main point to note here is that all the listed indices except the Teweles and Wobus index suggest the same conclusion when being used as a guide to the comparative standard of the two forecasting systems.

## CONCLUSIONS

This experiment tended to emphasize rather than clarify the problem of assessing the comparative usefulness of the verification indices computed. Such comparisons have little meaning unless they are linked closely to some definite objective purpose. It might have been supposed that, for many practical purposes, a verification index should be highly correlated with subjective assessments of forecast charts, but most of the indices tested had little apparent relationship with subjective assessment and, in any case, the value of the subjective assessments themselves is open to doubt.

Nevertheless, some reasonably firm conclusions can be gleaned from the experiment.

(i) Most of the indices tested could show whether a forecast chart was very good or not, but the degrees of quality below a high level were not adequately revealed by the indices.

(ii) Verification indices designed to show the broad-scale quality of the type of fields used in this experiment should be computed over areas not less than the 15 × 10 full area shown in Figure 2. Forecast fields with, say, very high verification correlations or very low root mean square height errors over the 15×10 area are likely to be fairly accurate over the whole area, whereas if the verification indices are poor these indices will give no indication whether the forecast is poor generally or extremely poor in a small sub-area.

(iii) When discussing geostrophic wind errors it is most important to specify the grid length over which the winds are computed. Results of this experiment suggest that root mean square wind errors are approximately inversely proportional to the square root of the grid length.

(iv) The distribution of the root mean square height error (item 3), the Bagrov scalar index (item 10), the root mean square vector error (item 15) and the Wasko westerly index (item 19) suggest that these indices are more suitable for verification purposes than the remaining indices whose distributions tend to be leptokurtic.

(v) Only the root mean square height error (item 3) and the Bagrov scalar index (item 10) are moderately well correlated with subjective assessments of forecast charts.

It is suggested that, unless a completely new and more suitable verification index is discovered, the most practical, convenient and meaningful indices to use are the root mean square height error and root mean square vector wind error, computed over the 15×10 or slightly larger area with geostrophic winds measured over one grid length.

## Bibliography


1. WASKO, P. E.; Verification of prognostic maps prepared by means of the control-line method. *Bull. Amer. met. Soc., Lancaster, Pa.*, **35**, 1954, p. 257.
2. TEWELES, S. and WOBUS, H. B.; Verification of prognostic charts. *Bull. Amer. met. Soc., Lancaster, Pa.*, **35**, 1954, p. 455.
3. BAGROV, N. A.; Experience in the application of the analogue principle for forecasting average monthly air temperatures. *Tsent. Inst. Progn., Leningrad*, T. Vyp. 49, 1957, p. 231.

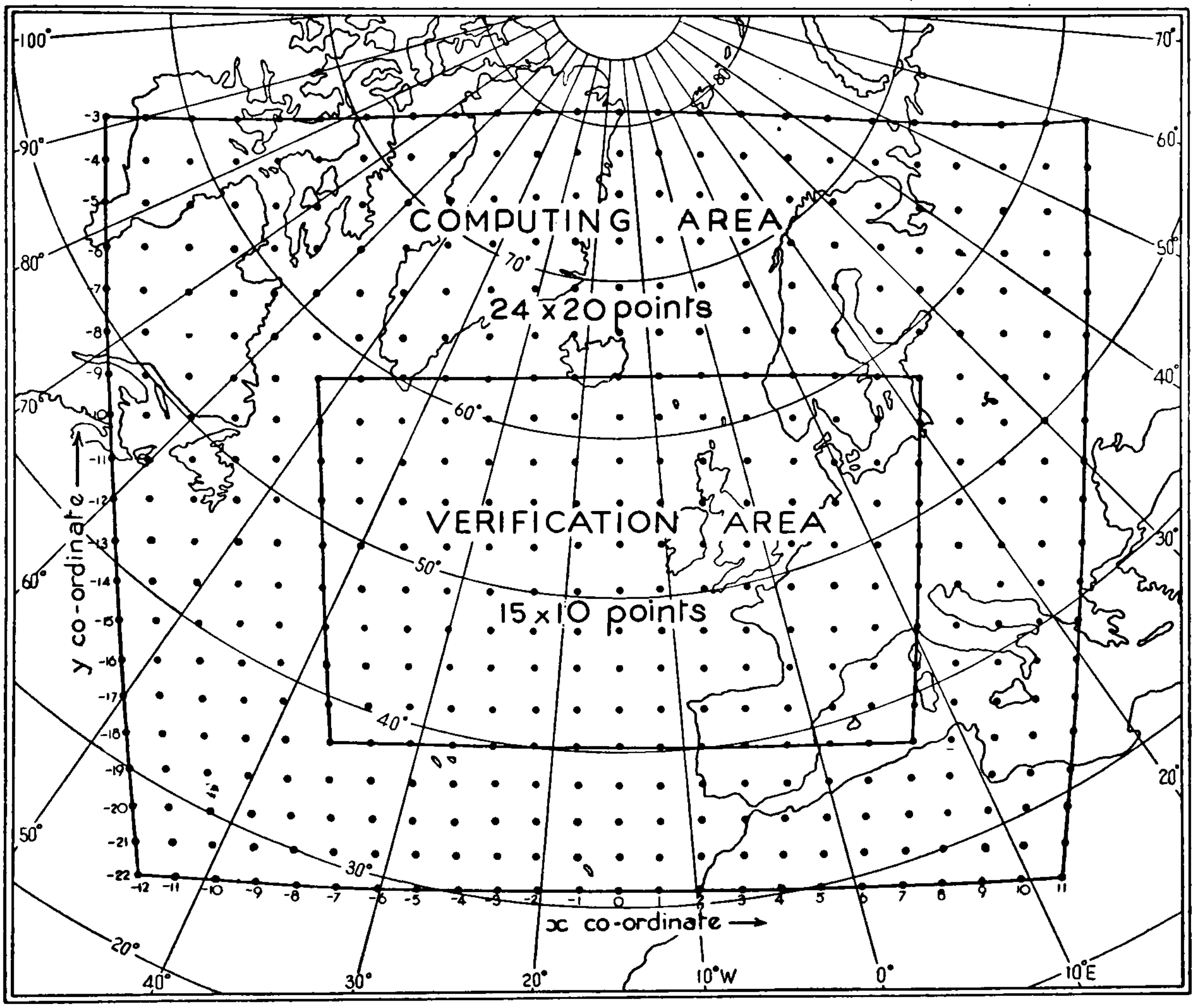


FIGURE 1. Computing grid for numerical forecasting experiment

The grid points form a square network on a stereographic projection from the South Pole on to a plane perpendicular to the earth's axis.

Grid length = 88·1412 (1 + sin $\phi$) n. miles where $\phi$ = latitude
= 176 n. miles (325 km) at North Pole
= 164 n. miles (303 km) at 60°N
= 132 n. miles (244 km) at 30°N

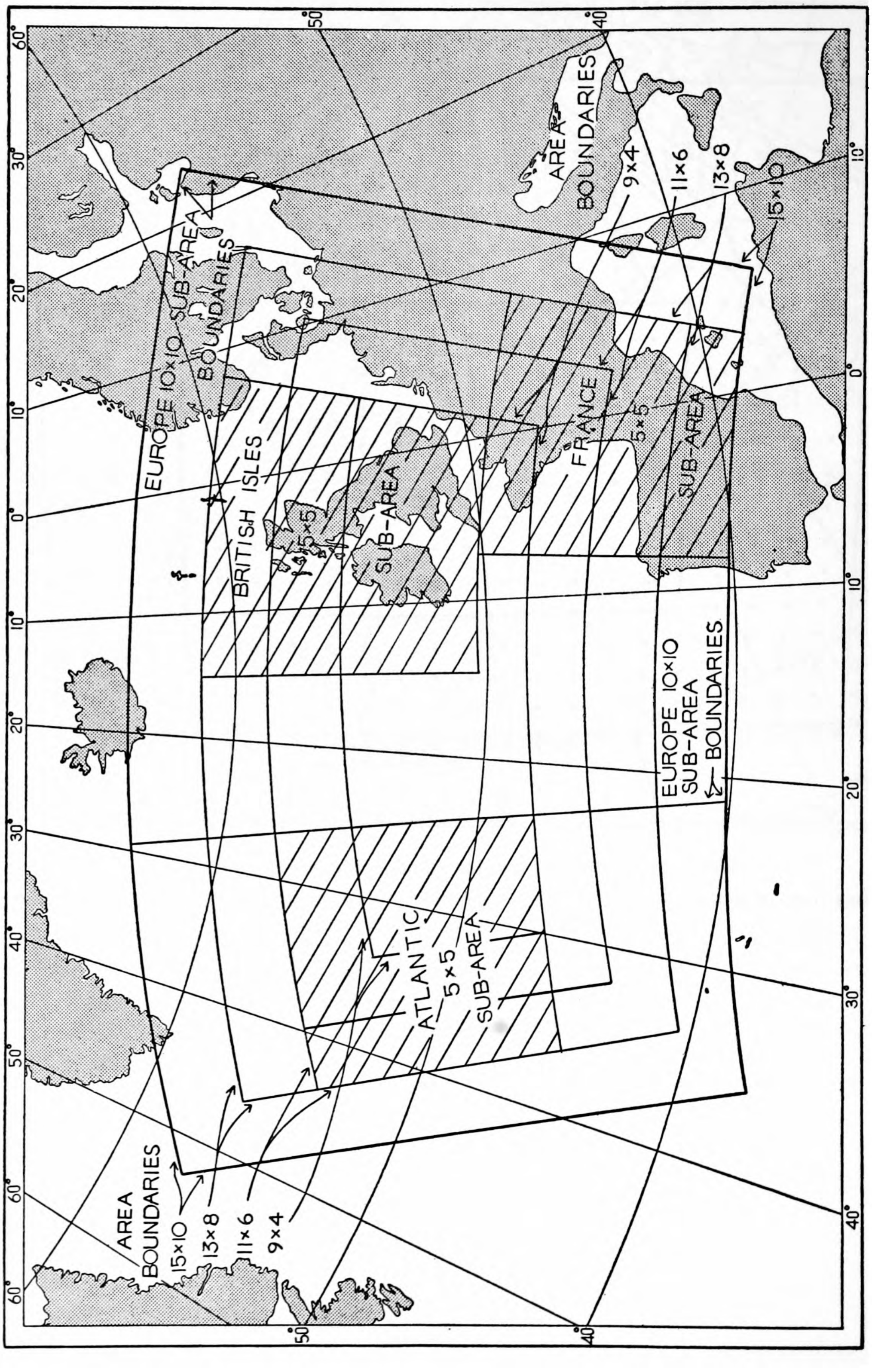


FIGURE 2. Sub-areas of the full size verification area

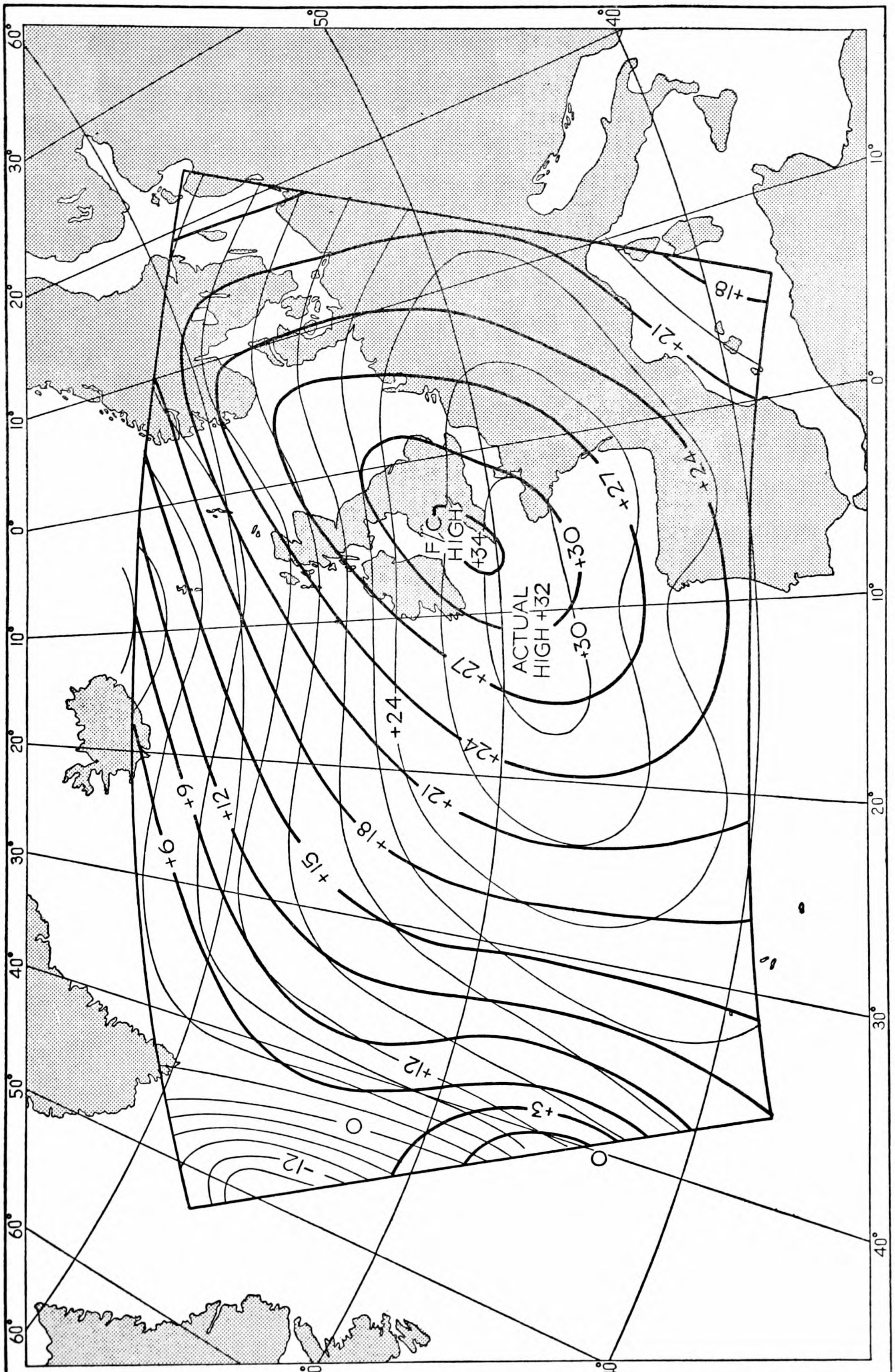


FIGURE 3(*a*). 1000-mb contours, in decametres, 1800 G.M.T., 11 January 1957
Thick and thin lines represent forecast and actual contours respectively.

