U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 385

Skill of Medium Range Forecasts

Francis D. Hughes
Prediction Branch, Climate Analysis Center

February 1992

385

PURPOSE

This paper depicts in a graphical manner the skill of the Medium Range (3-10 day) man and machine (numerical model guidance) forecasts. It will be updated each February in order to present the latest scores for each of the several forecast categories in the medium range forecast program (MRFP). Only scores with at least a 5-year period of record are presented. This paper contains the standardized and unstandardized mean sea level pressure and 500mb correlation; the Gilman precipitation skill; the minimum/maximum average absolute temperature error; and the 5-day mean normalized 500mb correlation, temperature, and precipitation skill scores.

Figure 1 depicts the North American (NOAM, 130 grid points) and United States (US, 88 grid points) mean sea level pressure (MSLP) and 500mb correlation score verification areas.

Figure 2 is a plot of the calendar year 1991 seasonal and 10 year (1981-1990) average seasonal (darker lines) standarized correlation scores for the man (official) NOAM area MSLP progs verifying on days 3, 4, and 5 after forecast day. (See appendix A for an explanation of this score).

Figure 3 is similar to Figure 2 except scores are for the NMC/NWP model progs.

Figure 4 is a plot of the 24/22 year (1968/70-1991) standardized correlation scores for the man (official) and NMC/NWP model NOAM area MSLP progs verifying on days 3, 4, and 5 after forecast day.

Figure 5 is a plot of the calendar year pentad standardized correlation scores for the NMC/NWP model and ECMWF model (darker line) NOAM area MSLP progs verifying on day 4 after forecast day.

Figure 6 is similar to Figure 3 except the level is 500mb.

Figure 7 is similar to Figure 4 except the level is 500mb.

Figure 8 is similar to Figure 5 except the level is 500mb.

Figure 9 is similar to Figure 8 except scores are for the blend (a concensus of the bias and linear regression corrected NMC/NWP model, the ECMWF model, and the multiple regression corrected NMC/NWP model and ECMWF model) and the NMC/NWP model (darker line) progs verifying on days 6-10 after forecast day.

Figure 10 is a plot of the 12 year (1980-1991) standardized correlation scores for the man (official), NMC/NWP model, and ECMWF model (82-91) NOAM area 500mb progs verifying on days 6-10 after forecast day.

Figure 11 depicts the 41/61 stations in the US where the 1-5/6-10 day temperature forecasts are verified.

Figure 12 is a plot of the calendar year 1991 bi-monthly and 10 year (1981-1990) average bi-monthly (darker lines) US area temperature forecasts verifying on days 3, 4, and 5 after forecast day.

Figure 13 is similar to Figure 12 except scores are for the maximum.

Figure 14 is a plot of the 21 year (1971-1991) absolute error minimum temperature scores for the man (official) and climatology US area temperature forecasts verifying on days 3, 4, and 5 after forecast day.

Figure 15 is similar to Figure 14 except scores are for the maximum.

Figure 16 is a plot of the 21 year (1971-1991) absolute error (minimum + maximum)/2 temperature scores for the man (official), perfect prog, and climatology US area temperature forecasts verifying on days (3+4+5)/3 after forecast day.

Figure 17 is a plot of the 14 year (1978-1991) 3-class and 5-class temperature skill scores for the man (official) and persistence US area categorical temperature forecasts verifying on days 6-10 after forecast day. (See appendix B for an explanation of this score).

Figure 18 depicts the 100 stations in the US where the precipitation forecasts are verified.

Figure 19 is an example of a day 3, 4, or 5 precipitation forecast. The dashed lines are the 24-hour departure from probability of precipitation (DN POP) forecast for January 3. The solid lines are the 24-hour climatological (normal) probability of precipitation (NPOP) for the first 15 days of January. A total of (DN POP + NPOP) > 30 is considered "yes" forecast of precipiation (> 0.01 inch). All stations with an (NPOP) > 30 are considered as a "yes" climatological forecast of precipitation.

Figure 20 is a plot of the calendar year 1991 seasonal and 11 year (1981-1990) average seasonal (darker lines) Gilman precipitation skill scores for the man (official) US area precipitation forecasts verifying on days 3, 4, and 5 after forecast day. (See appendix C for an explanation of this score).

Figure 21 is a plot of the 22 year (1970-1991) Gilman precipitation skill scores for the man (official) and climatology US area precipitation forecasts verifying on days 3, 4, and 5 after forecast day.

Figure 22 is a plot of the 14 year (1978-1991) 3-class precipitation skill scores for the man (official), NMC/NWP model, and climatology US area categorical precipitation forecasts verifying on days 1-5 after forecast day. (See appendix D for an explanation of this score).

Figure 23 is similar to Figure 22 except the forecasts are for days 6-10 after forecast day.

SECTION 1

Man and Machine (NMC/NWP Model Guidance)

Mean Sea Level Pressure and 500mb Correlation Scores

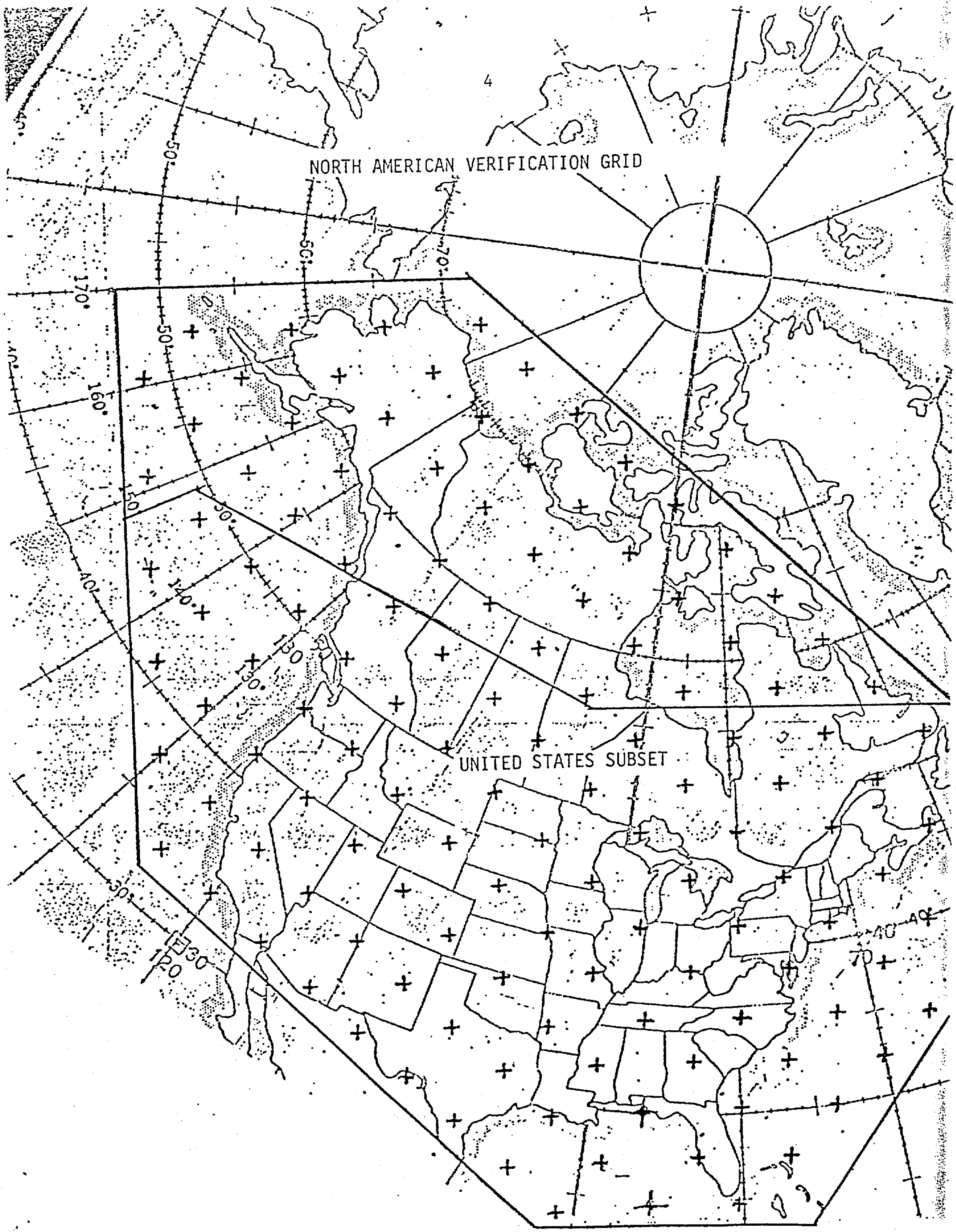


Figure 1

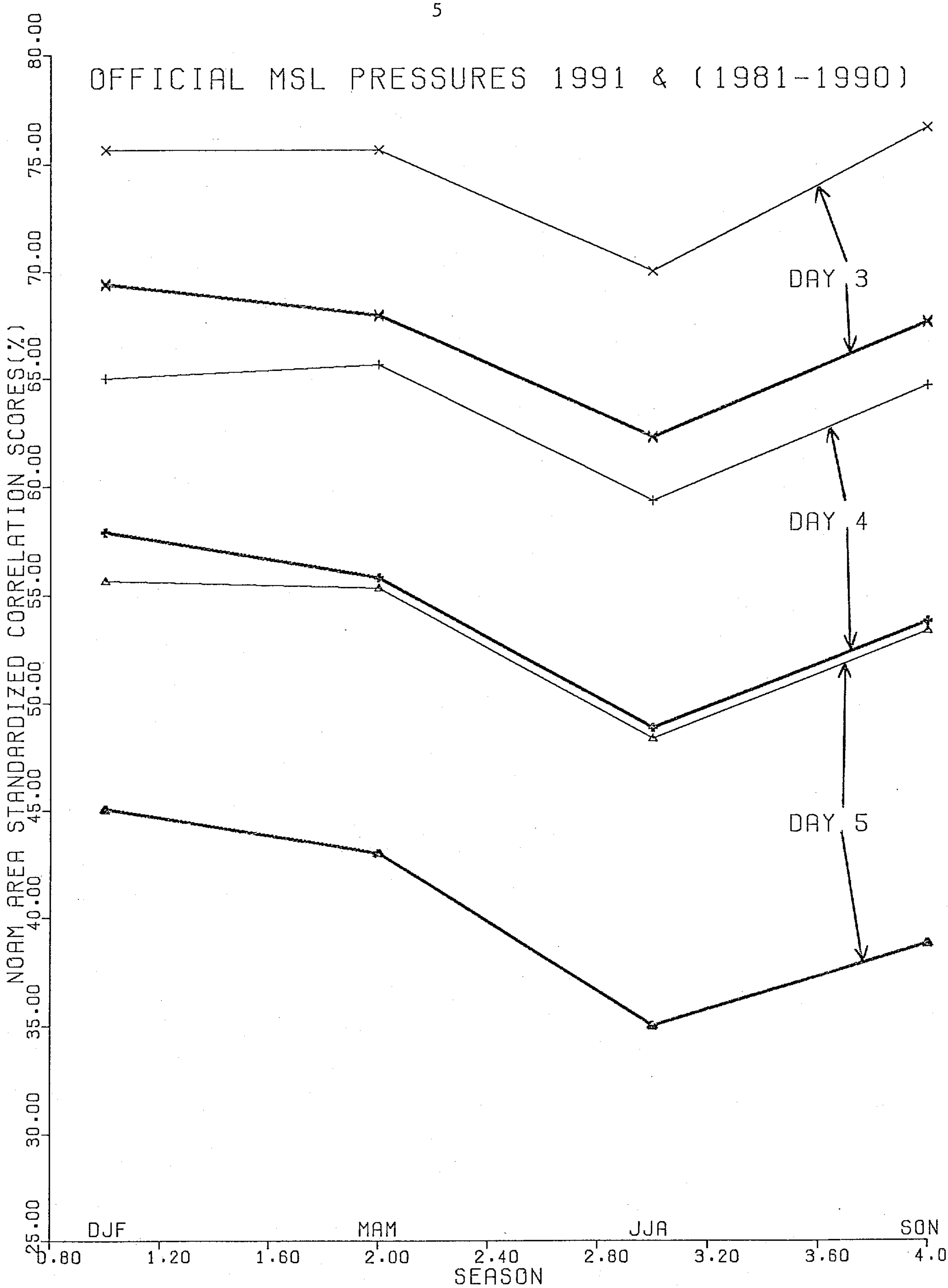

OFFICIAL MSL PRESSURES 1991 & (1981-1990)
DAY 3
DAY 4
DAY 5
NOAM AREA STANDARDIZED CORRELATION SCORES(%)
80.00
75.00
70.00
65.00
60.00
55.00
50.00
45.00
40.00
35.00
30.00
25.00
DJF
MAM
JJA
SON
0.80
1.20
1.60
2.00
2.40
2.80
3.20
3.60
4.00
SEASON