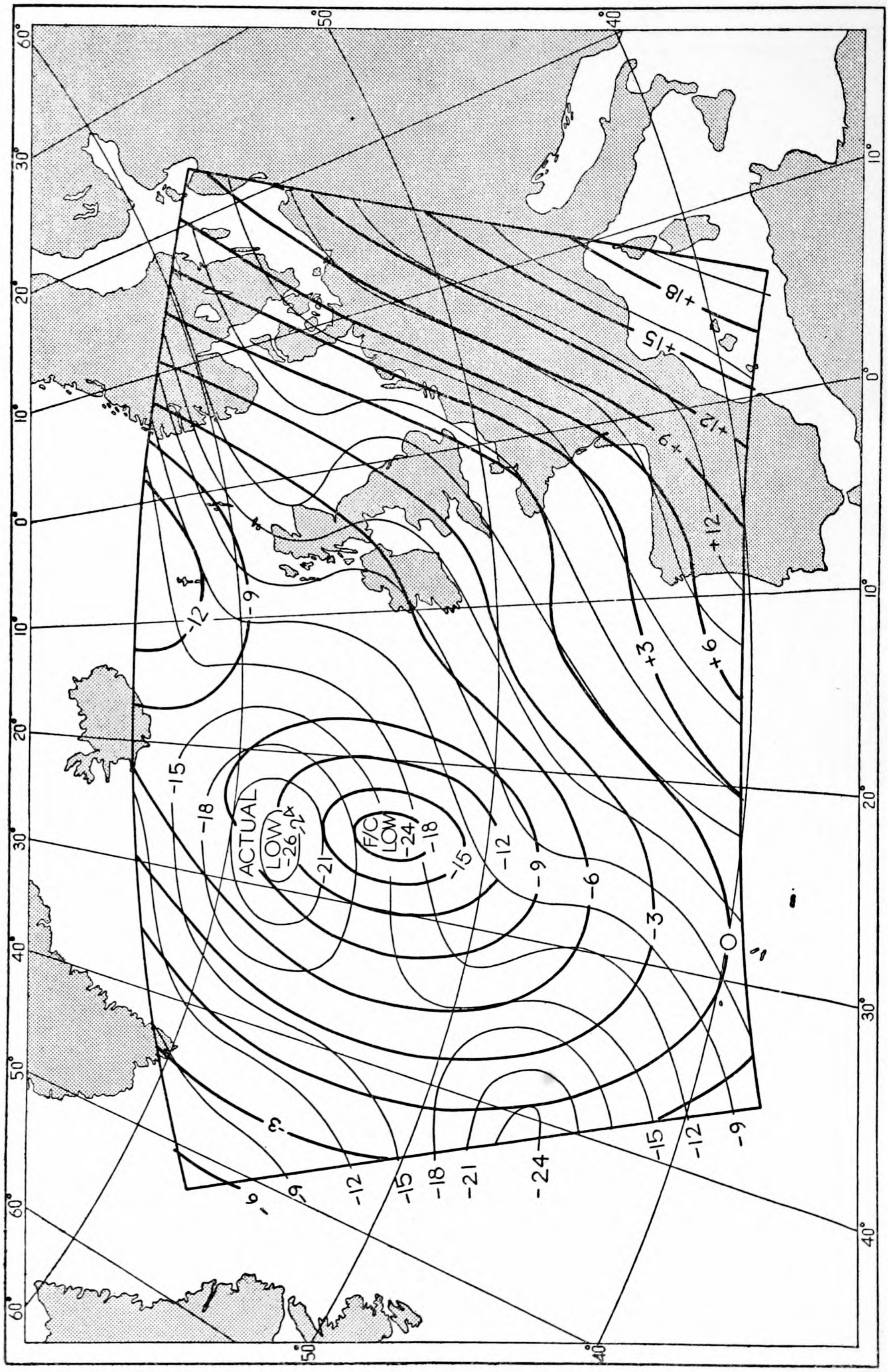


FIGURE 3(*b*). 1000-mb contours, in decametres, 0001 G.M.T., 6 February 1957
Thick and thin lines represent forecast and actual contours respectively.

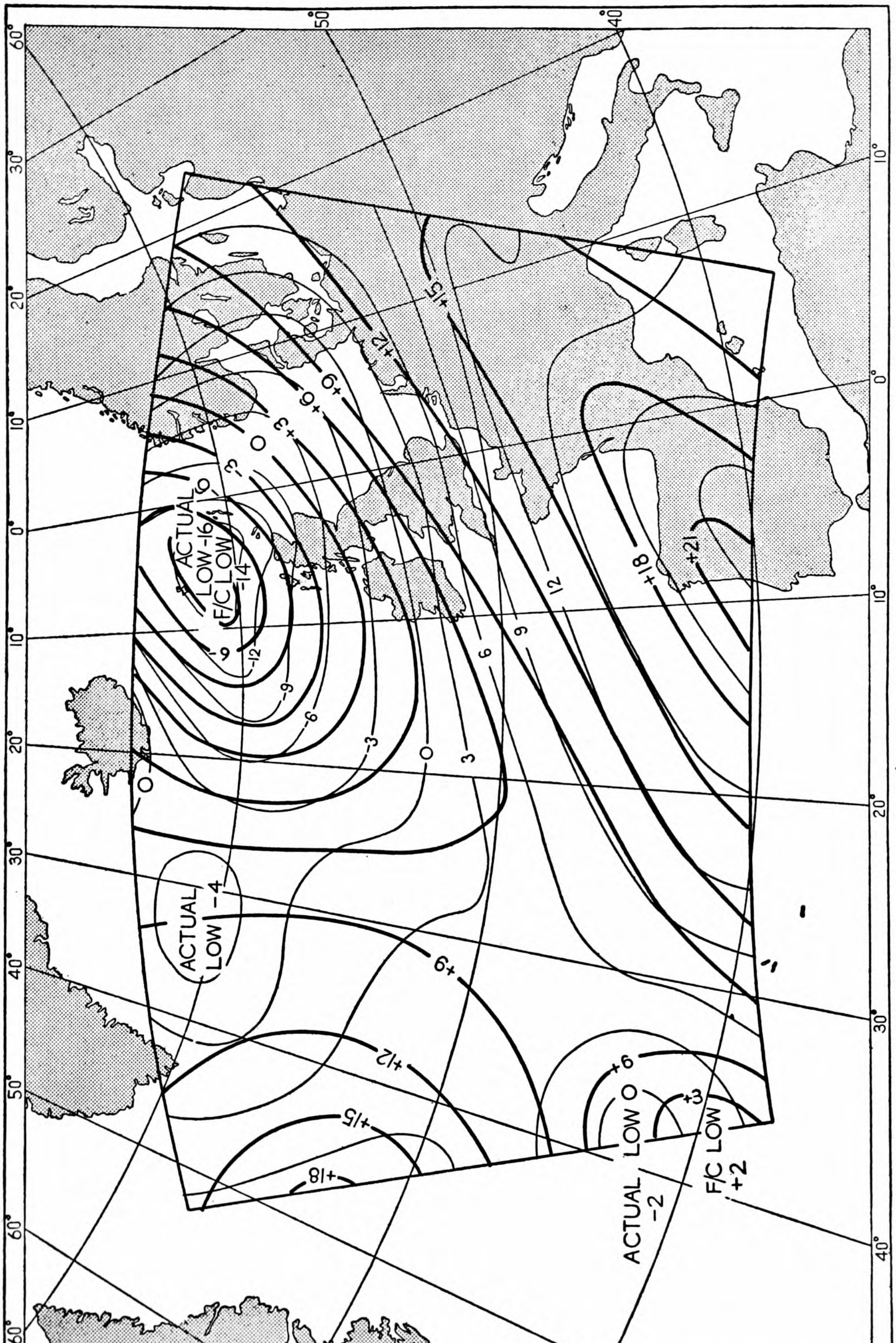


FIGURE 3(c). 1000-mb contours, in decametres, 1200 G.M.T., 17 March 1957
Thick and thin lines represent forecast and actual contours respectively.

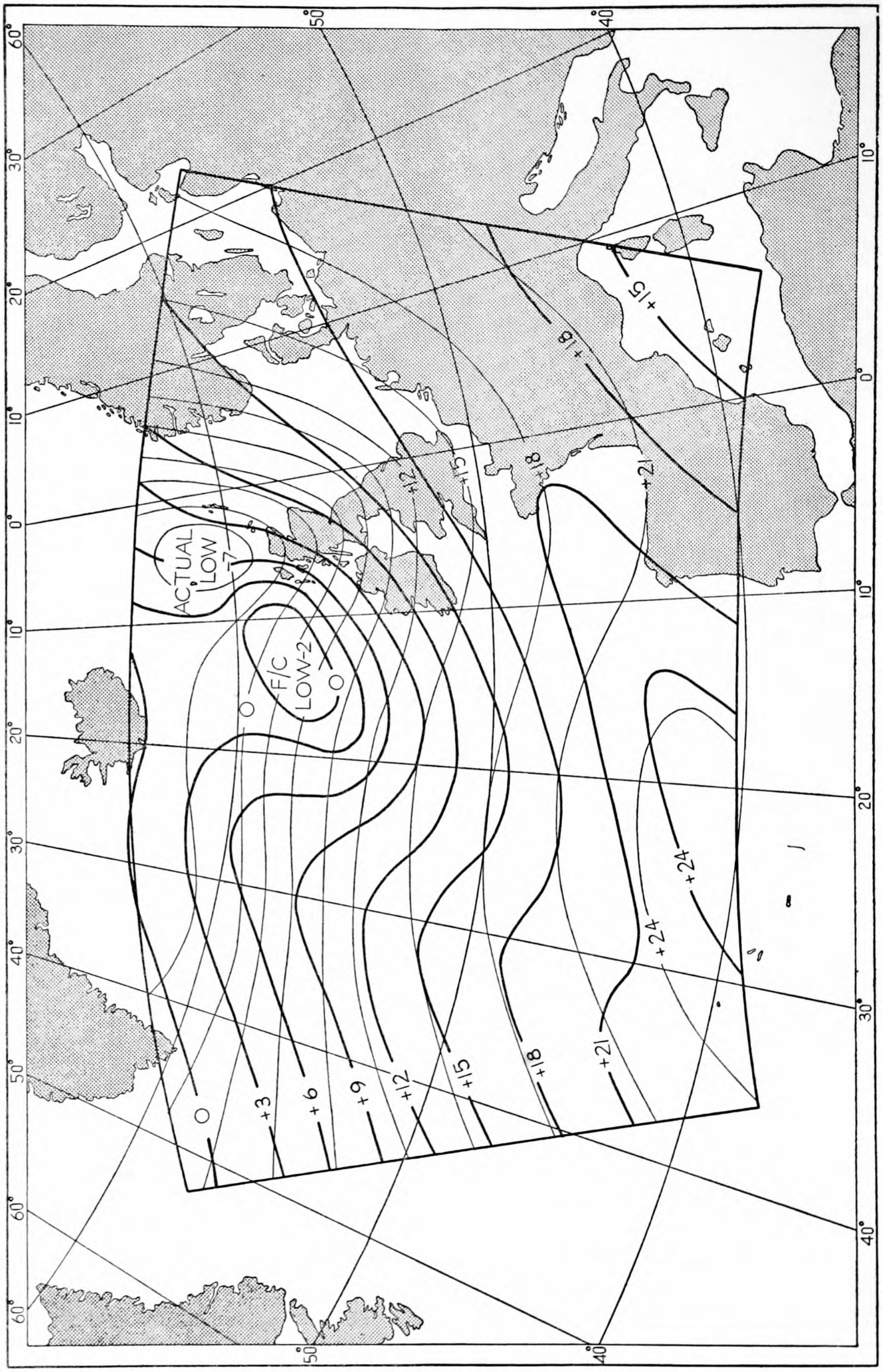


FIGURE 3(*d*). 1000-mb contours, in decametres, 1200 G.M.T., 18 April 1957
Thick and thin lines represent forecast and actual contours respectively.

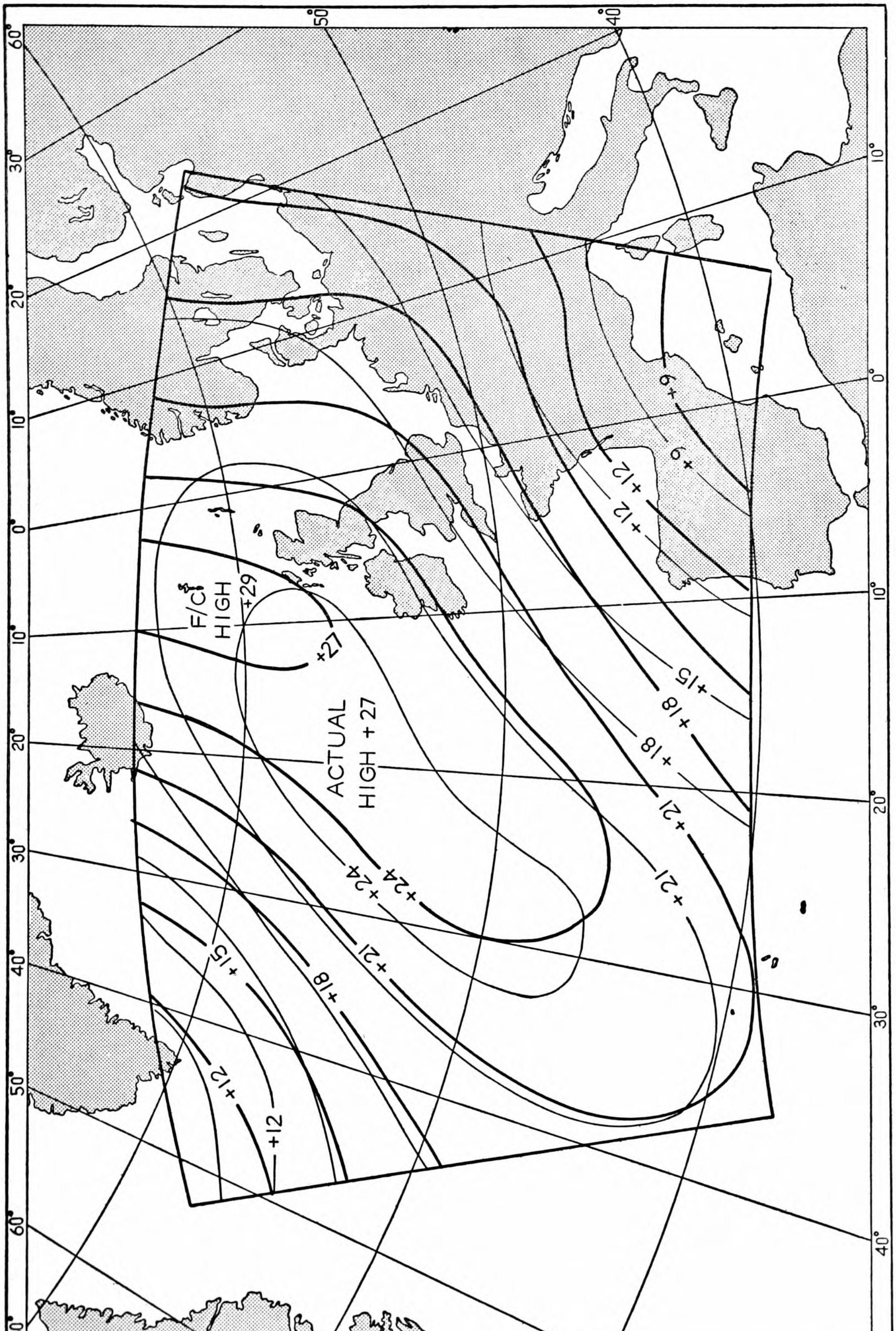


FIGURE 3(*e*). 1000-mb contours, in decametres, 1200 G.M.T., 28 May 1957
Thick and thin lines represent forecast and actual contours respectively.