Figure 2

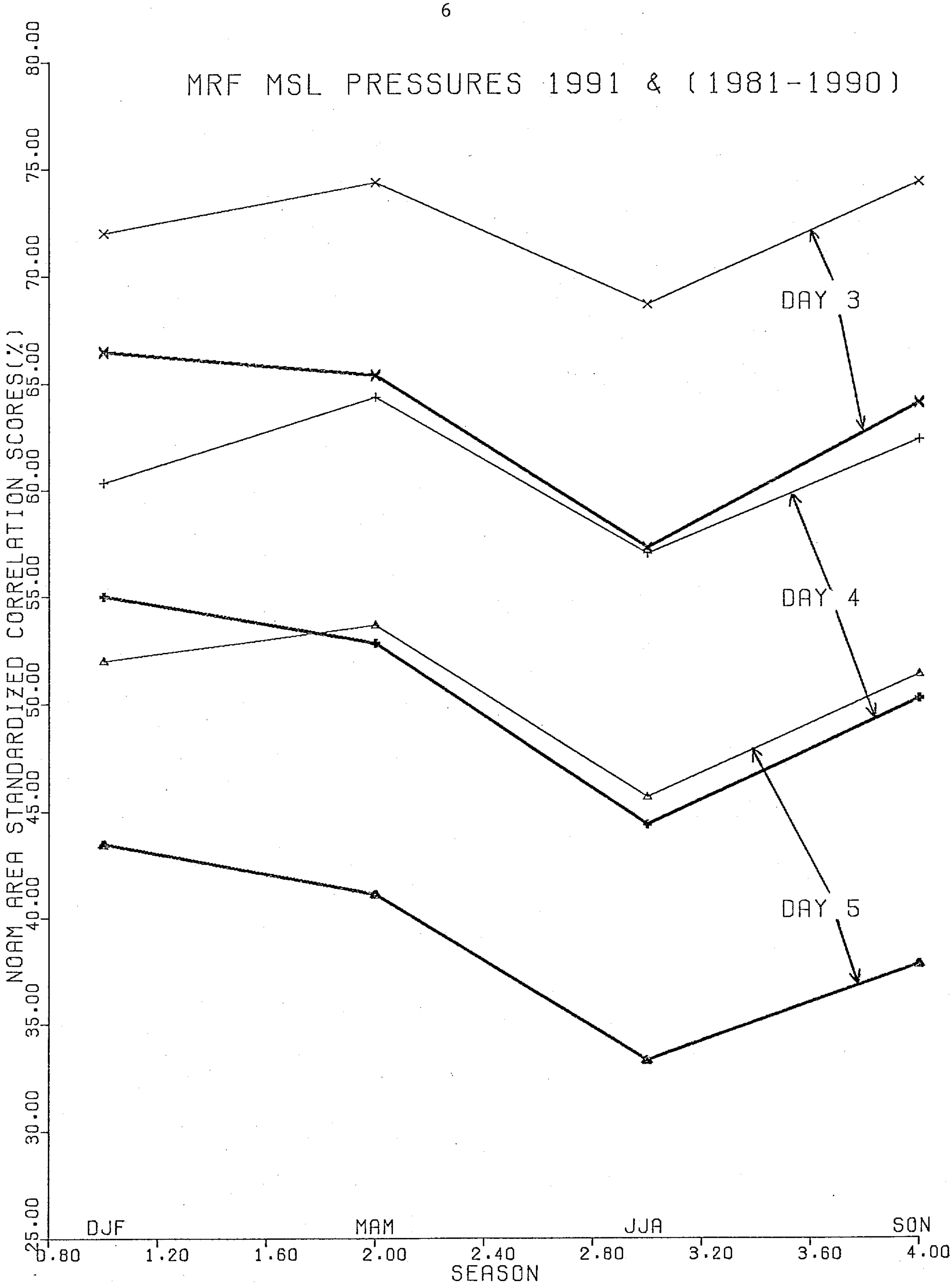

MRF MSL PRESSURES 1991 & (1981-1990)
NOAM AREA STANDARDIZED CORRELATION SCORES(%)
80.00
75.00
70.00
65.00
60.00
55.00
50.00
45.00
40.00
35.00
30.00
25.00
DAY 3
DAY 4
DAY 5
DJF
MAM
JJA
SON
0.80
1.20
1.60
2.00
2.40
2.80
3.20
3.60
4.00
SEASON

Figure 3

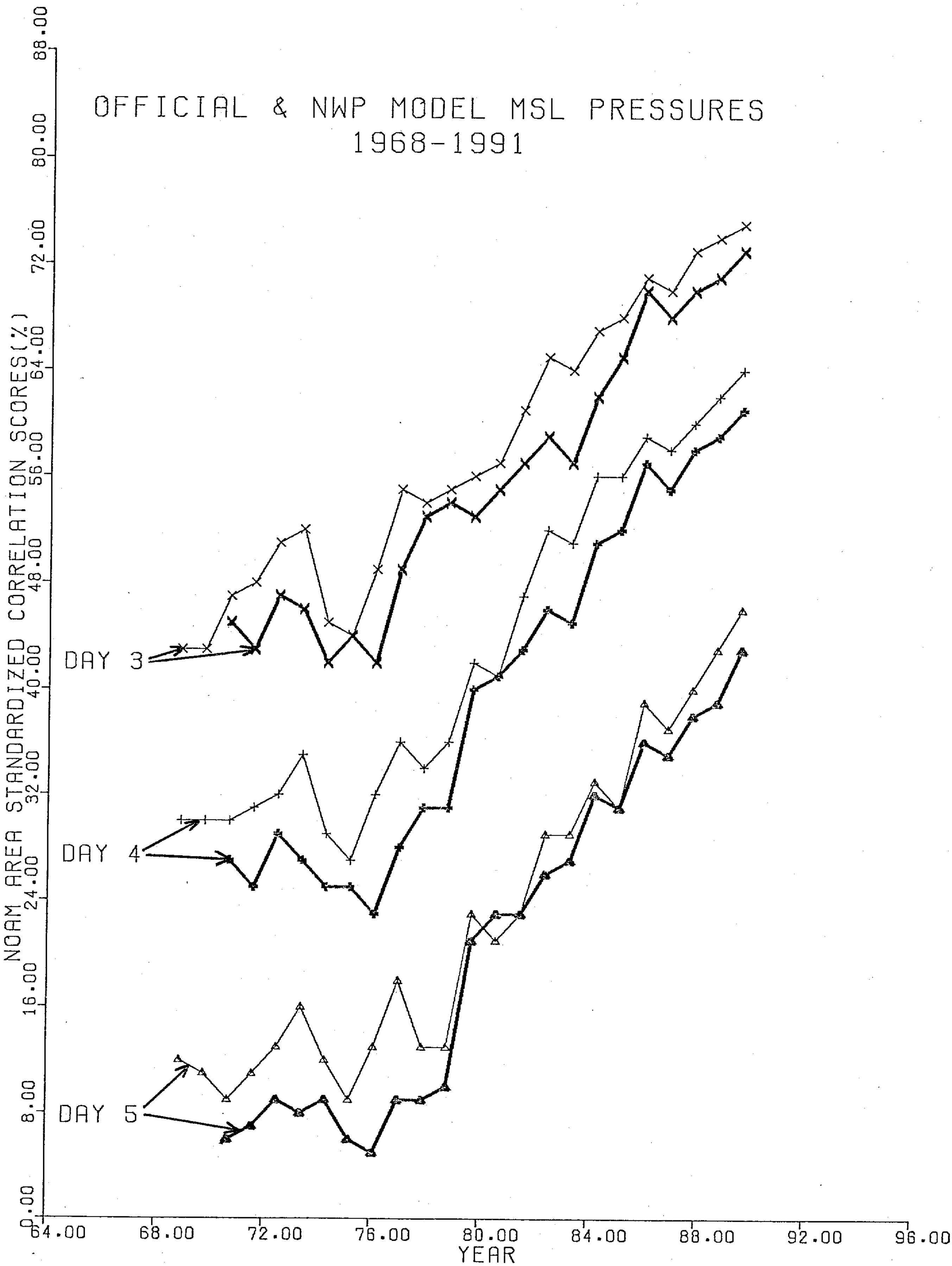

OFFICIAL & NWP MODEL MSL PRESSURES
1968-1991
NOAM AREA STANDARDIZED CORRELATION SCORES(%)
YEAR
DAY 3
DAY 4
DAY 5
0.00
8.00
16.00
24.00
32.00
40.00
48.00
56.00
64.00
72.00
80.00
88.00
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00

Figure 4

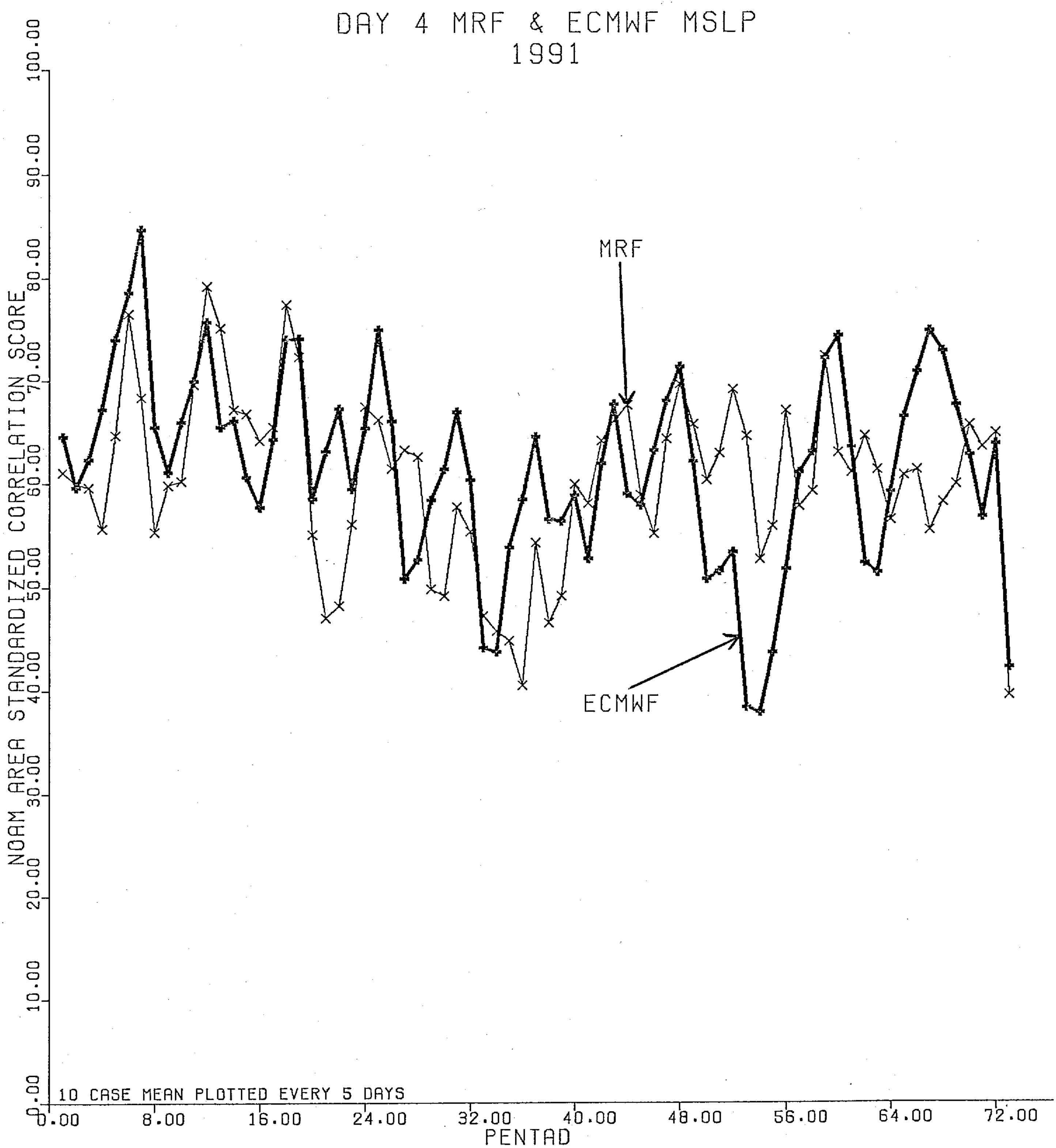

DAY 4 MRF & ECMWF MSLP
1991
NOAM AREA STANDARDIZED CORRELATION SCORE
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
100.00
MRF
ECMWF
10 CASE MEAN PLOTTED EVERY 5 DAYS
0.00
8.00
16.00
24.00
32.00
40.00
48.00
56.00
64.00
72.00
PENTAD