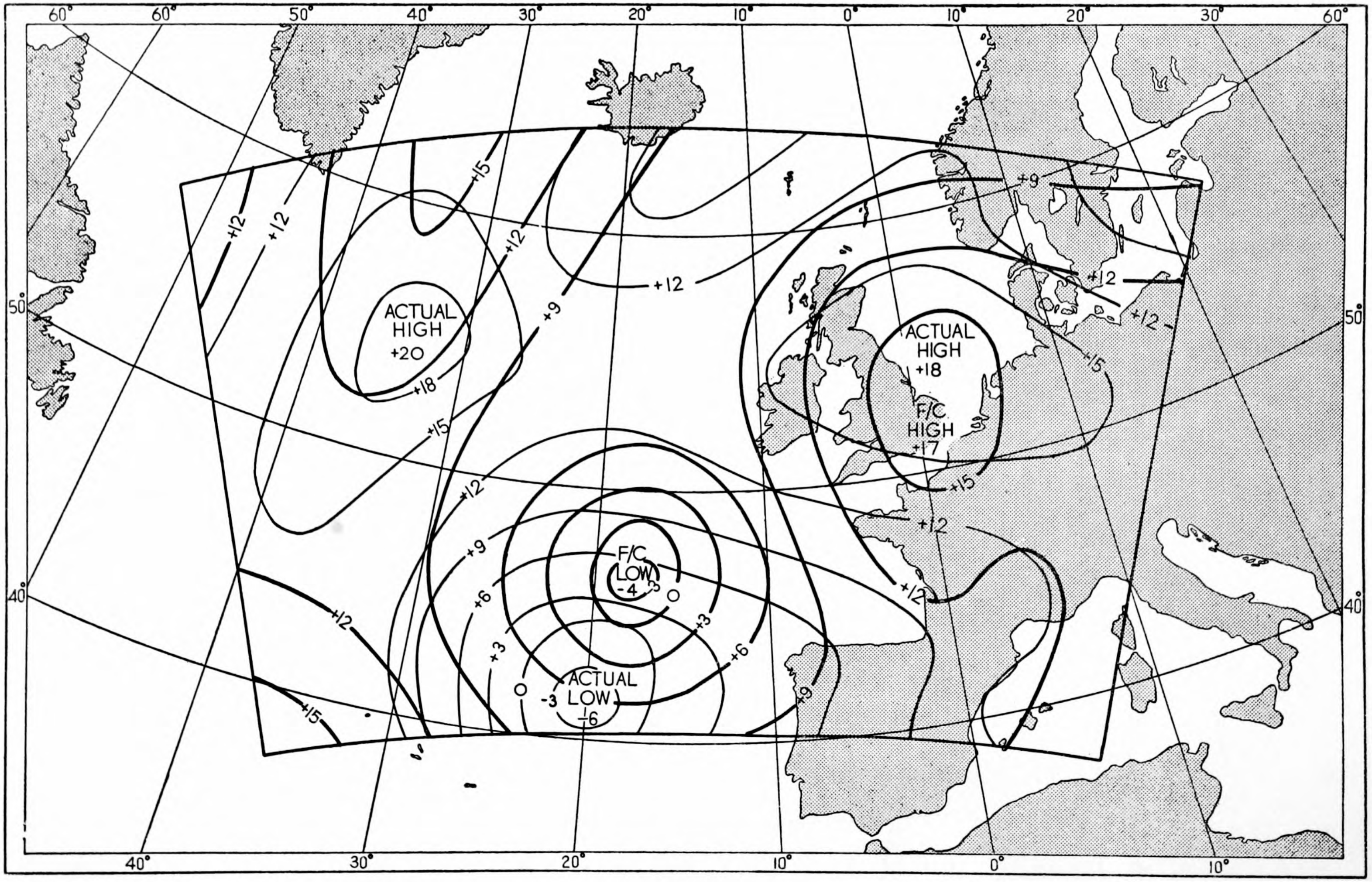


FIGURE 3(*f*). 1000-mb contours, in decametres, 1200 G.M.T., 2 July 1957
Thick and thin lines represent forecast and actual contours respectively.

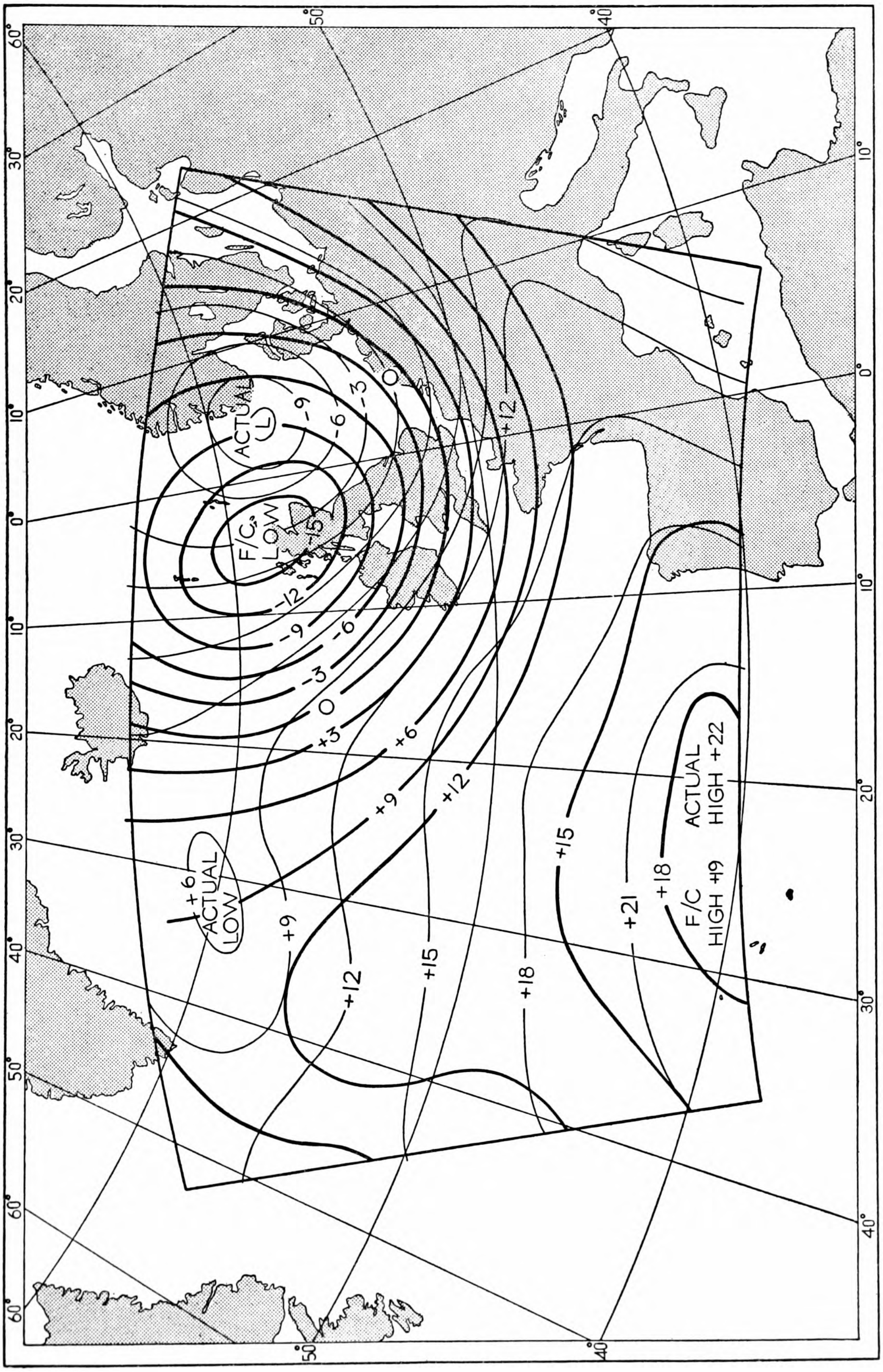


FIGURE 3(*g*). 1000-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thick and thin lines represent forecast and actual contours respectively.

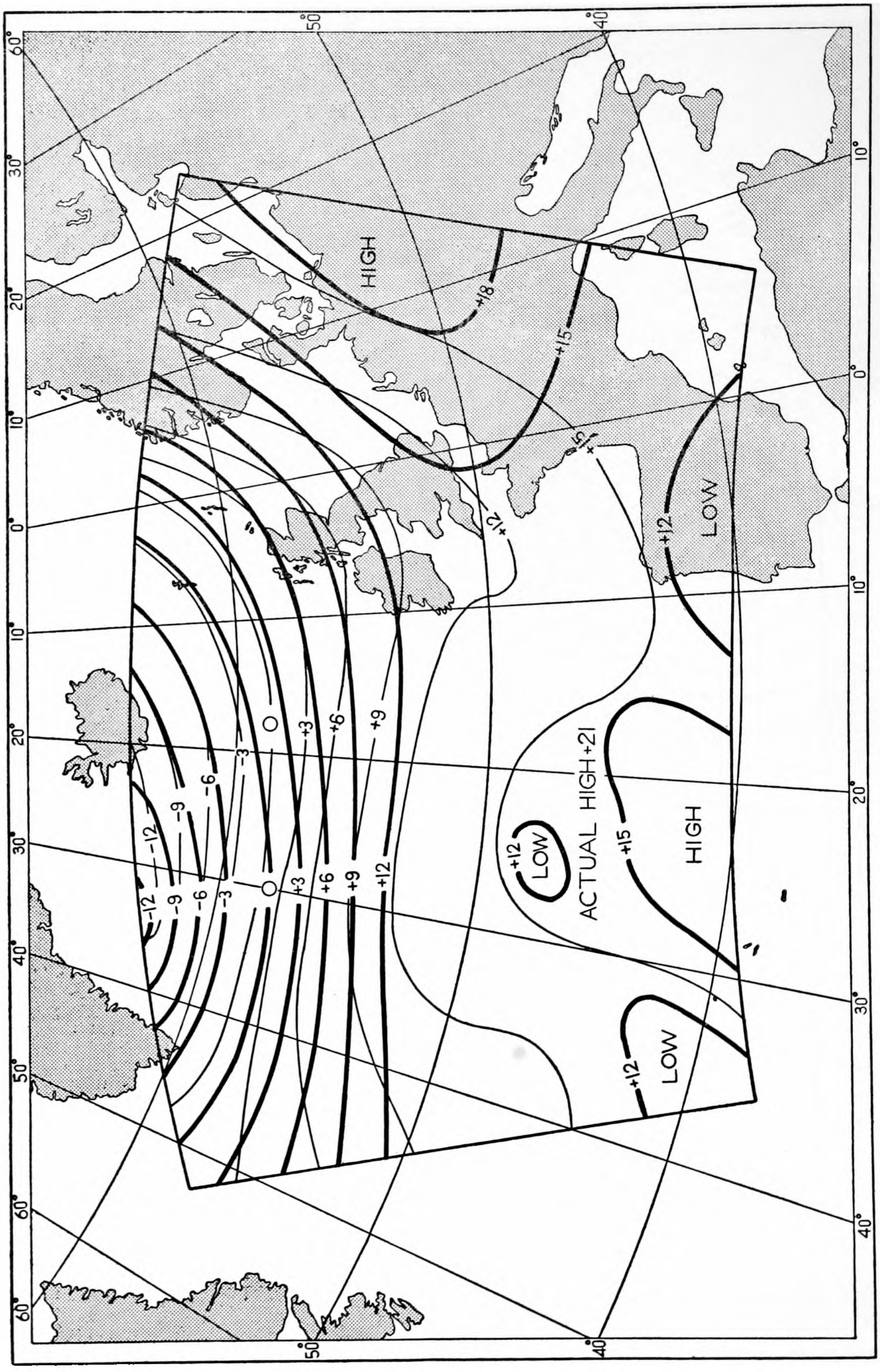


FIGURE 3(*h*). 1000-mb contours, in decametres, 0600 G.M.T., 16 October 1957
Thick and thin lines represent forecast and actual contours respectively.