Figure 5

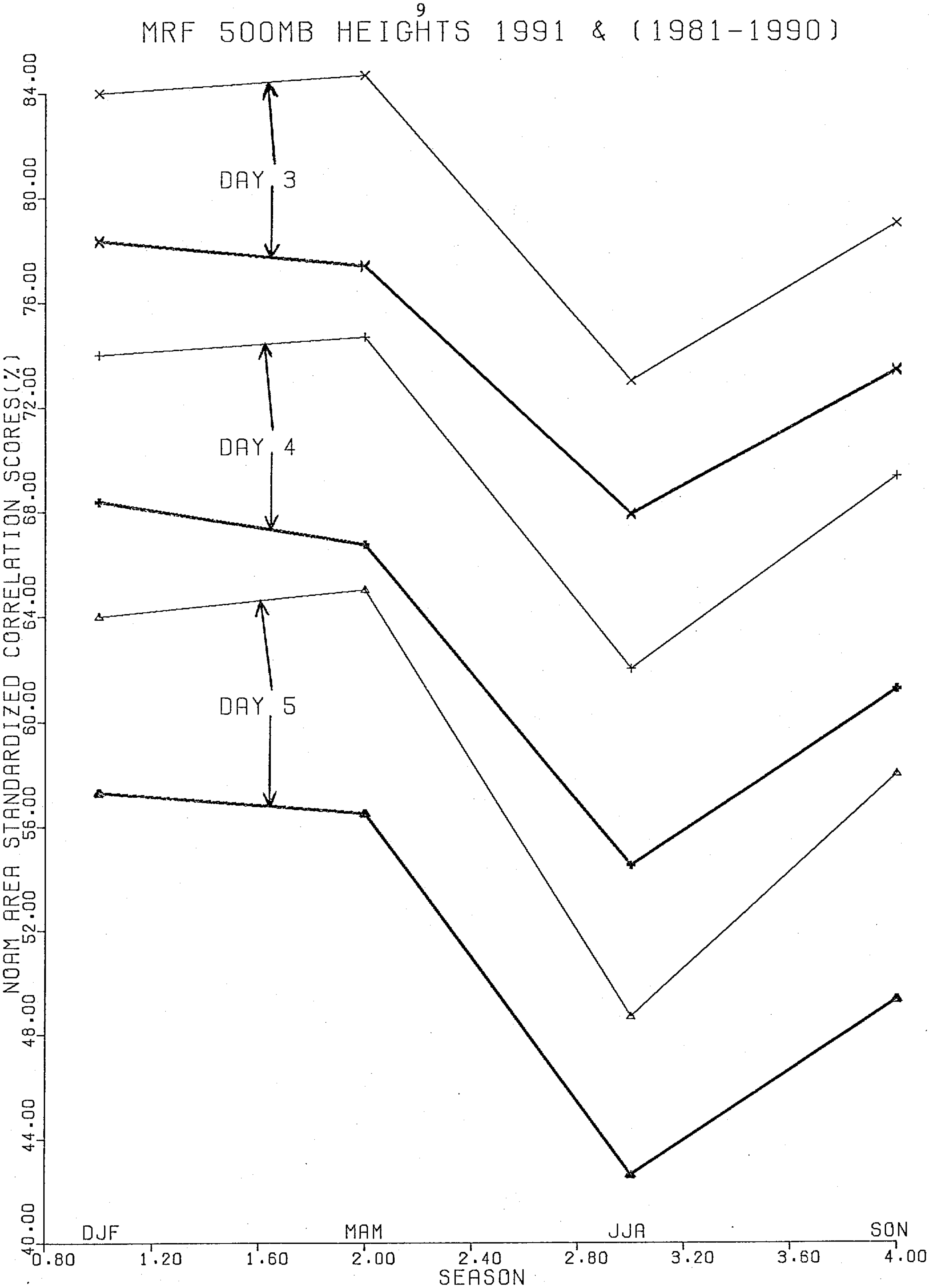

MRF 500MB HEIGHTS 1991 & (1981-1990)
NOAM AREA STANDARDIZED CORRELATION SCORES(%)
84.00
80.00
76.00
72.00
68.00
64.00
60.00
56.00
52.00
48.00
44.00
40.00
DAY 3
DAY 4
DAY 5
DJF
MAM
JJA
SON
0.80
1.20
1.60
2.00
2.40
2.80
3.20
3.60
4.00
SEASON

Figure 6

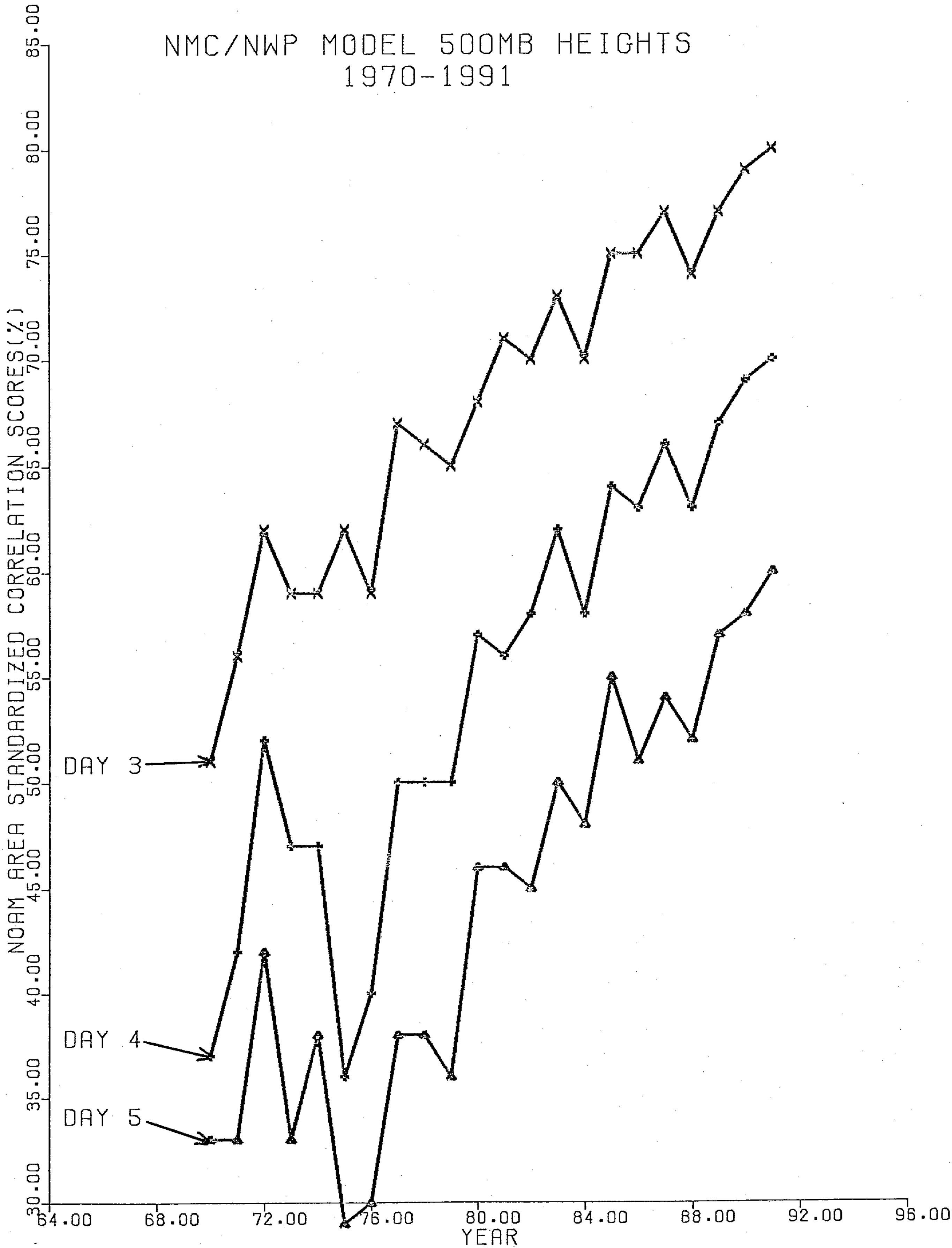

NMC/NWP MODEL 500MB HEIGHTS
1970-1991
NOAM AREA STANDARDIZED CORRELATION SCORES(%)
YEAR
DAY 3
DAY 4
DAY 5
30.00
35.00
40.00
45.00
50.00
55.00
60.00
65.00
70.00
75.00
80.00
85.00
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00

Figure 7

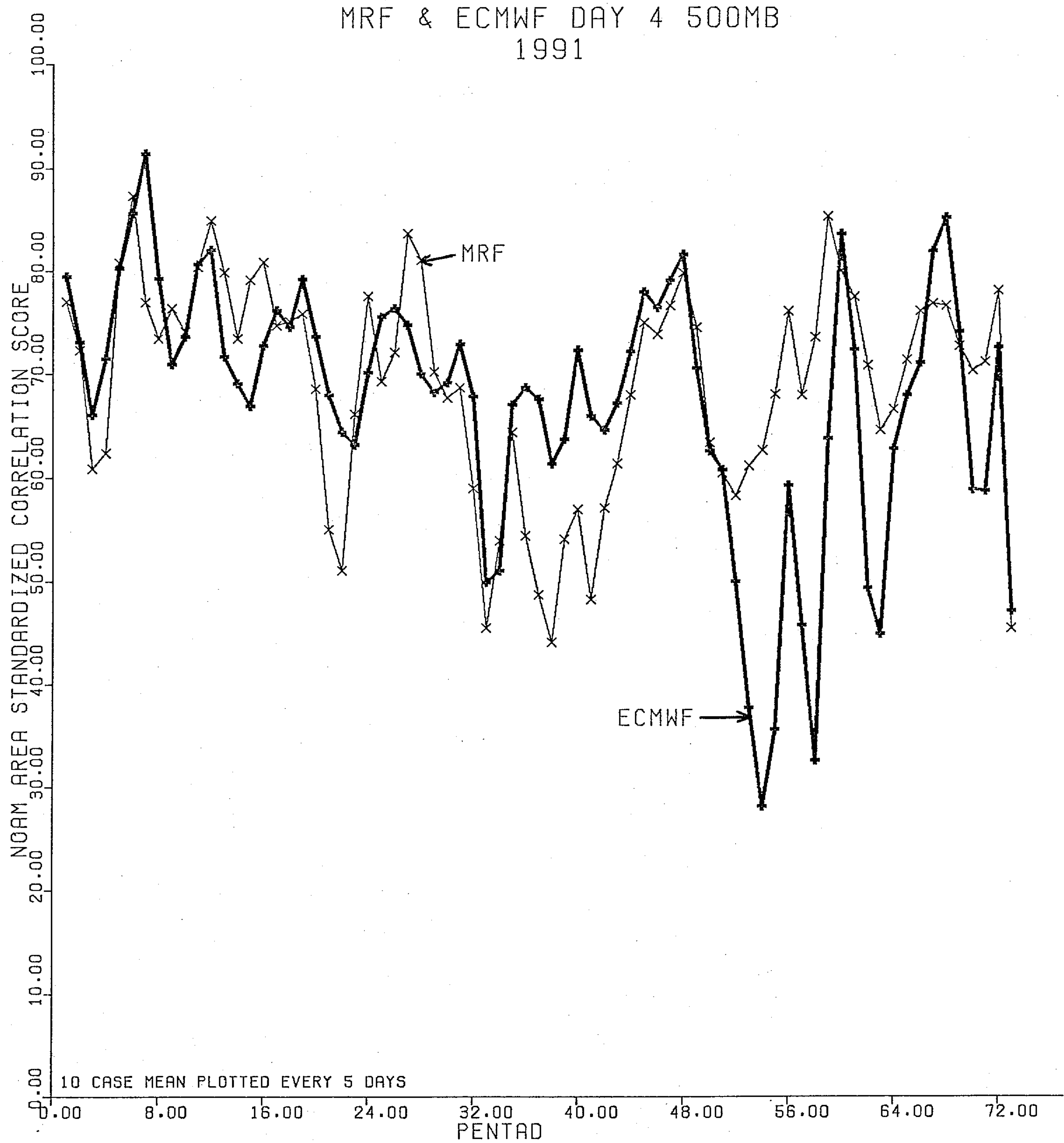

MRF & ECMWF DAY 4 500MB
1991
NOAM AREA STANDARDIZED CORRELATION SCORE
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
100.00
MRF
ECMWF
10 CASE MEAN PLOTTED EVERY 5 DAYS
0.00
8.00
16.00
24.00
32.00
40.00
48.00
56.00
64.00
72.00
PENTAD