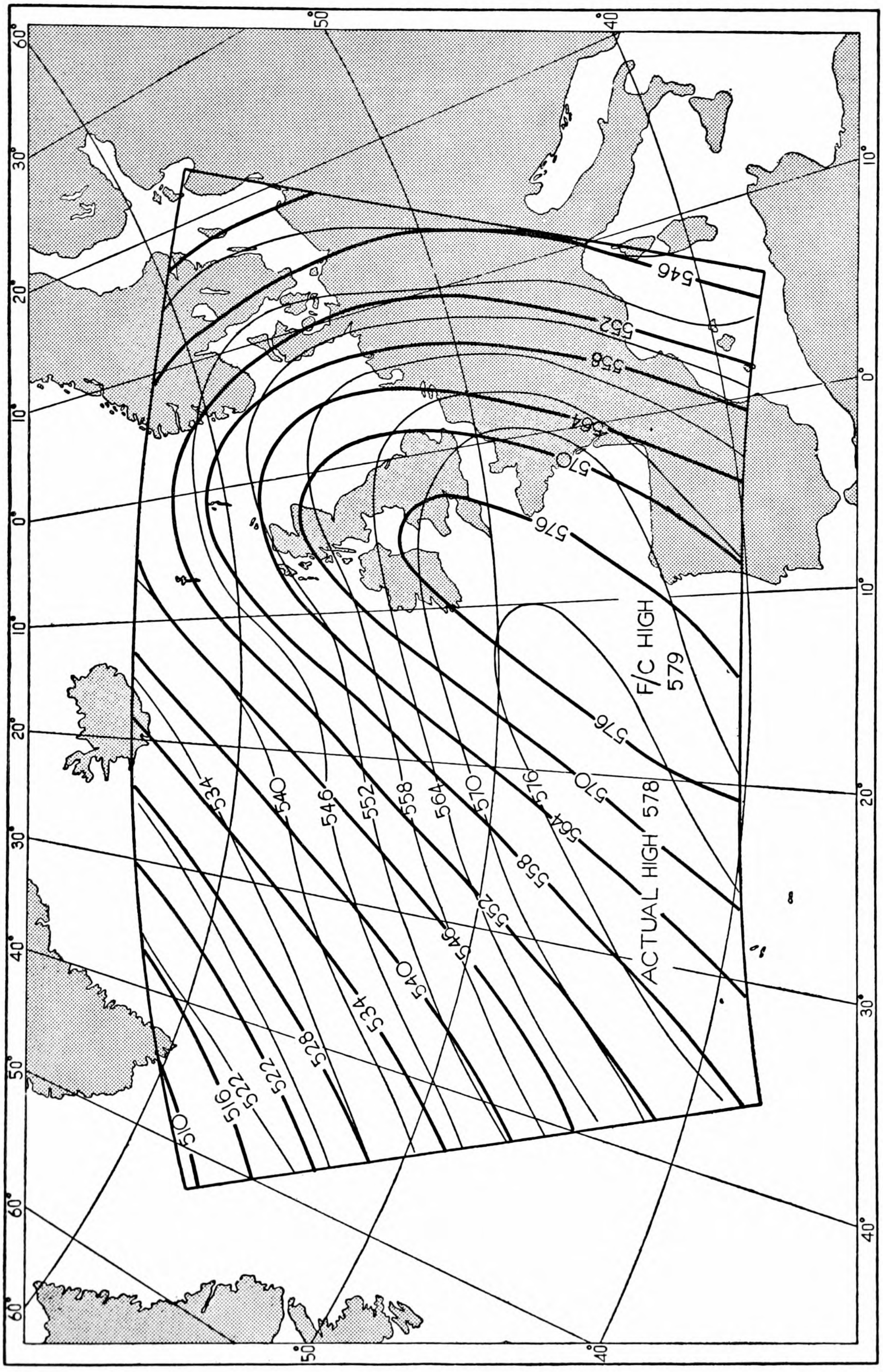


FIGURE 4(*a*). 500-mb contours in decametres, 1800 G.M.T., 11 January 1957
Thick and thin lines represent forecast and actual contours respectively.

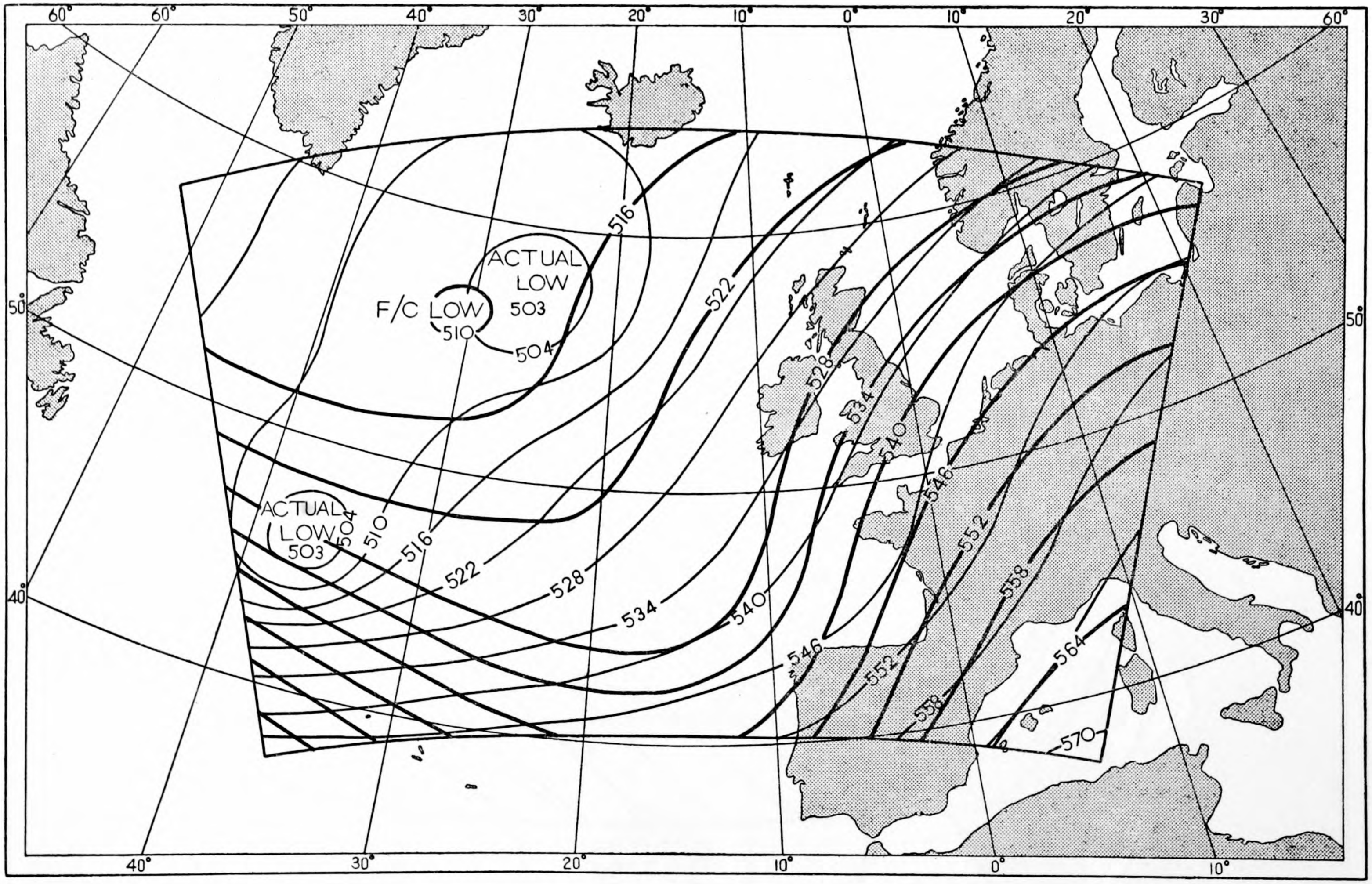


FIGURE 4(*b*). 500-mb contours, in decametres, 0001 G.M.T., 6 February 1957
Thick and thin lines represent forecast and actual contours respectively.

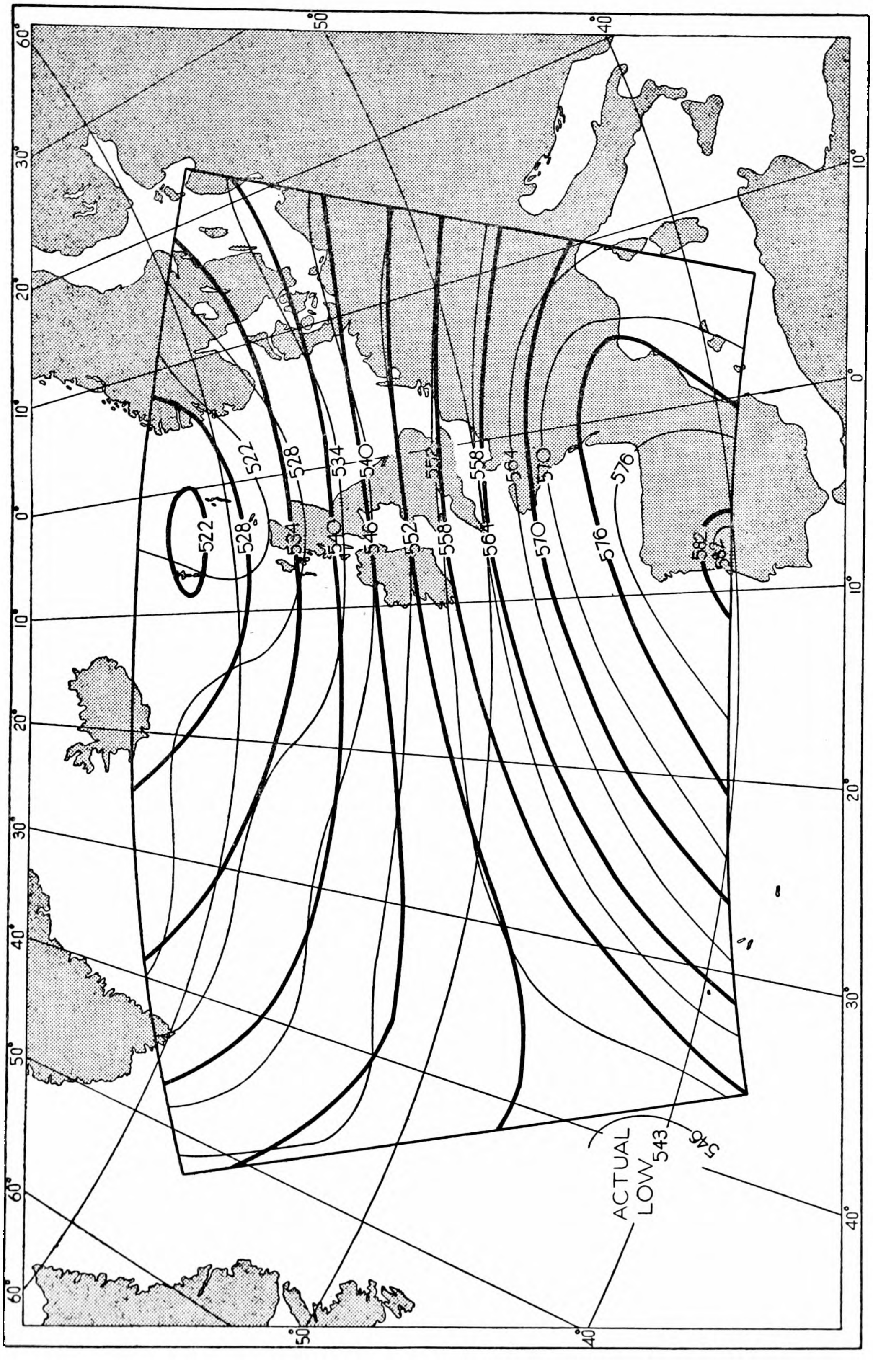


FIGURE 4(*c*). 500-mb contours, in decametres, 1200 G.M.T., 17 March 1957
Thick and thin lines represent forecast and actual contours respectively.

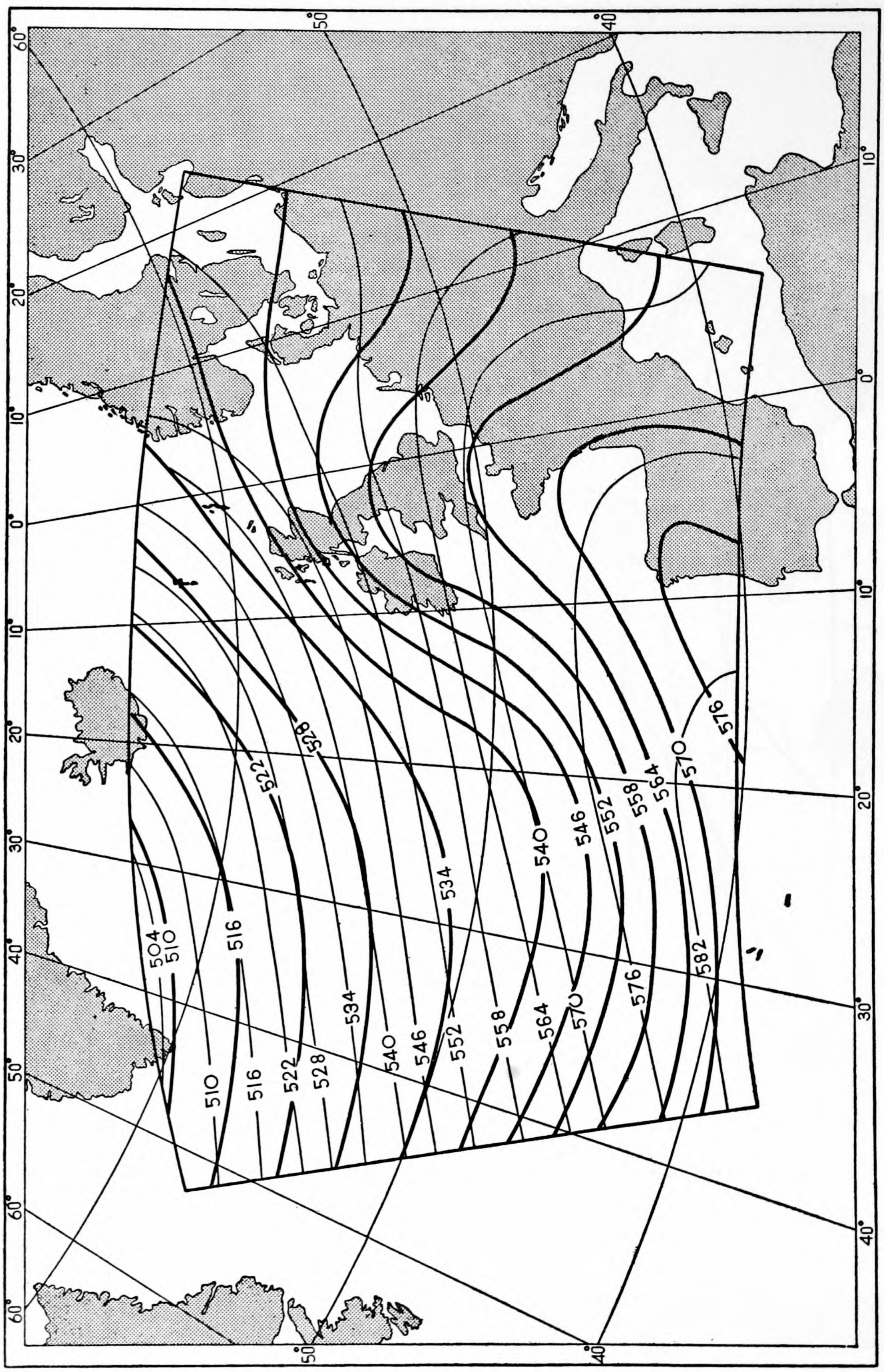


FIGURE 4(*d*). 500-mb contours, in decametres, 1200 G.M.T., 18 April 1957
Thick and thin lines represent forecast and actual contours respectively.

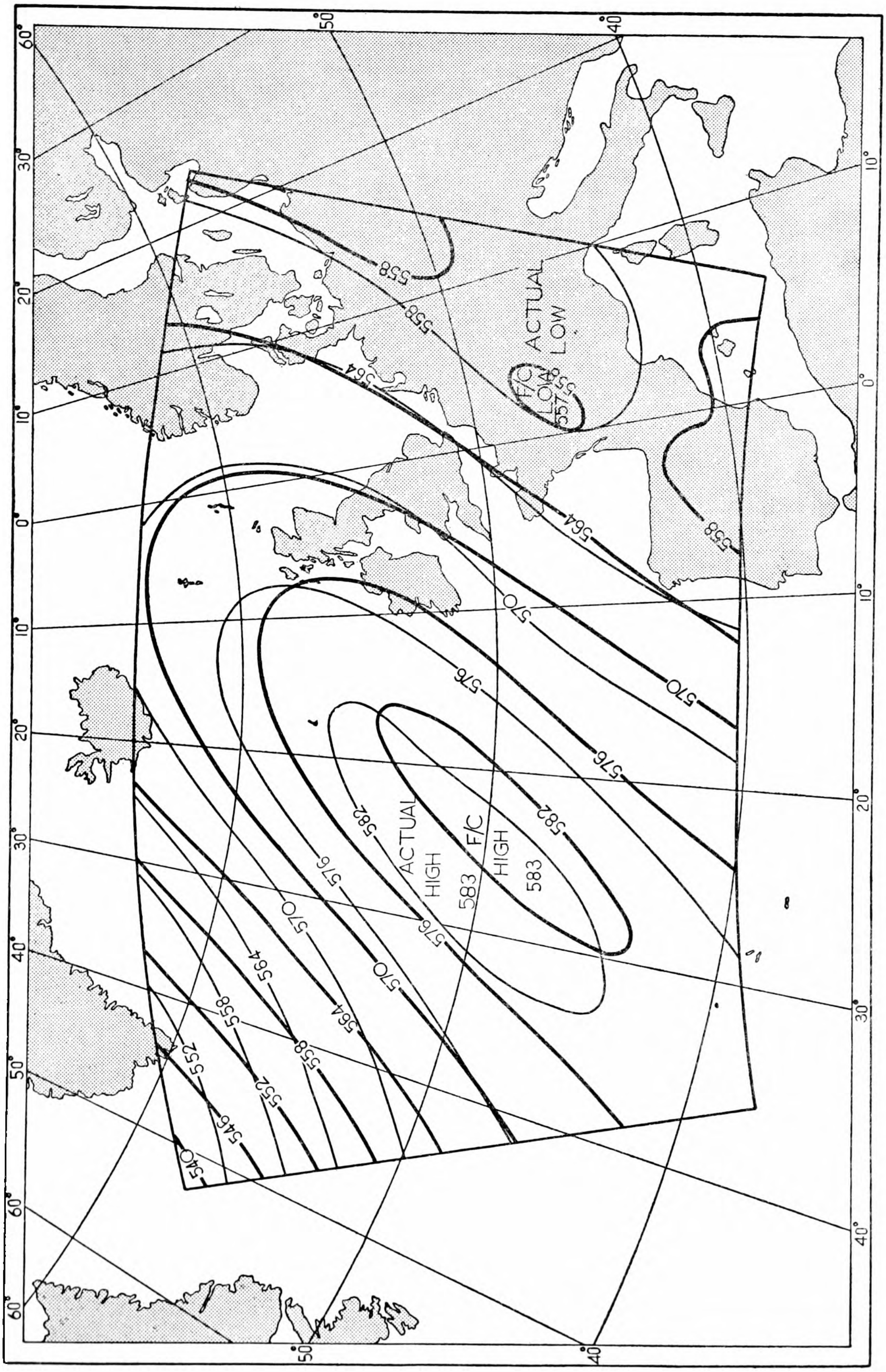


FIGURE 4(*e*). 500-mb contours, in decametres, 1200 G.M.T., 28 May 1957
Thick and thin lines represent forecast and actual contours respectively.

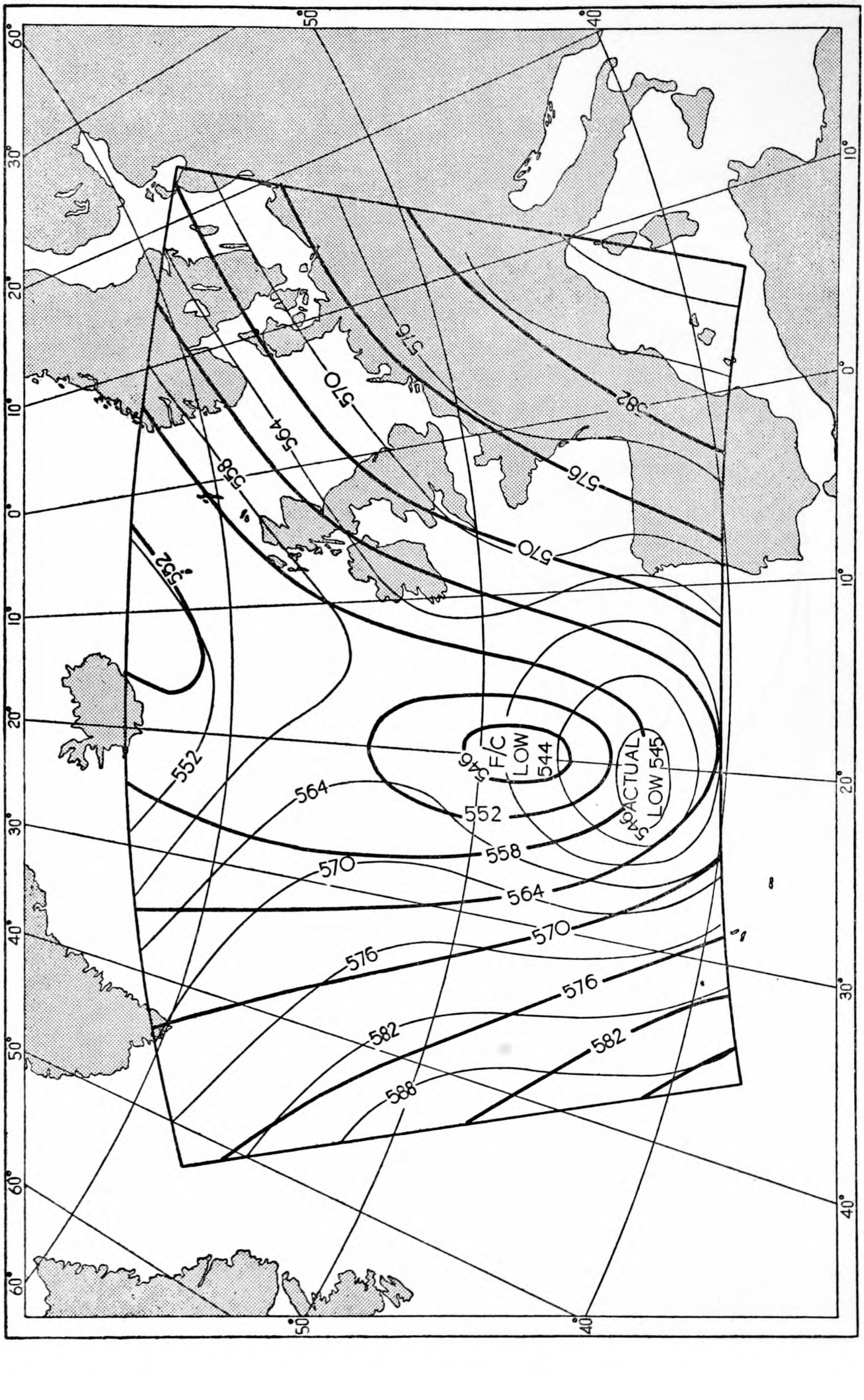


FIGURE 4(*f*). 500-mb contours, in decametres, 1200 G.M.T., 2 July 1957
Thick and thin lines represent forecast and actual contours respectively.

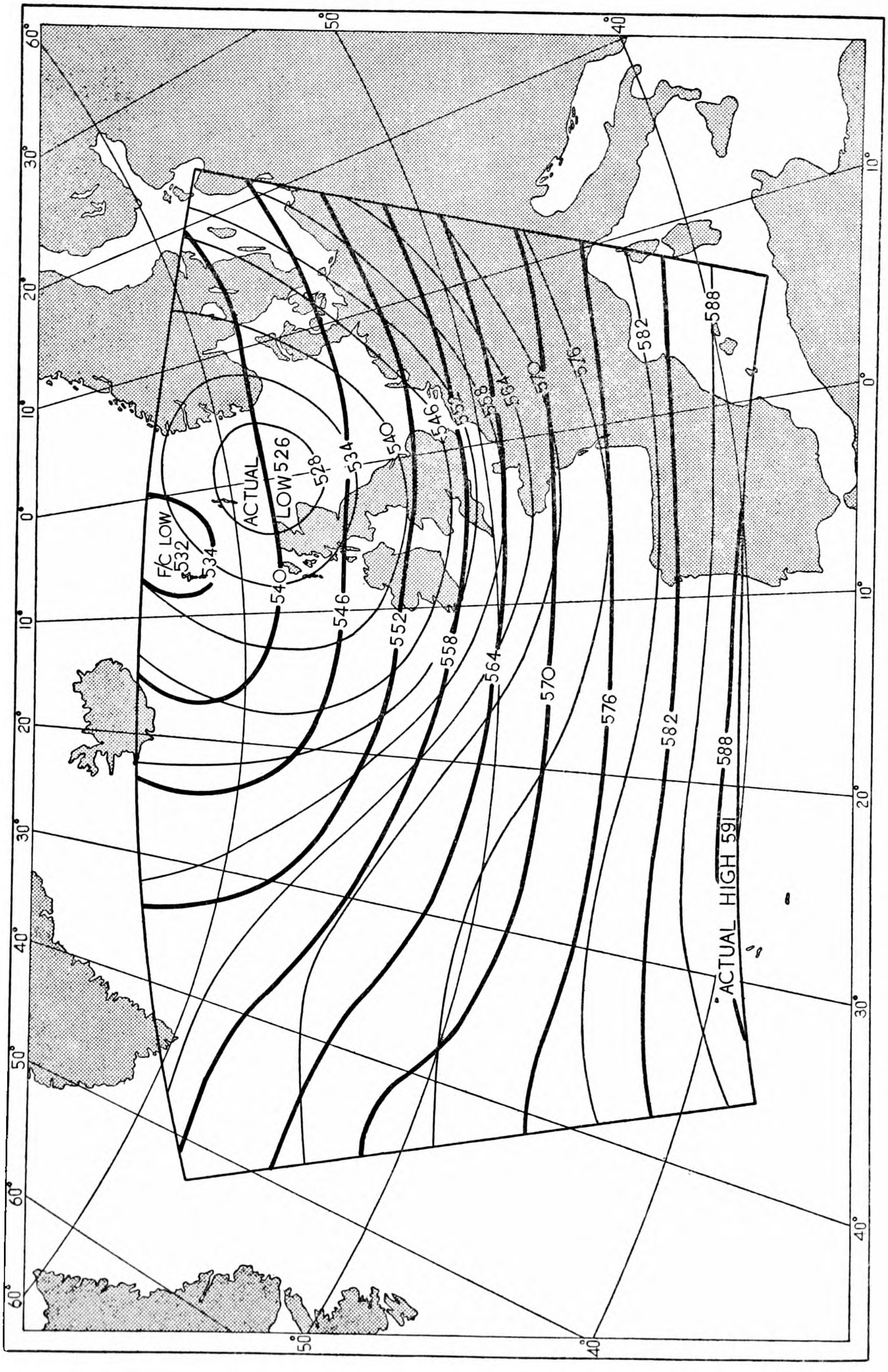


FIGURE 4(*g*). 500-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thick and thin lines represent forecast and actual contours respectively.