Figure 8

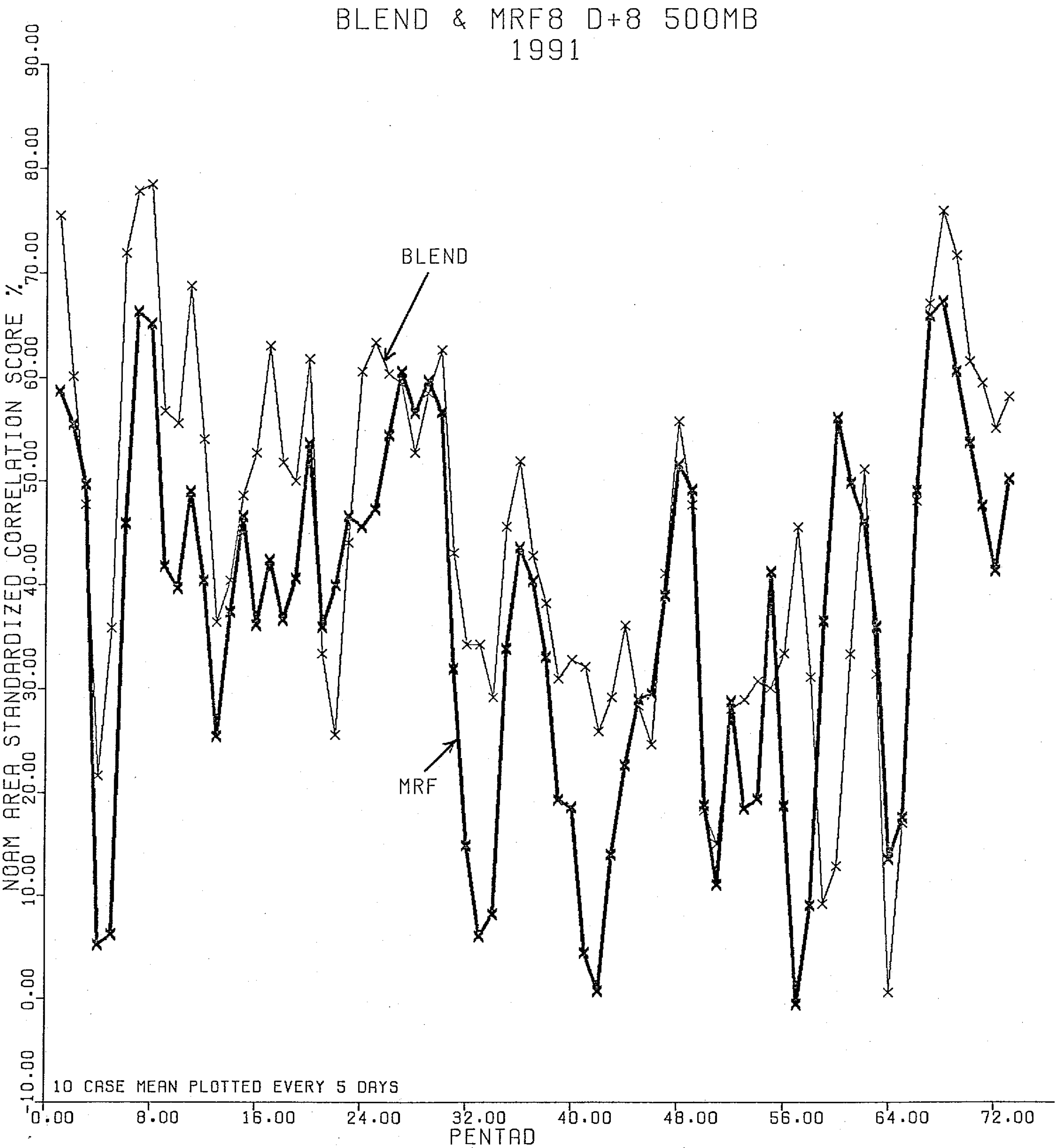

BLEND & MRF8 D+8 500MB
1991
BLEND
MRF
NOAM AREA STANDARDIZED CORRELATION SCORE %
10 CASE MEAN PLOTTED EVERY 5 DAYS
PENTAD
-10.00
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
0.00
8.00
16.00
24.00
32.00
40.00
48.00
56.00
64.00
72.00

Figure 9

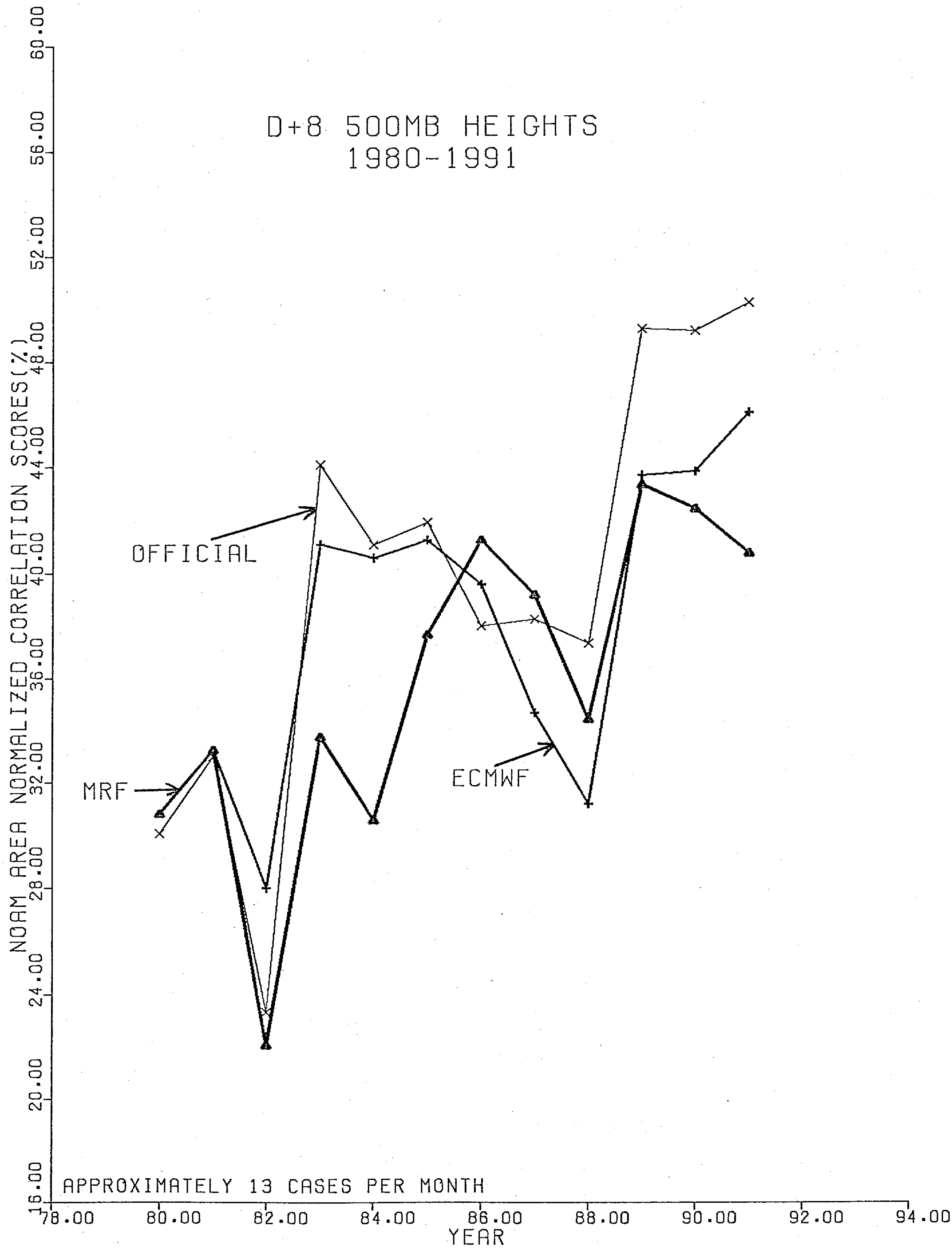

D+8 500MB HEIGHTS
1980-1991
NOAM AREA NORMALIZED CORRELATION SCORES(%)
60.00
56.00
52.00
48.00
44.00
40.00
36.00
32.00
28.00
24.00
20.00
16.00
OFFICIAL
MRF
ECMWF
APPROXIMATELY 13 CASES PER MONTH
78.00
80.00
82.00
84.00
86.00
88.00
90.00
92.00
94.00
YEAR

Figure 10

SECTION 2

Man and Machine (Klein-Lewis Perfect Prog Guidance) Average Absolute Error Temperature Scores

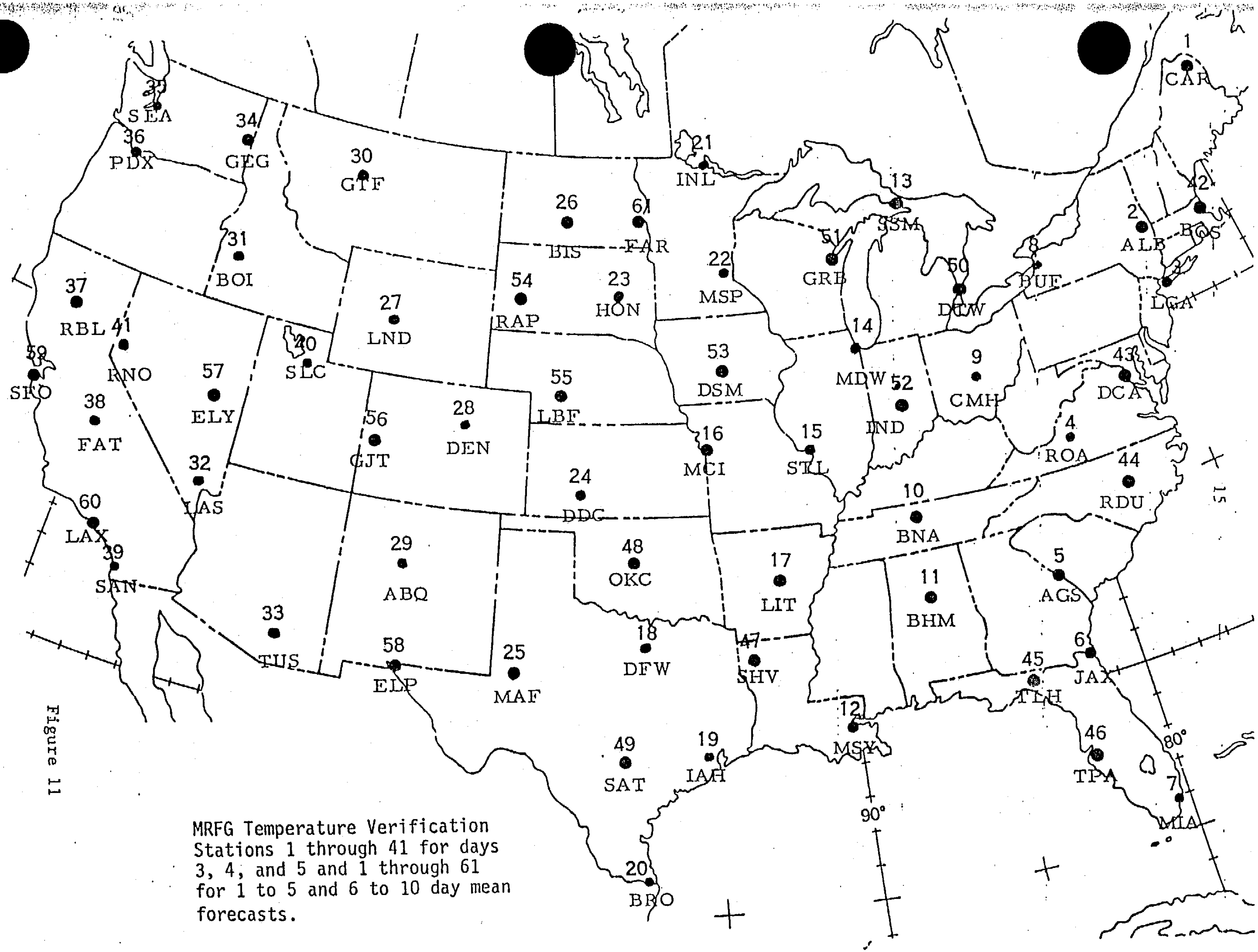
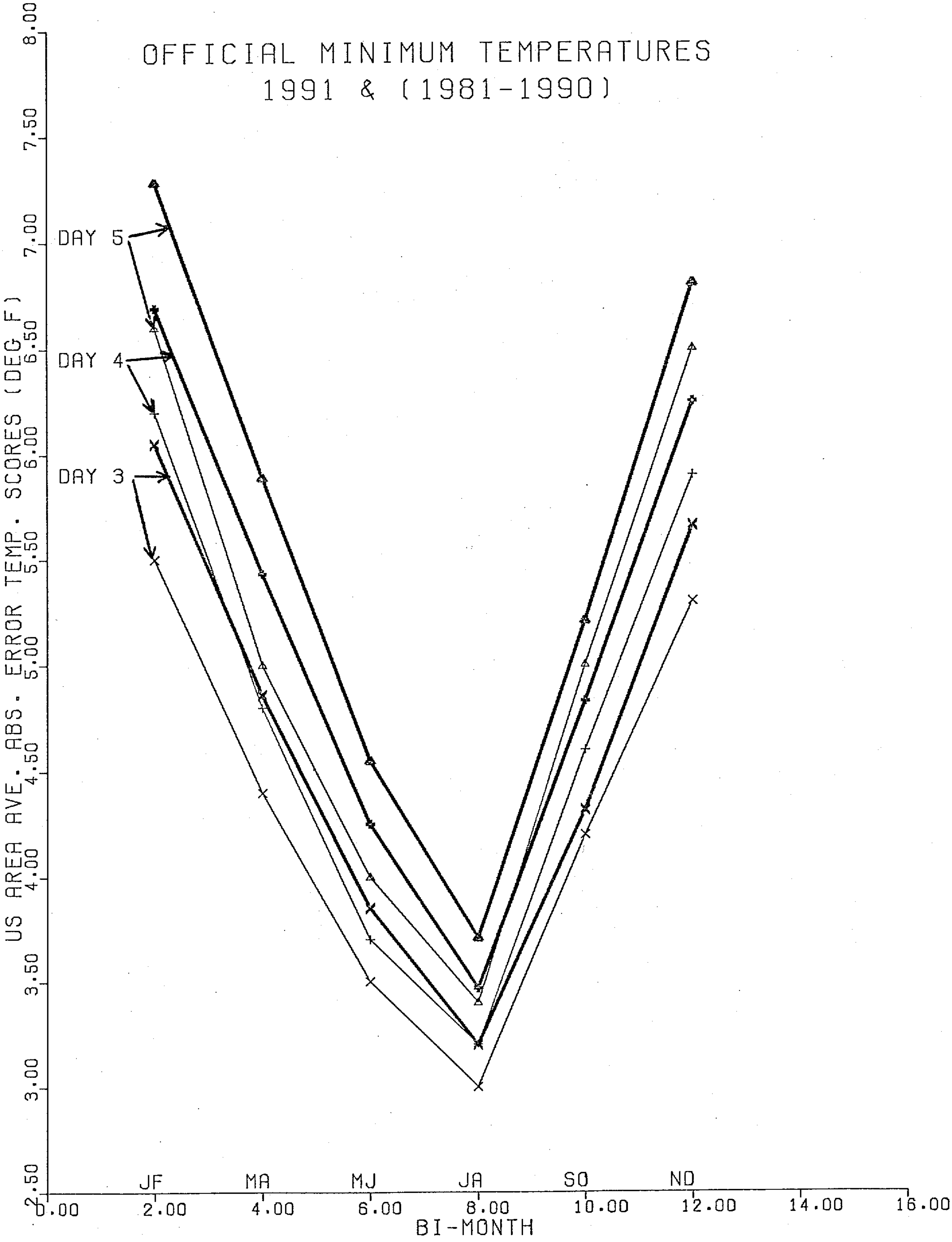


Figure 11

MRFG Temperature Verification Stations 1 through 41 for days 3, 4, and 5 and 1 through 61 for 1 to 5 and 6 to 10 day mean forecasts.