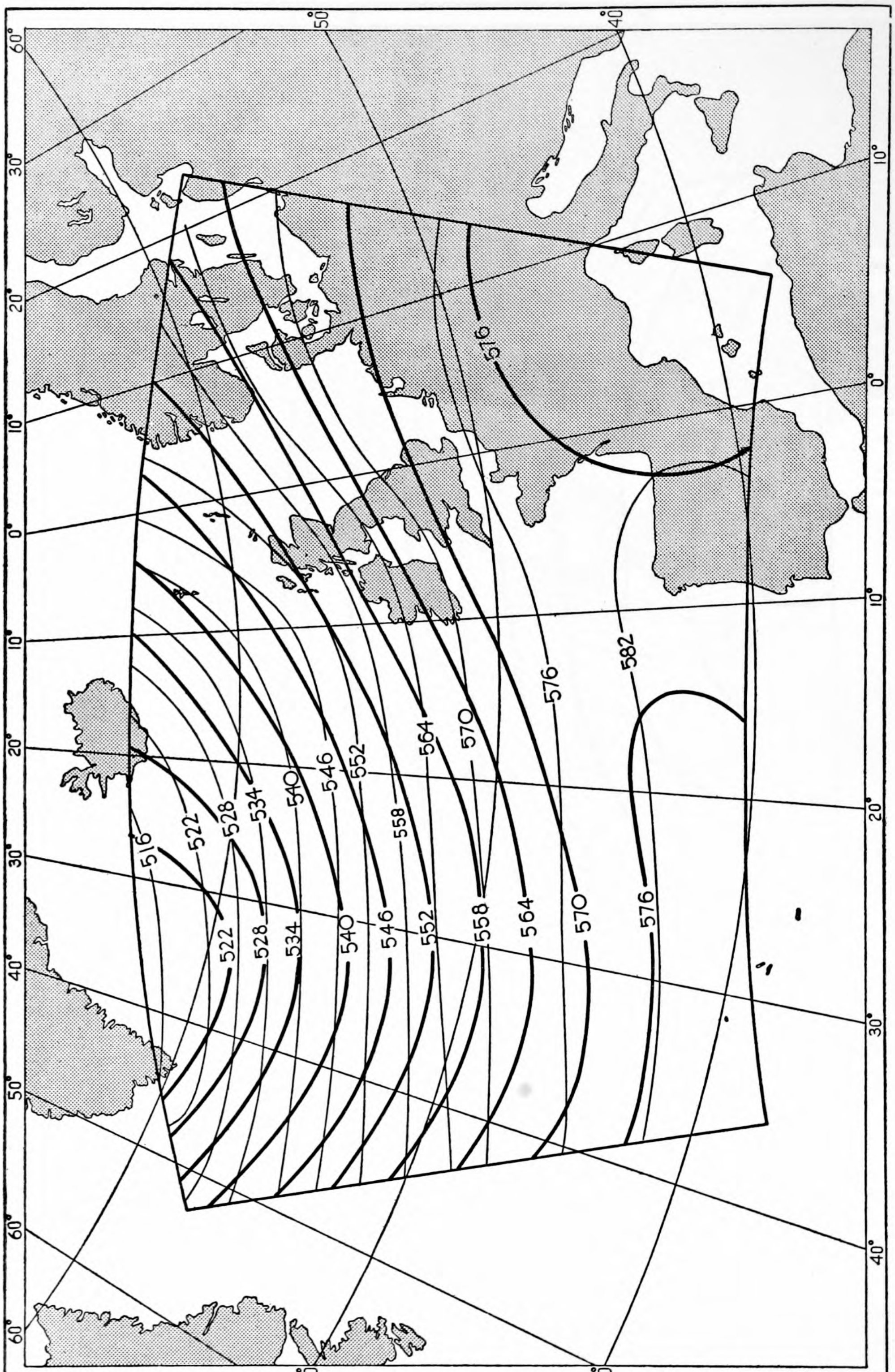


FIGURE 4(*h*). 500-mb contours, in decametres, 0600 G.M.T., 16 October 1957
Thick and thin lines represent forecast and actual contours respectively.

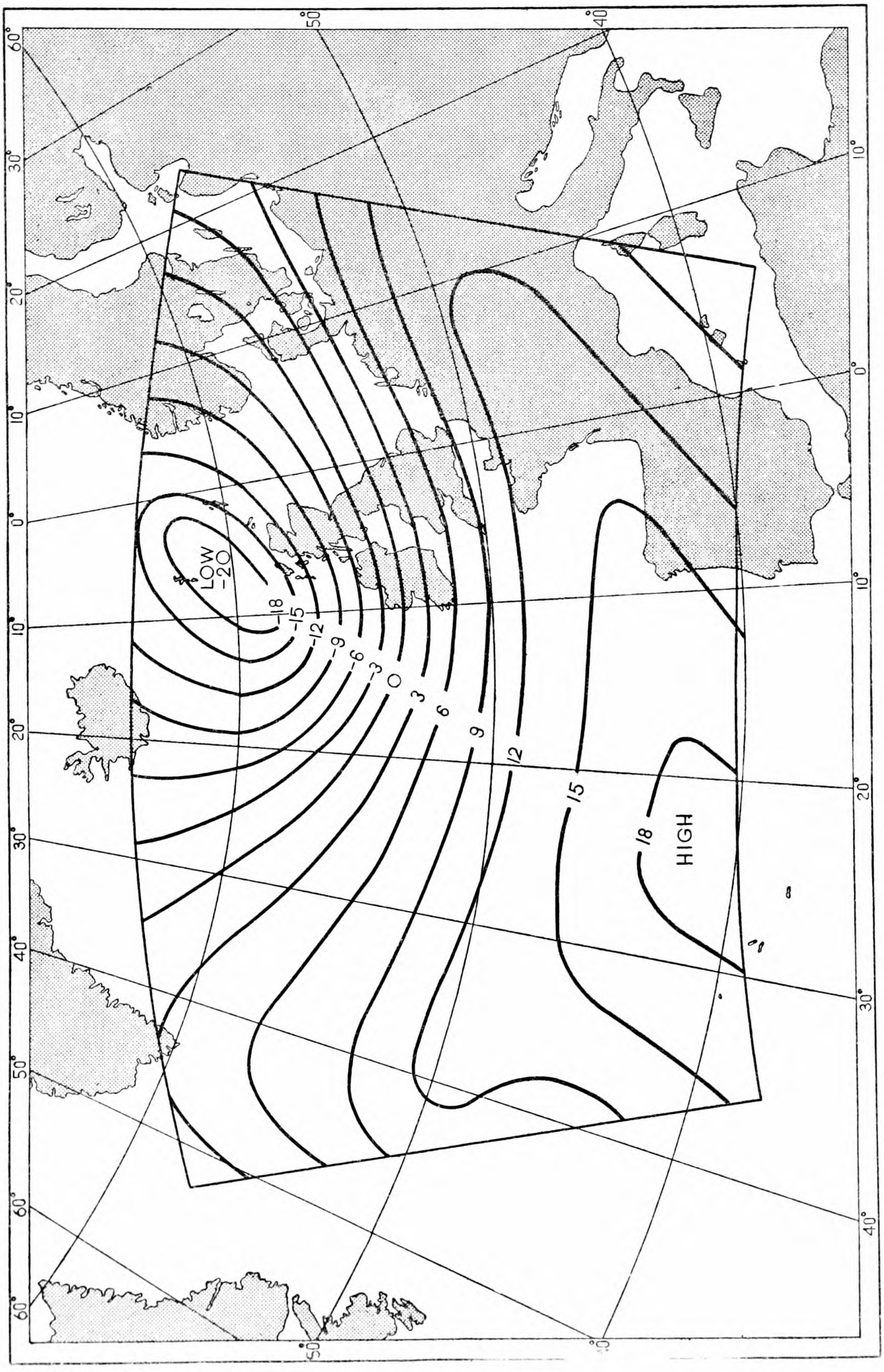


FIGURE 5(*a*). Actual 1000-mb contours, in decametres, 1800 G.M.T., 24 August 1957

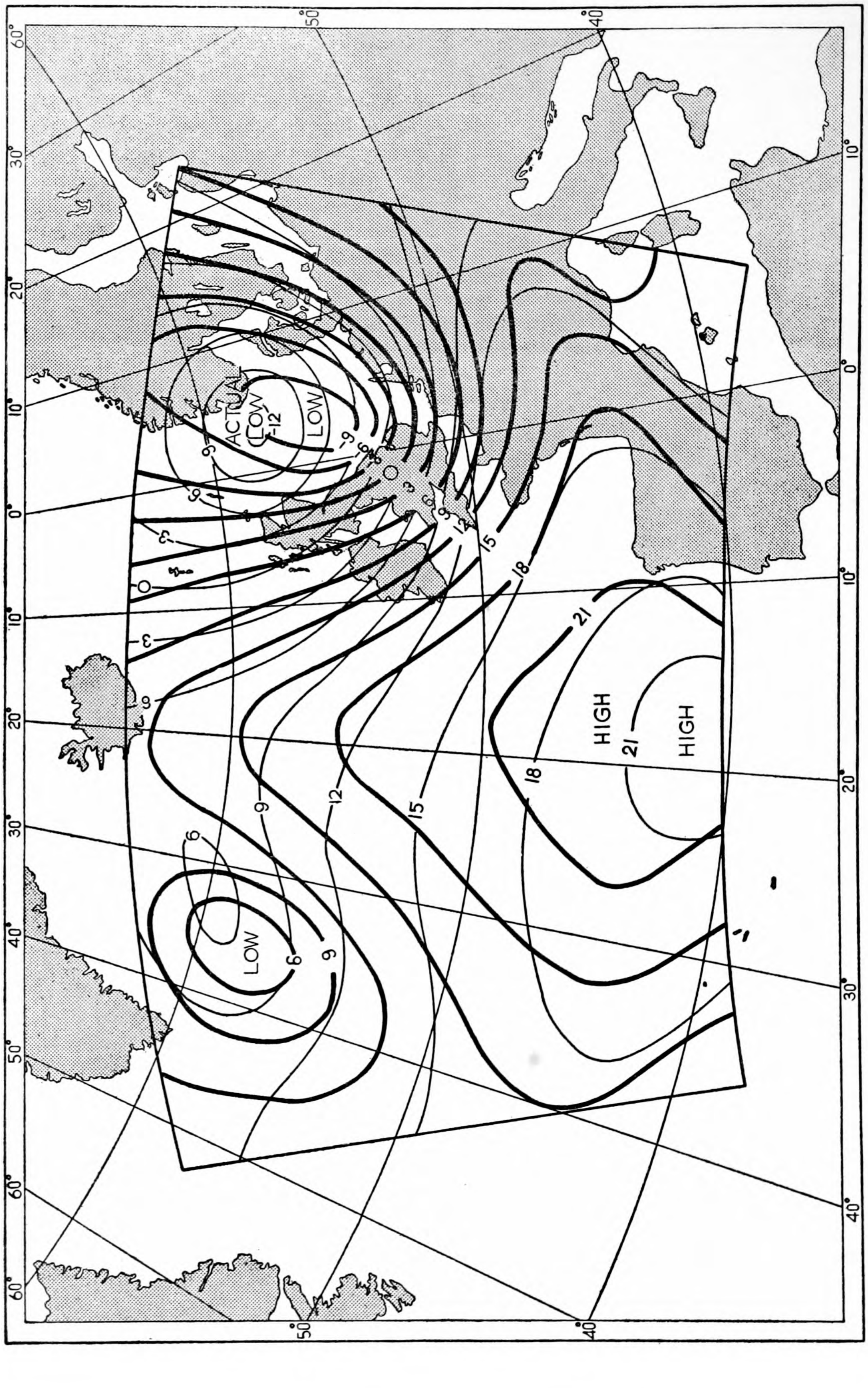


FIGURE 5(*b*). 1000-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thin lines represent actual contours and thick lines represent contours obtained by using $1\frac{1}{2}$ times the actual changes during the previous 24 hours.

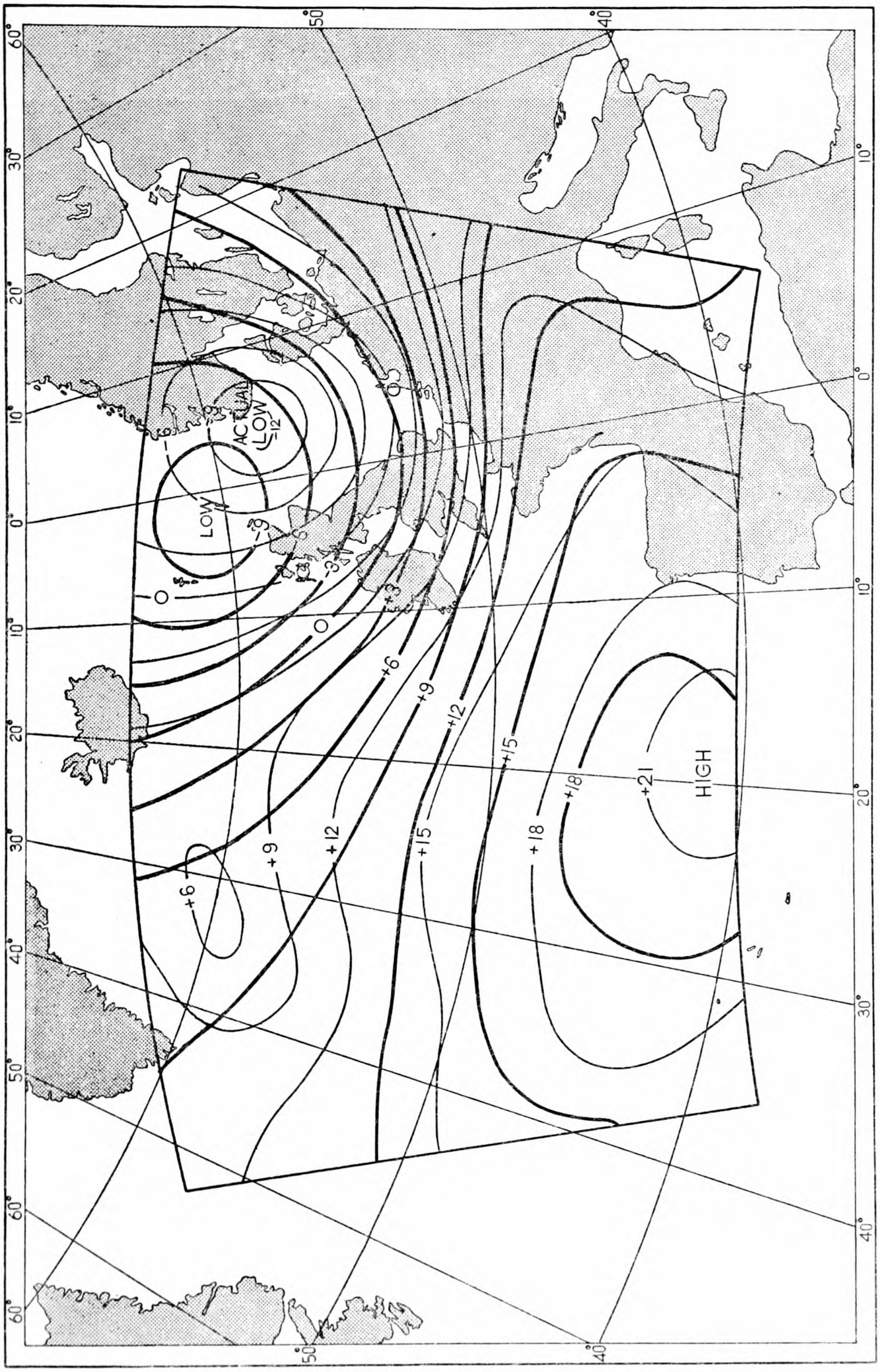


FIGURE 5(*c*). 1000-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thin lines represent actual contours and thick lines represent contours obtained by using half the actual changes during the previous 24 hours.

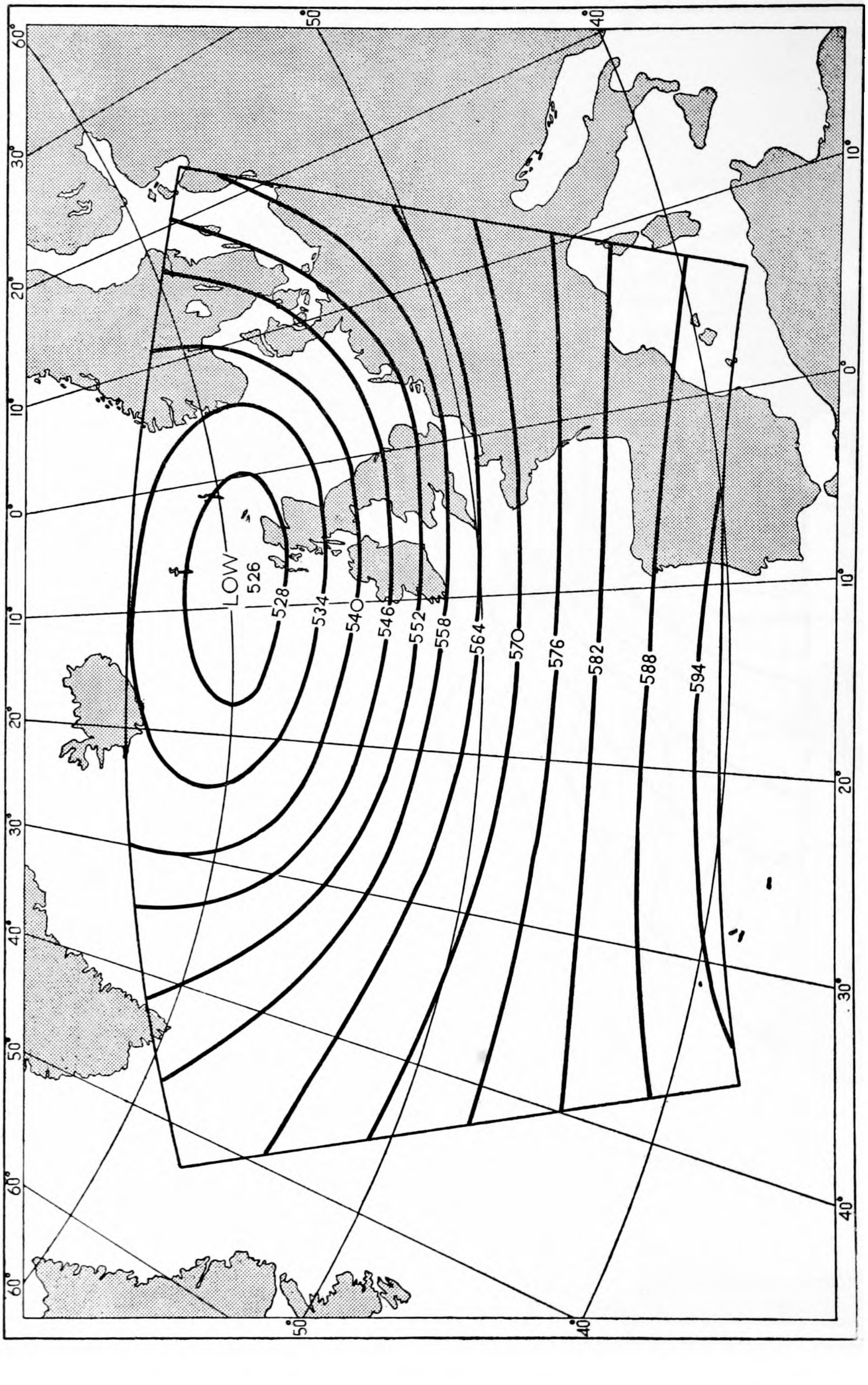


FIGURE 6(*a*). Actual 500-mb contours, in decametres, 1800 G.M.T., 24 August 1957

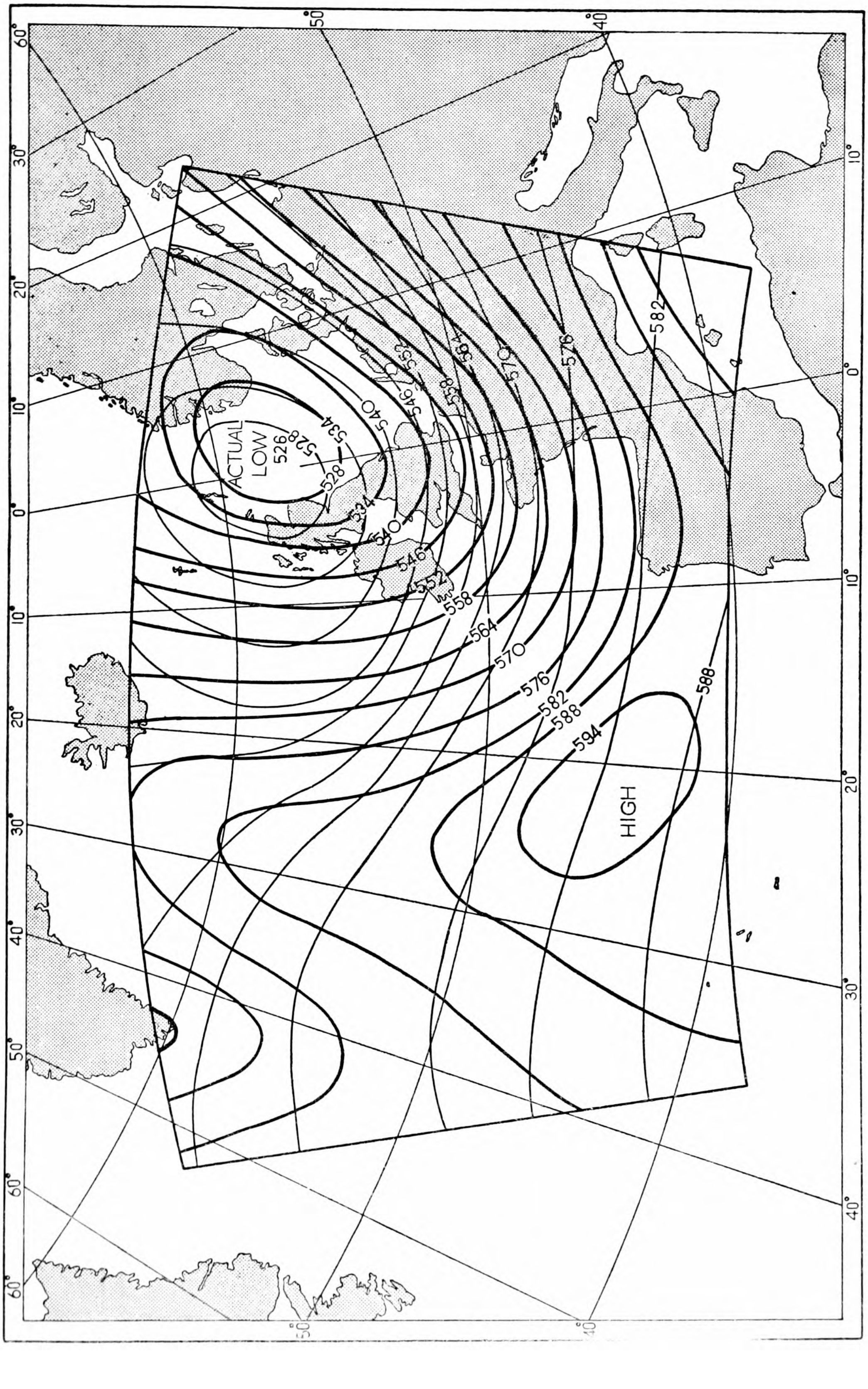


FIGURE 6(*b*). 500-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thin lines represent actual contours and thick lines represent contours obtained by using twice the actual changes during the previous 24 hours.

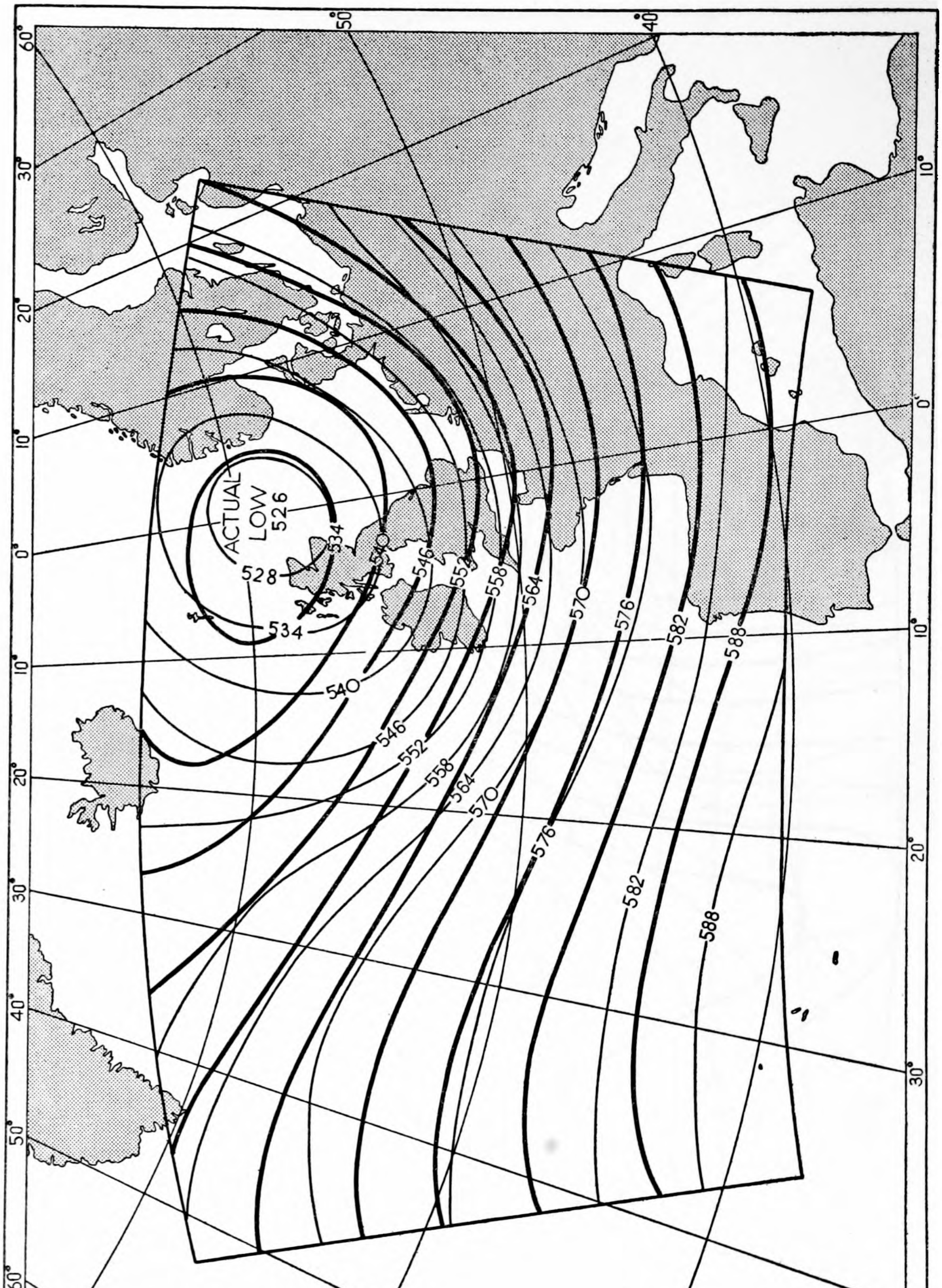


FIGURE 6(*c*). 500-mb contours, in decametres, 1800 G.M.T., 25 August 1957
Thin lines represent actual contours and thick lines represent contours obtained by using half the actual changes during the previous 24 hours.

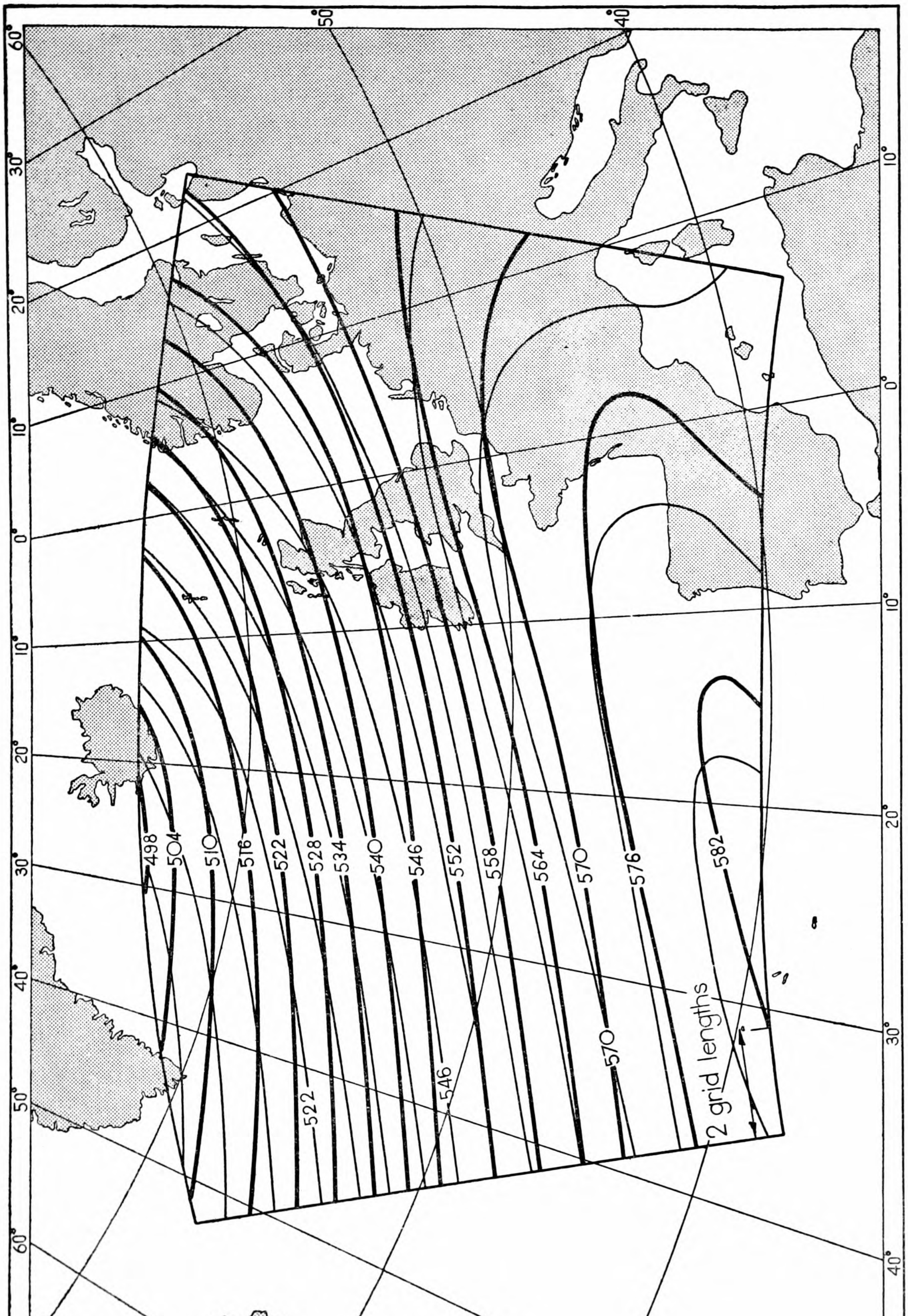


FIGURE 7. 500-mb contours, in decametres, 1200 G.M.T., 18 April 1957

Thin lines represent actual contours and thick lines represent actual contours moved two grid lengths east.