OFFICIAL MINIMUM TEMPERATURES
1991 & (1981-1990)
DAY 5
DAY 4
DAY 3
US AREA AVE. ABS. ERROR TEMP. SCORES (DEG F)
8.00
7.50
7.00
6.50
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
JF
MA
MJ
JA
SO
ND
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
BI-MONTH

Figure 12

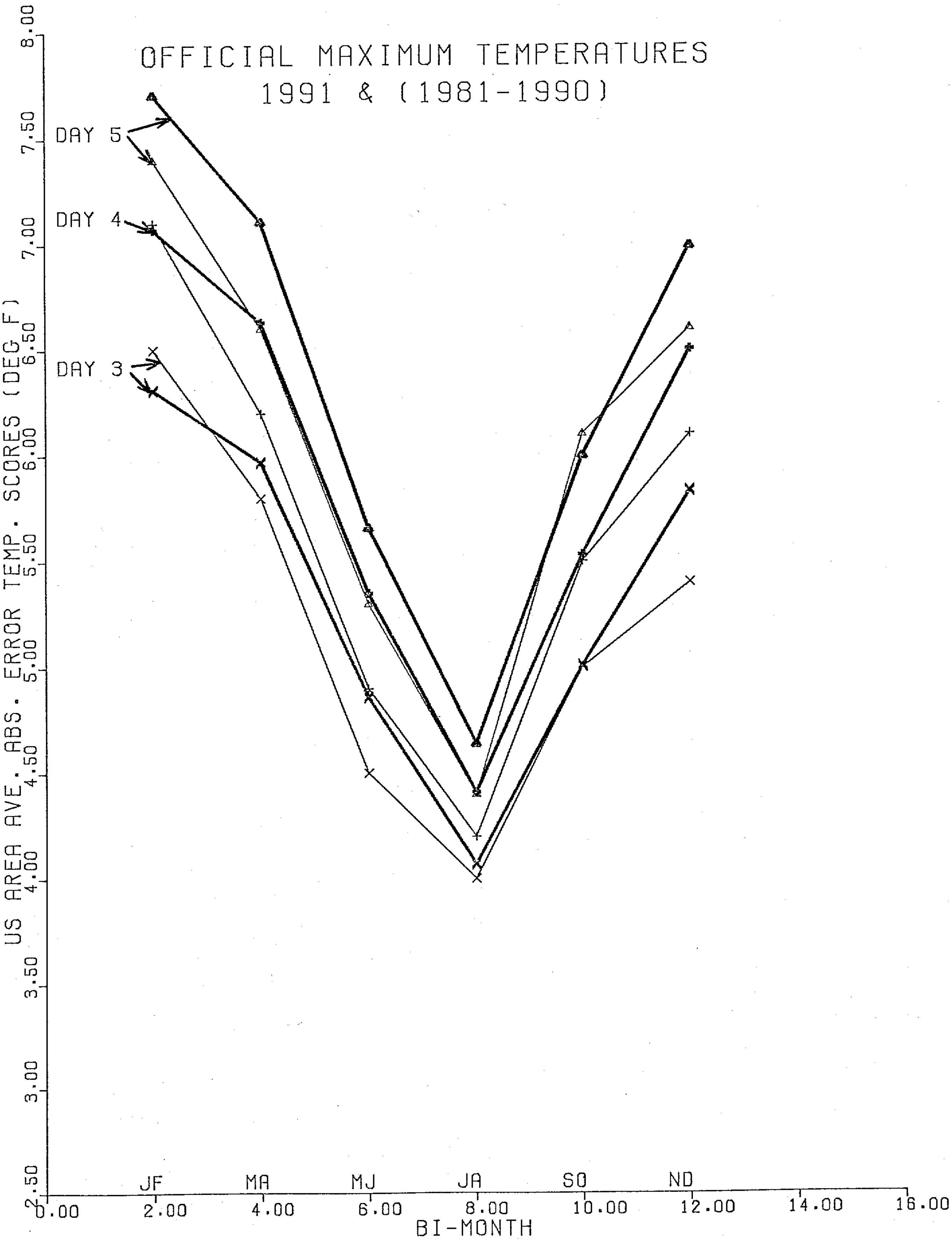

OFFICIAL MAXIMUM TEMPERATURES
1991 & (1981-1990)
DAY 5
DAY 4
DAY 3
US AREA AVE. ABS. ERROR TEMP. SCORES (DEG F)
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
6.50
7.00
7.50
8.00
JF
MA
MJ
JA
SO
ND
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
BI-MONTH

Figure 13

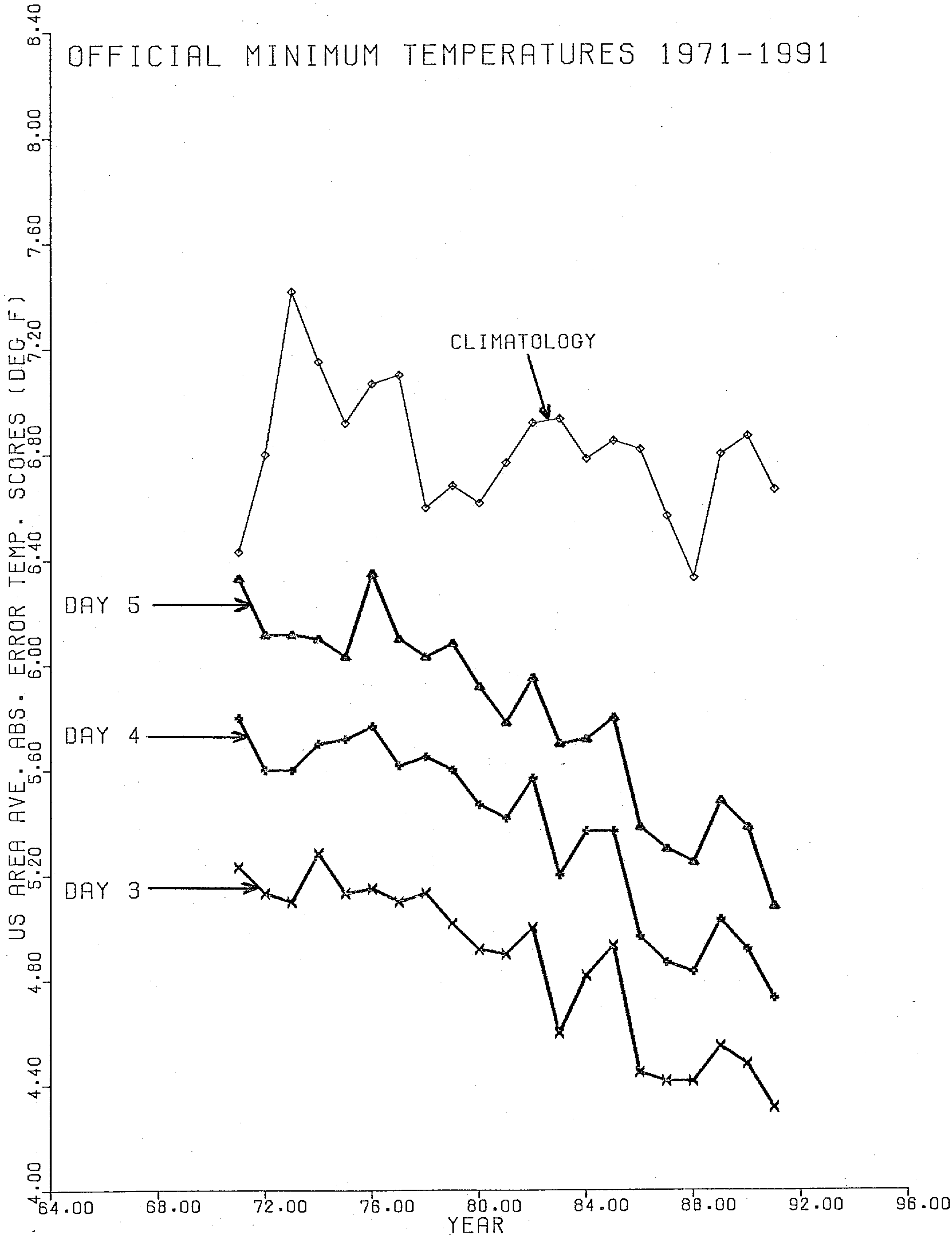

OFFICIAL MINIMUM TEMPERATURES 1971-1991
CLIMATOLOGY
DAY 5
DAY 4
DAY 3
US AREA AVE. ABS. ERROR TEMP. SCORES (DEG F)
8.40
8.00
7.60
7.20
6.80
6.40
6.00
5.60
5.20
4.80
4.40
4.00
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00
YEAR

Figure 14

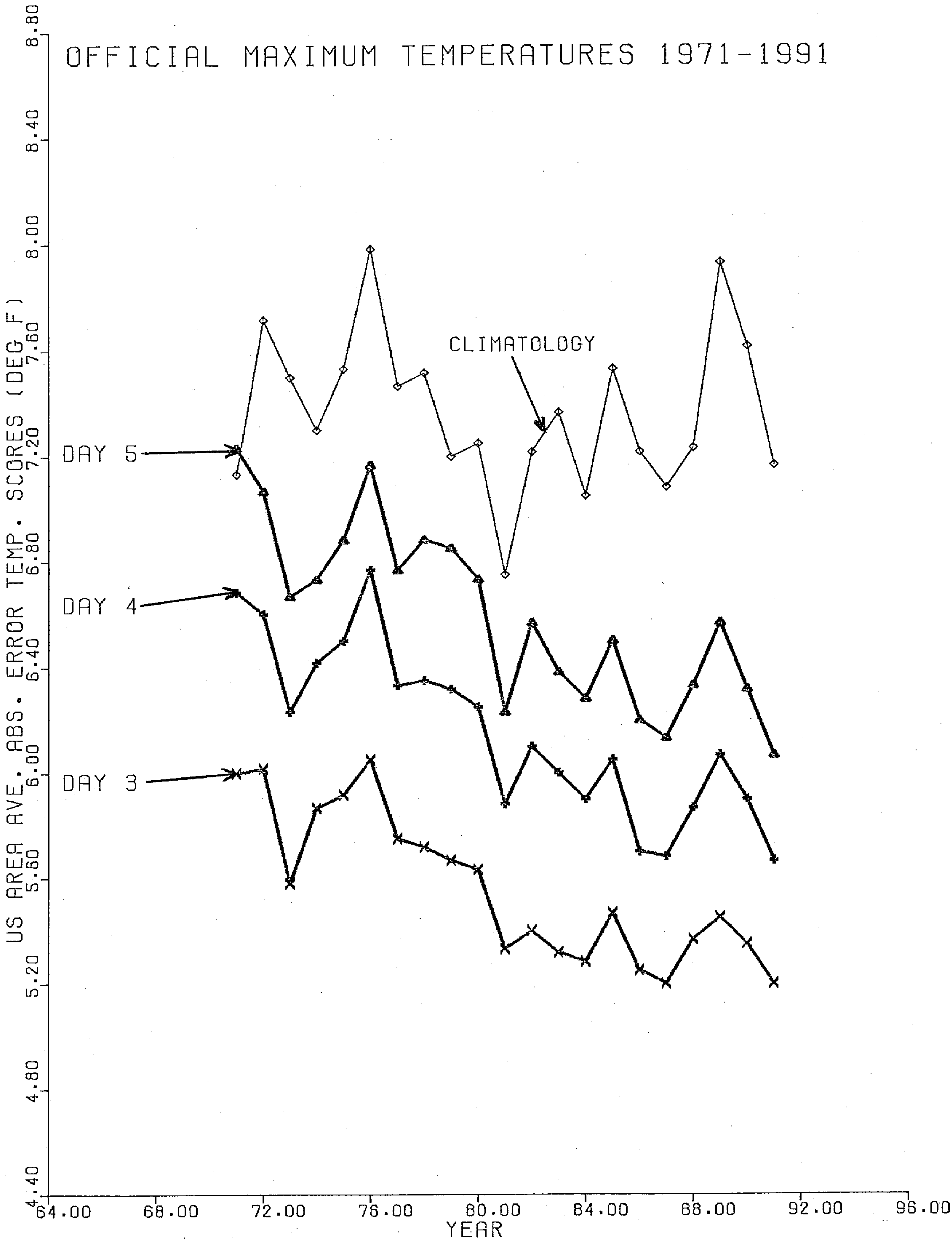

OFFICIAL MAXIMUM TEMPERATURES 1971-1991
CLIMATOLOGY
DAY 5
DAY 4
DAY 3
US AREA AVE. ABS. ERROR TEMP. SCORES (DEG F)
YEAR
8.80
8.40
8.00
7.60
7.20
6.80
6.40
6.00
5.60
5.20
4.80
4.40
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00

Figure 15

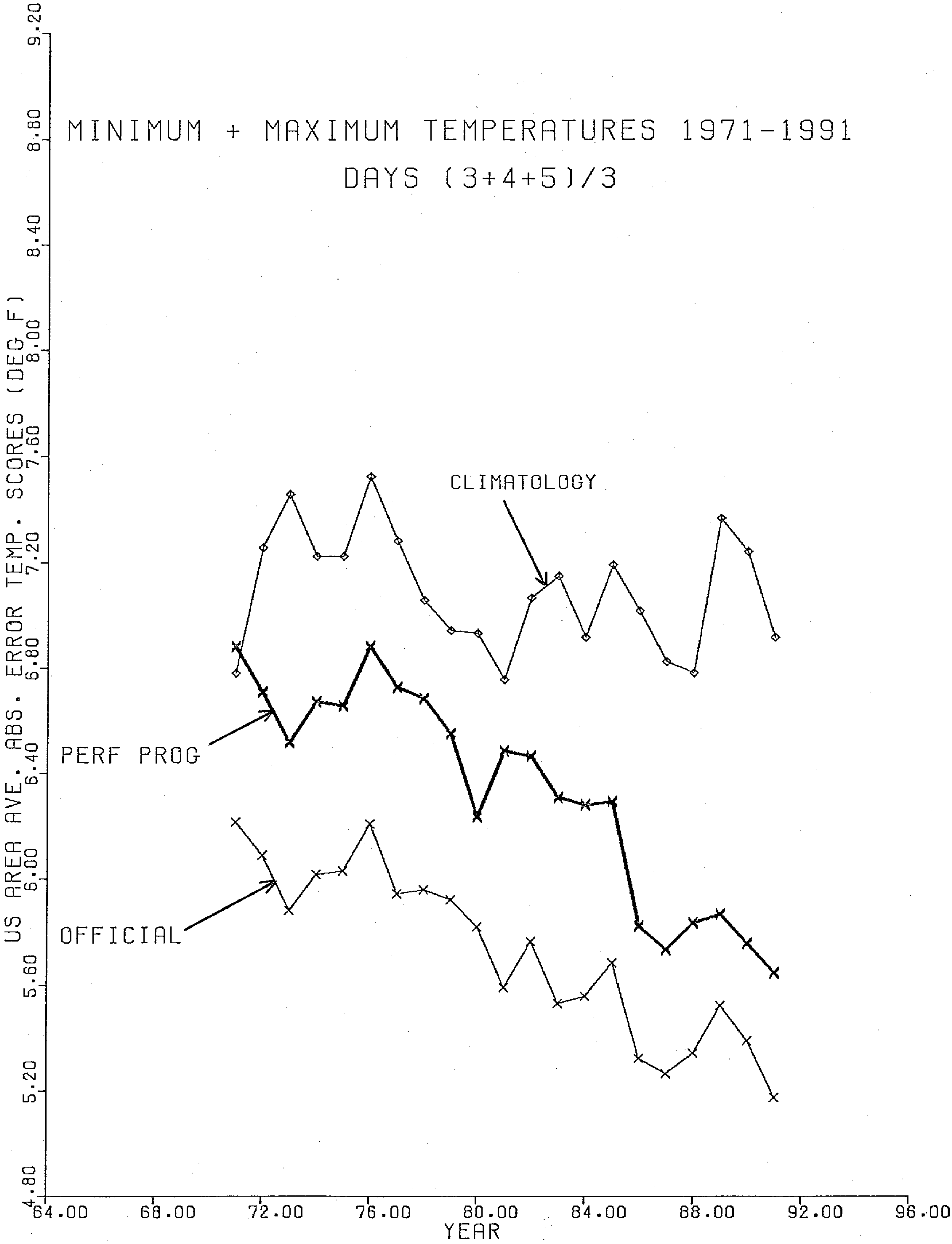

MINIMUM + MAXIMUM TEMPERATURES 1971-1991
DAYS (3+4+5)/3
CLIMATOLOGY
PERF PROG
OFFICIAL
US AREA AVE. ABS. ERROR TEMP. SCORES (DEG F)
9.20
8.80
8.40
8.00
7.60
7.20
6.80
6.40
6.00
5.60
5.20
4.80
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00
YEAR

Figure 16

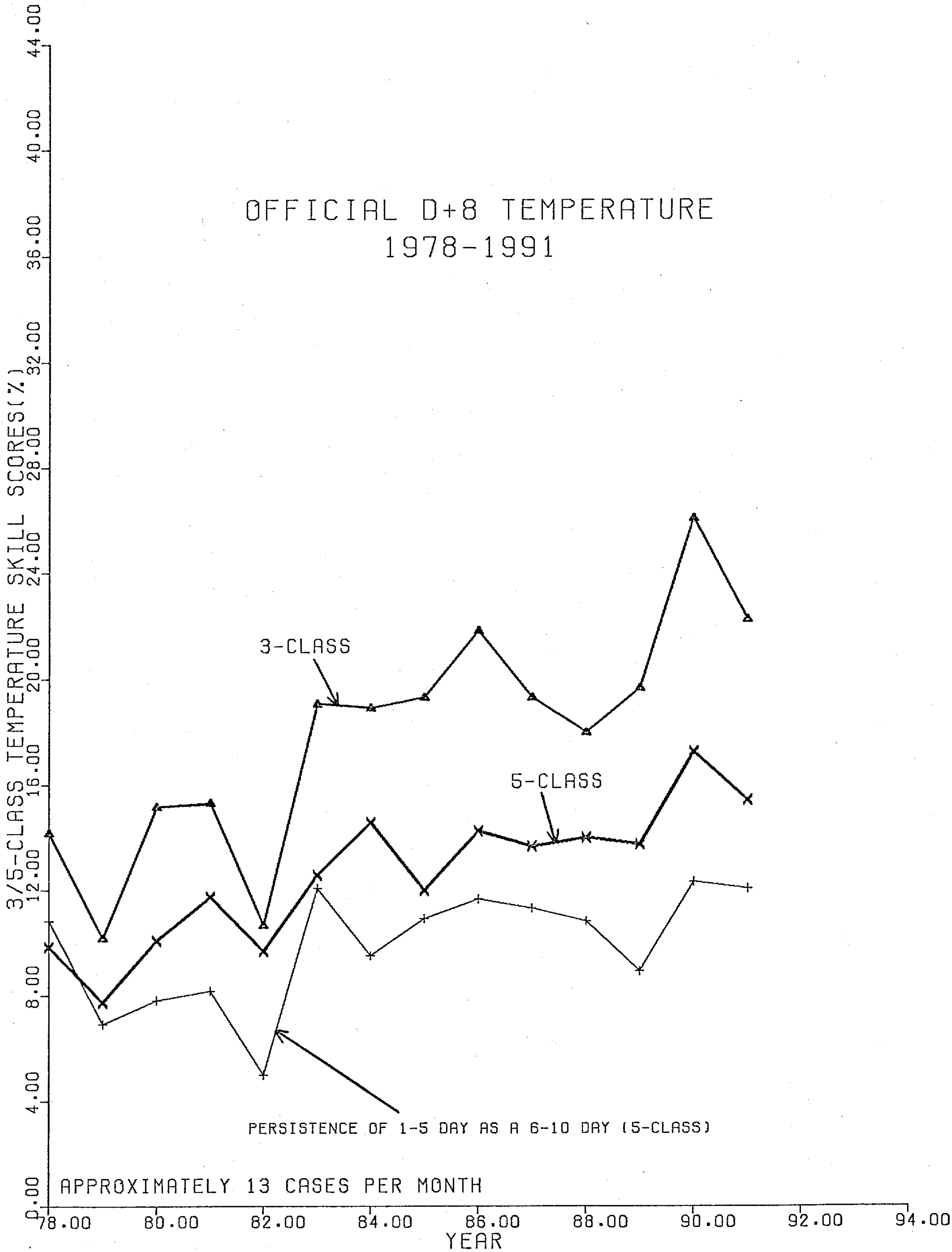
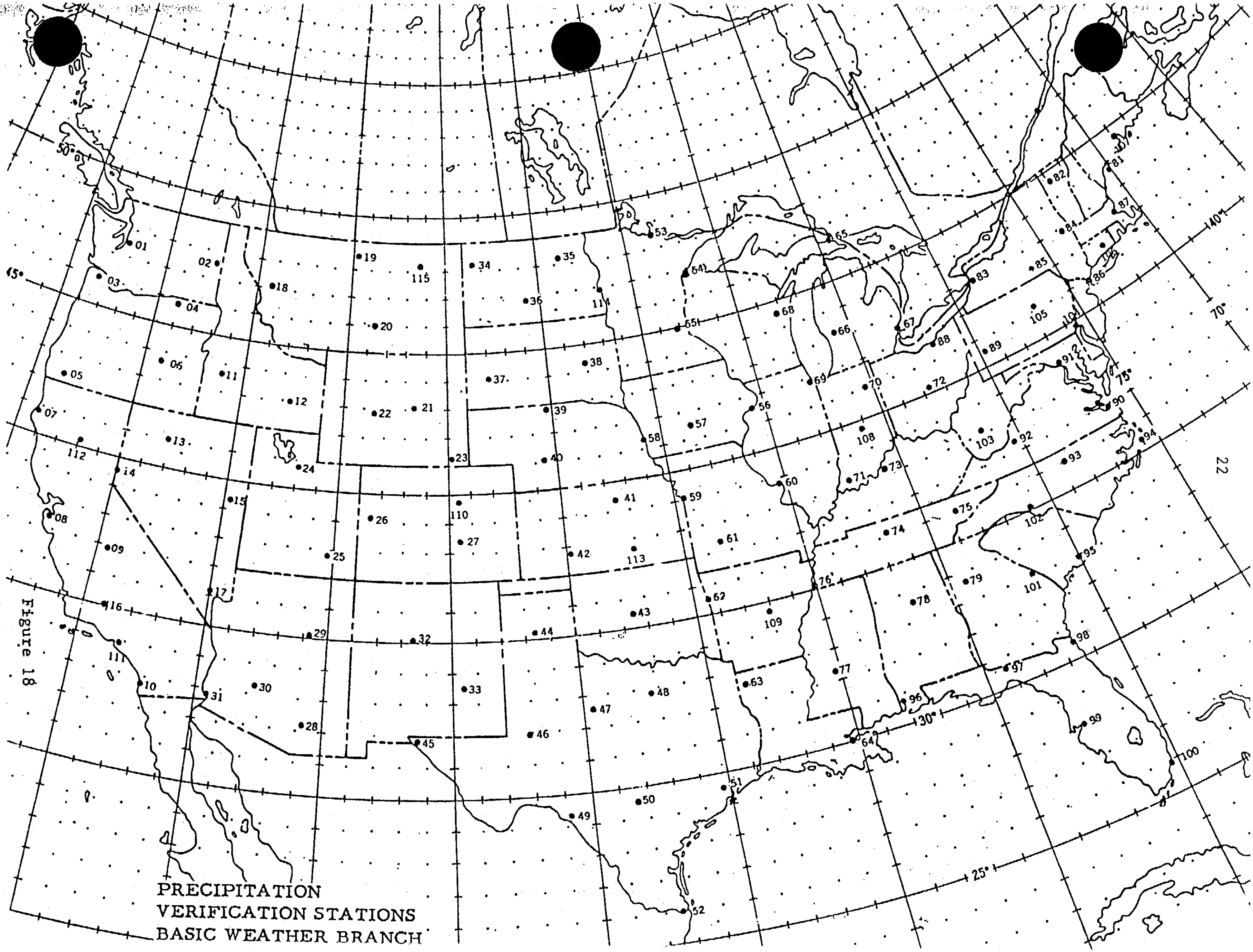

OFFICIAL D+8 TEMPERATURE
1978-1991
3/5-CLASS TEMPERATURE SKILL SCORES(%)
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
40.00
44.00
3-CLASS
5-CLASS
PERSISTENCE OF 1-5 DAY AS A 6-10 DAY (5-CLASS)
APPROXIMATELY 13 CASES PER MONTH
78.00
80.00
82.00
84.00
86.00
88.00
90.00
92.00
94.00
YEAR