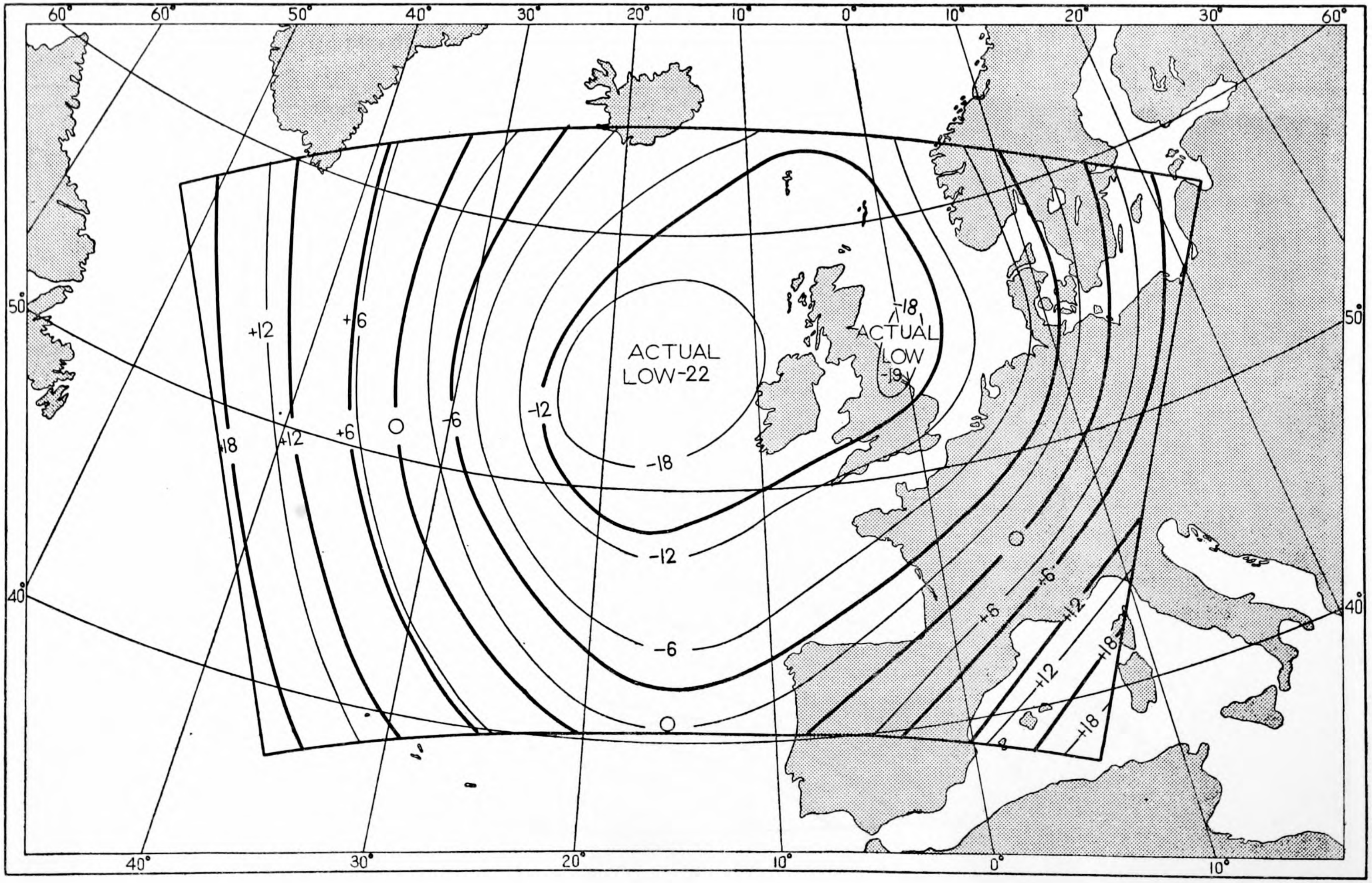


FIGURE 8(*a*). 1000-mb contours, in decametres, 0600 G.M.T., 4 November 1957

Thin lines represent actual contours and thick lines represent contours obtained by using the formula for the forecast heights given on p. 14 and substituting $\alpha = +\frac{1}{2}$.

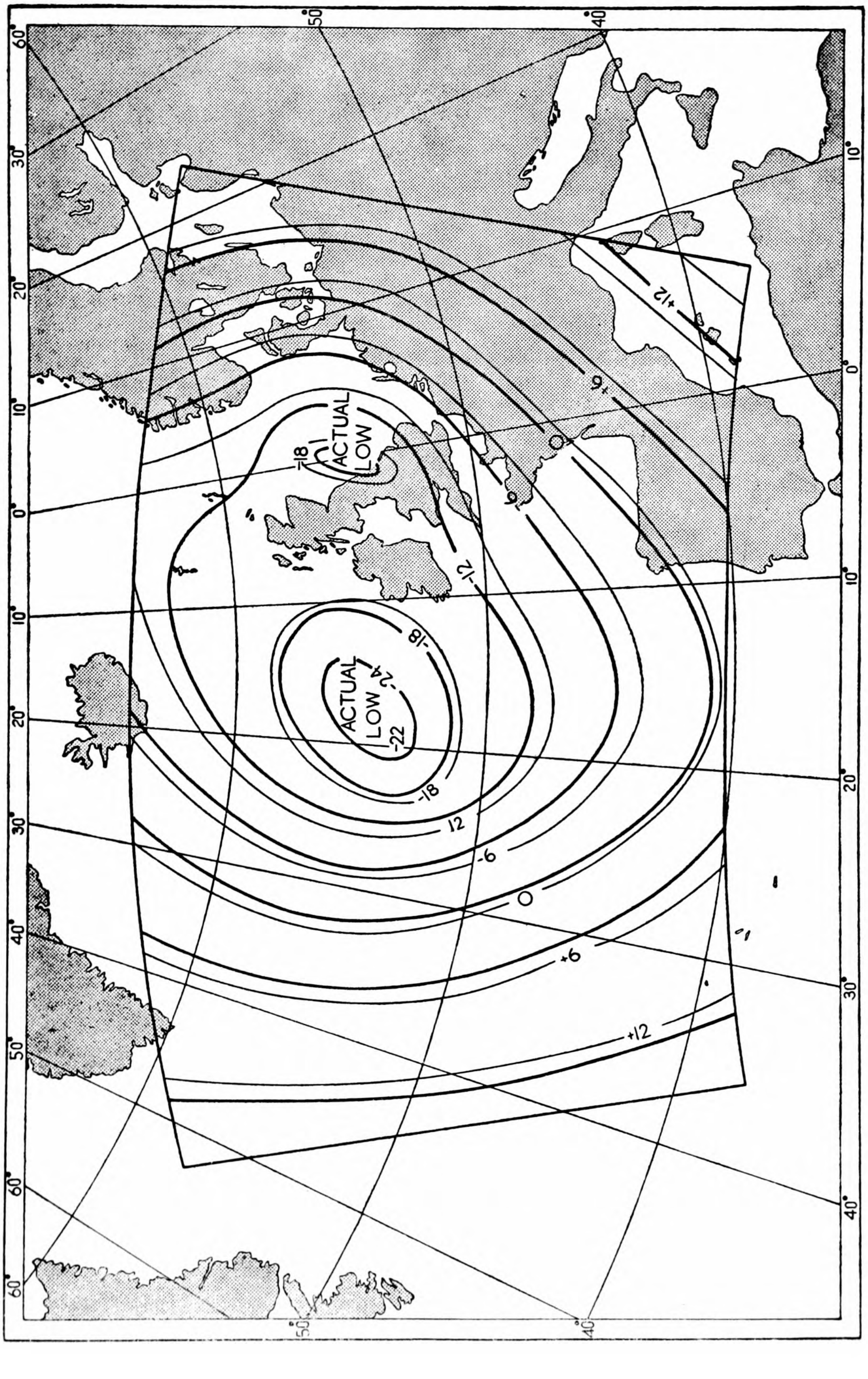


FIGURE 8(*b*). 1000-mb contours, in decametres, 0600 G.M.T., 4 November 1957
Thin lines represent actual contours and thick lines represent contours obtained by using the formula for the forecast heights given on p. 14 and substituting $\alpha = -1$.

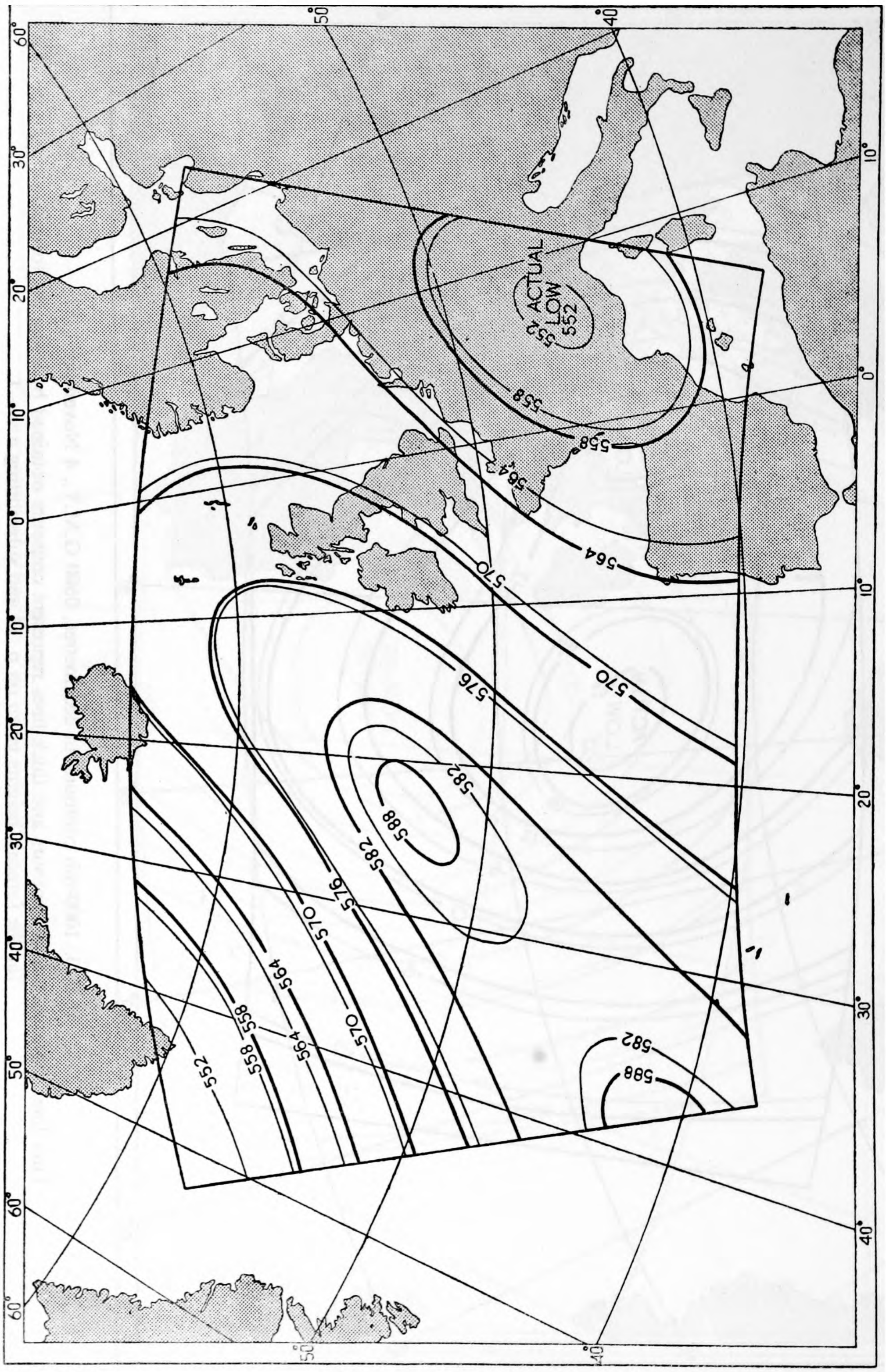


FIGURE 9. 500-mb contours, in decametres, 1200 G.M.T., 28 May 1957

Thin lines represent actual contours and thick lines represent contours obtained by using the formula for the forecast heights given on p. 14 and substituting $\alpha = +1$.

Printed in England under the authority of HER MAJESTY'S STATIONERY OFFICE by
James Townsend & Sons, Ltd., Exeter

Wt.2423 BM4227 K10 12/60

Published by
HER MAJESTY'S STATIONERY OFFICE

To be purchased from
York House, Kingsway, London W.C.2
423 Oxford Street, London W.1
13A Castle Street, Edinburgh 2
109 St. Mary Street, Cardiff
39 King Street, Manchester 2
50 Fairfax Street, Bristol 1
2 Edmund Street, Birmingham 3
80 Chichester Street, Belfast 1
or through any bookseller

*Printed in England*

S. O. Code No. 40-164-9