Figure 17

SECTION 3

Man and Machine (NMC/NWP Model Guidance)

Precipitation Skill Scores



PRECIPITATION
VERIFICATION STATIONS
BASIC WEATHER BRANCH

Figure 18

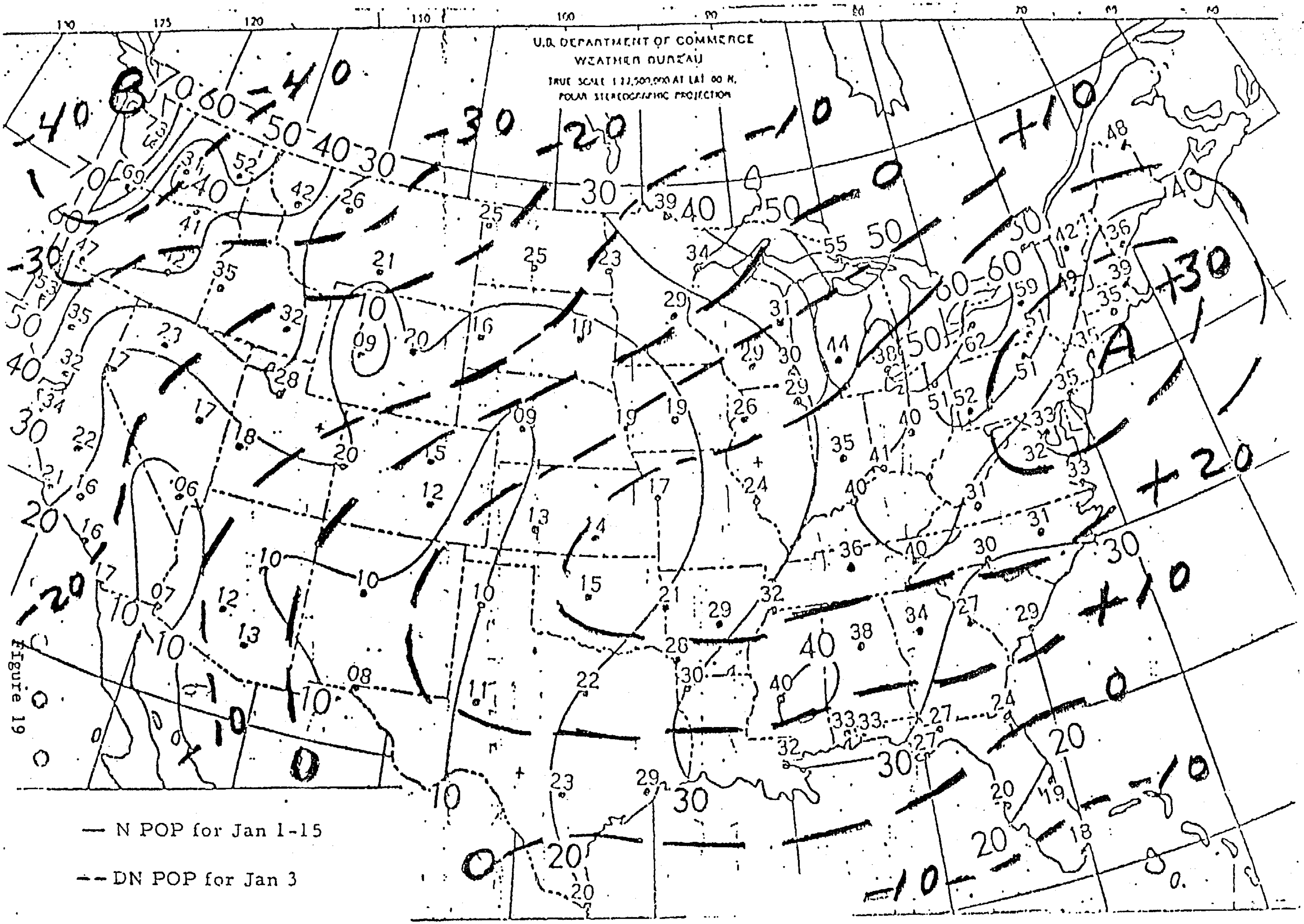

U.S. DEPARTMENT OF COMMERCE
WEATHER BUREAU
— N POP for Jan 1-15
-- DN POP for Jan 3

Figure 19

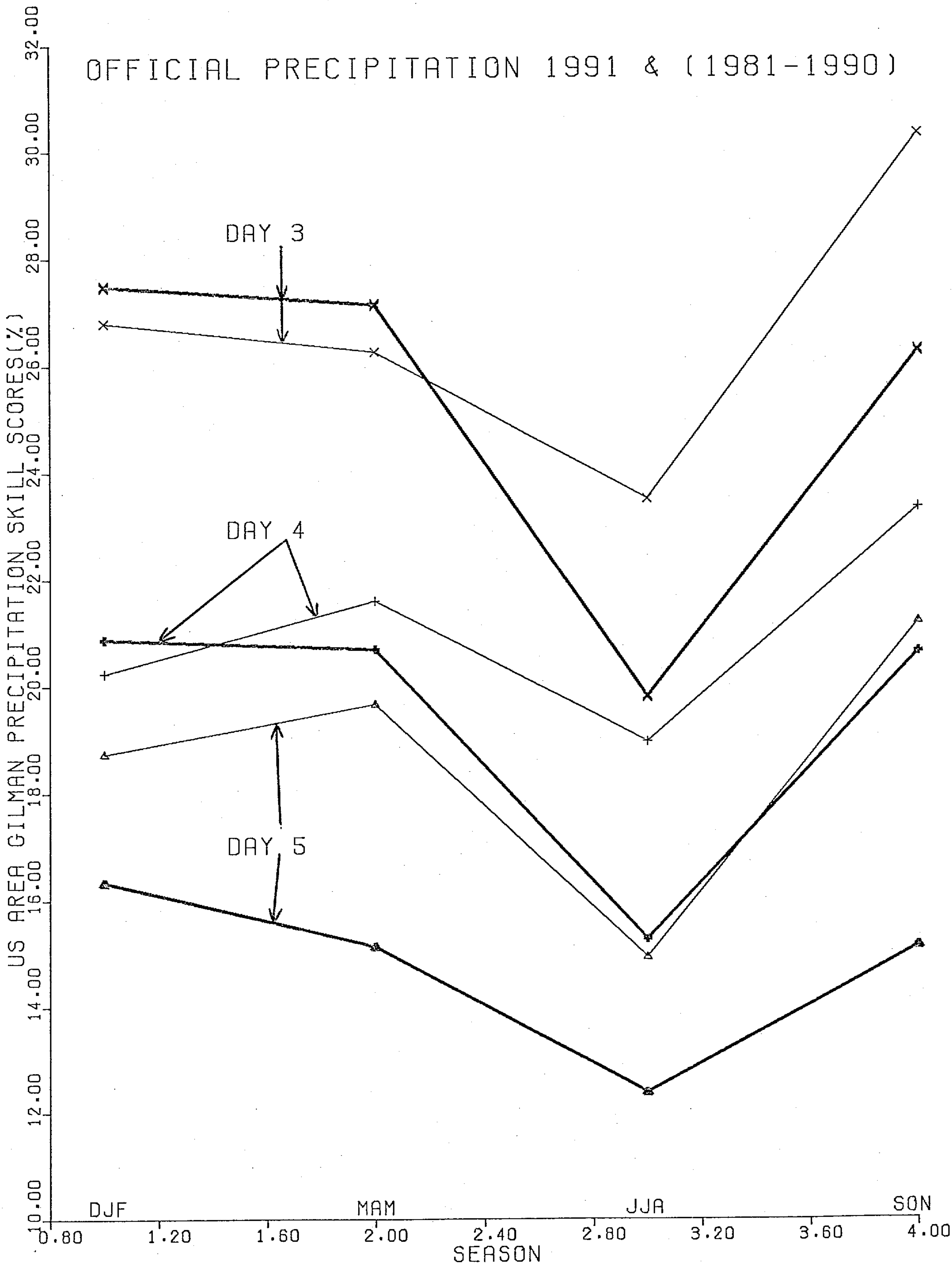

OFFICIAL PRECIPITATION 1991 & (1981-1990)
US AREA GILMAN PRECIPITATION SKILL SCORES(%)
DAY 3
DAY 4
DAY 5
DJF
MAM
JJA
SON
SEASON
0.80
1.20
1.60
2.00
2.40
2.80
3.20
3.60
4.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
24.00
26.00
28.00
30.00
32.00

Figure 20

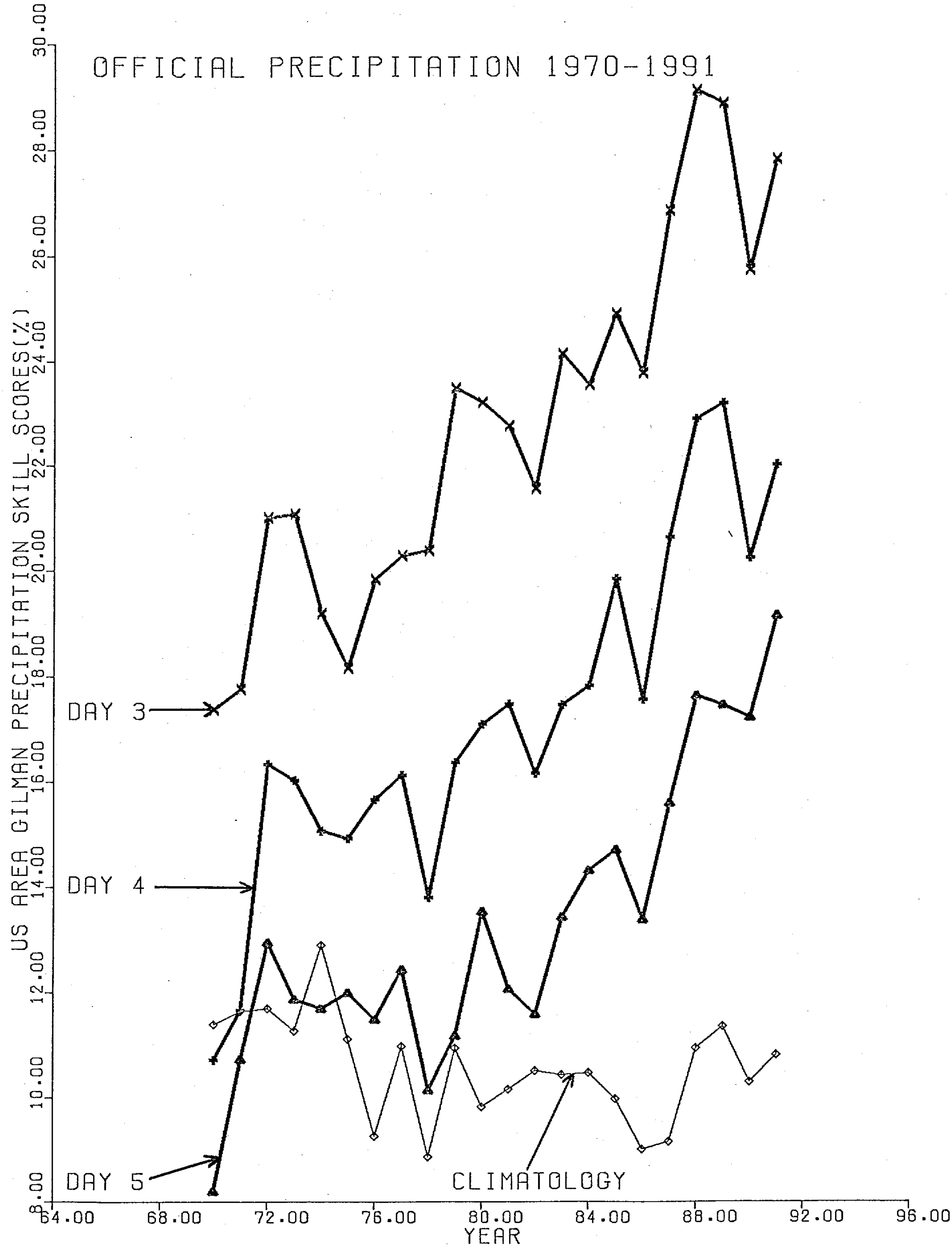

OFFICIAL PRECIPITATION 1970-1991
US AREA GILMAN PRECIPITATION SKILL SCORES(%)
30.00
28.00
26.00
24.00
22.00
20.00
18.00
16.00
14.00
12.00
10.00
8.00
DAY 3
DAY 4
DAY 5
CLIMATOLOGY
64.00
68.00
72.00
76.00
80.00
84.00
88.00
92.00
96.00
YEAR

Figure 21

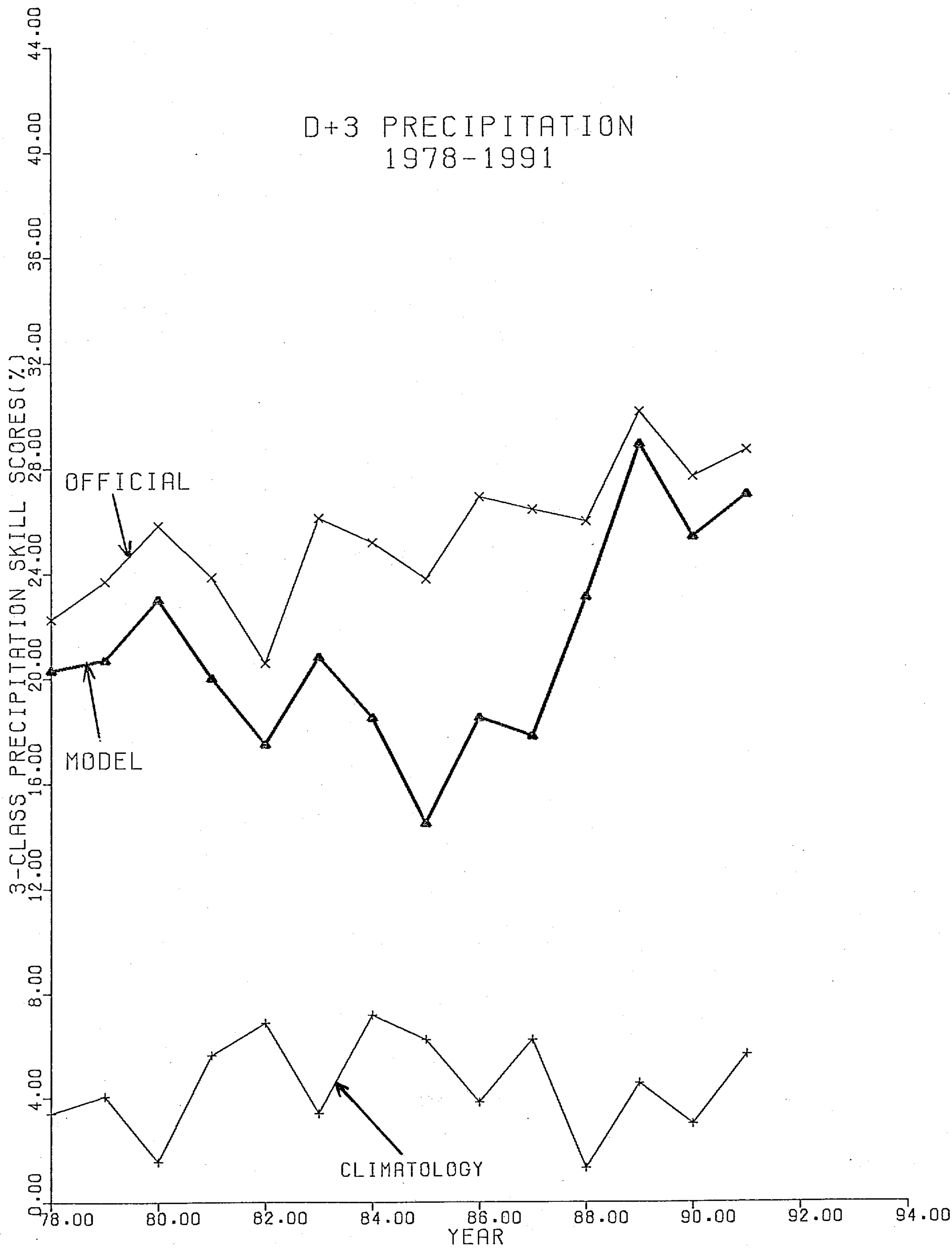

D+3 PRECIPITATION
1978-1991
OFFICIAL
MODEL
CLIMATOLOGY
3-CLASS PRECIPITATION SKILL SCORES(%)
YEAR
0.00
4.00
8.00
12.00
16.00
20.00
24.00
28.00
32.00
36.00
40.00
44.00
78.00
80.00
82.00
84.00
86.00
88.00
90.00
92.00
94.00

Figure 22

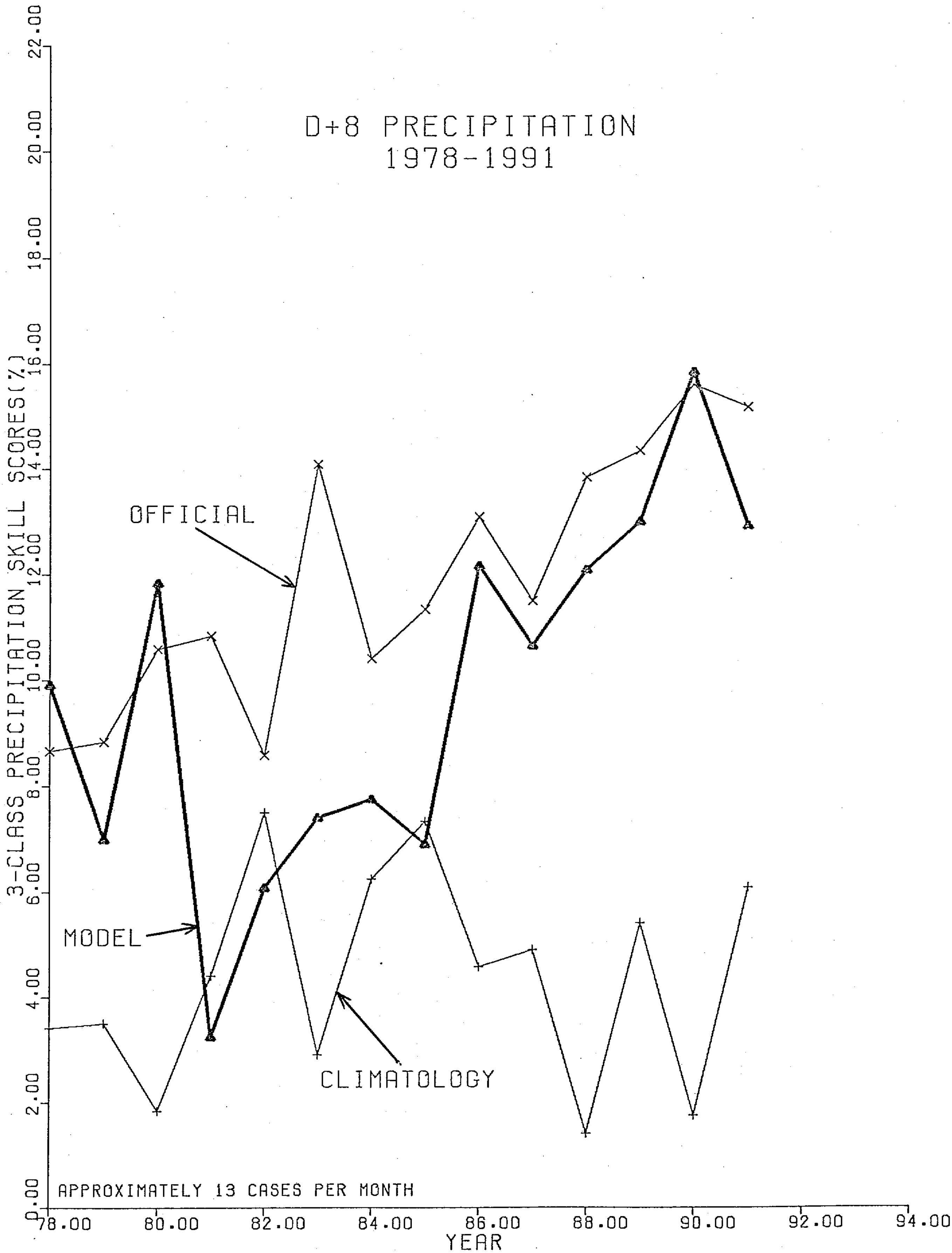

D+8 PRECIPITATION
1978-1991
3-CLASS PRECIPITATION SKILL SCORES(%)
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
22.00
OFFICIAL
MODEL
CLIMATOLOGY
APPROXIMATELY 13 CASES PER MONTH
78.00
80.00
82.00
84.00
86.00
88.00
90.00
92.00
94.00
YEAR

Figure 23

Appendix A

The standardized mean sea level pressure correlation score is used to determine the skill of the man and machine days 3, 4 and 5 mean sea level pressure forecasts. The correlation score is employed because the phasing instead of the intensity of systems primarily determines how well the various weather parameters can be forecast. The standardizing procedure prevents the contribution of the high variability (higher latitude) grid points from overwhelming the low variability grid points (lower latitude).

f = forecast mean sea level pressure at a grid point

o = observed mean sea level pressure at a grid point

σ = standard deviation at a grid point

n = normal mean sea level pressure at a grid point

$$F = \frac{f-n}{\sigma} \qquad O = \frac{o-n}{\sigma}$$

$\overline{F}$ = average standardized forecast across n grid points

$\overline{O}$ = average standardized observed across n grid points

$$\text{RMS } F = \sqrt{\overline{F^2}} \qquad \text{RMS } O = \sqrt{\overline{O^2}}$$

$$\text{RMS Error} = \sqrt{\overline{(F-O)^2}}$$

$$\text{Average Absolute Error} = \overline{|F-O|}$$

$$\text{Correlation} = \frac{\overline{FO} - \overline{F}\,\overline{O}}{\sqrt{(\overline{F^2} - \overline{F}^2)(\overline{O^2} - \overline{O}^2)}} \times 100$$

Since the normal mean sea level pressure is subtracted from the forecast/observed pressure at each grid point, it is assumed that the correlation of the normal to the observed is always zero. Therefore, any positive score is considered

APPENDIX A (cont'd)

to have skill over the normal. Some doubts have been raised about this assumption, however, and for the past 5 years the unstandardized correlation score also has been calculated. This procedure allows a correlation score to be computed for the normal. This score then is simply the correlation of the forecast to the observed mean sea-level pressure.

APPENDIX B

The 5 day mean temperature skill score is a generalization of the Heidke skill score where the expected values are derived from the observed temperature

$$\text{Heidke Skill} = \frac{C-E}{N-E}$$

C = total correct (hits)
N = total number of forecasts (61)
E = expected number of hits

The expected value is calculated as follows from the number of stations in each of the observed temperature categories:

E = 1/8 x Much Below + 1/8 x Much Above +
1/4 x Below + 1/4 x Above + 1/4 x Normal

The 5-day mean 3-class temperature skill score simply "lumps" together the much below with the below and the much above with the above. The expected (E) then is equal to 1/4 X Below + 1/4 X Normal + 1/4 Above.

Appendix C

The Gilman skill score is a generalization of the Heidke skill score where the expected values are derived from a randomized version of the precipitation forecast.

Heidke Skill = $\frac{C-E}{N-E}$

C = total correct (hits)

N = total number of forecasts (100)

E = expected number of hits

However, for a randomized forecast allowance must be made for stations having far different precipitation climate (N POP) across the United States. Therefore, to compute and score an expected chance forecast, climatology must be considered. The procedure for this is as follows:

First, the actual number of forecasts of precipitation are distributed randomly taking into account station climatology. The expected number of chance hits is then given by:

$$E = \sum^{N} (p_i r_i + (1 - p_i)(1 - r_i)) \text{ or}$$

$$E = 2\sum^{N} p_i r_i + N - \sum^{N} p_i - \sum^{N} r_i \quad (a)$$

where $r_i = 1$ for precipitation ($\geq .01$ inch) and 0 for no precipitaiton ($< .01$ inch). Now an expression for p_i, which is the probability that after the forecast precipitation events are redistributed randomly a forecast precipitation event will fall at point "i" is given approximately by $p_i \leq F\frac{a_i}{\sum^{N} a_i}$ (b). Here F = total number of forecasted precipitation events and a_i = climatic precipitation probability (N POP). This approximate value for p_i is most valid for small values of F and ($a_i / \sum a_i$) and is unstable at times. Because of this instability the less sophisticated but more stable Hughes skill score was developed.

Substituting the expression (b) into (a) gives $E = \frac{2F\sum^{N} a_i r_i}{\sum a_i} + N - F - R$, where E = the approximate expected value of a randomized forecast, R = total precipitation cases, and N = total number of stations. If the climatic probabilities are uniform ($a_i = a_j = a$), then the approximate value of E reduces to the standard Heidke value given by: $E = \frac{(N-F)(N-R)+FR}{N}$.

APPENDIX D

The 5-Day mean precipitation skill score is a generalization of the Heidke skill score where the expected values are derived from the observed precipitation:

$$\text{Heidke Skill} = \frac{C-E}{N-E}$$

C = total correct (hits)
N = total number of forecasts (100)
E = expected number of hits

For example, in January the number of stations in the area covered by the (NP/P), (NP/M/H) and (L/M/H) categories is 21, 28 and 51 respectively. The average value of the probability of NP for the stations in the (NP/P) area is 70% and 40% in the (NP/M/H) area. Now if (NP/L) is coded as 1, M as 2 and (P/H) as 3, then the number of stations expected to have coded value 1 thru 3 is as follows:

33% of (L/M/H) = 51 x .33 = 17 stations coded 1, 2, 3

40% of (NP/M/H) = 28 x .40 = 11 stations coded as 1 and 8.5 coded as 2,3

70% of (NP/P) = 21 x .70 = 14.7 stations coded as 1 and 6.3 coded as 3

Thus, code 1 = 17 + 11 + 14.7 = 42.7 stations
code 2 = 17 + 8.5 = 25.5 stations
code 3 = 17 + 8.5 + 6.3 = 31.8 stations
100.0 stations

Therefore, the expected value = .427a + .255b + .318c

where a, b and c are the number of coded values 1, 2 and 3 observed.