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

OFFICE NOTE 140

Humidity Analyses for Operational Prediction Models at the National Meteorological Center

Development Division

Part 1	LFM	Roland Chu
Part 2	PE	Dave Parrish

MARCH 1977

This is an informal unreviewed manuscript primarily intended for the exchange of information among NMC staff members.

PART I

A description of humidity analysis for the limited-area fine-mesh (LFM) model

Roland Chu

Development Division

1. Introduction

This paper gives a description of humidity analysis for the Limited-area Fine-mesh (LFM) model, an operational forecast model currently employed at the National Meteorological Center, NWS, NOAA. The LFM model analyzes mean relative humidity (RH) in three σ-coordinate layers that are closest to the Earth's surface. They are: the boundary layer with a fixed pressure depth of 50 mbs, and the first and second tropospheric layers whose pressure depth (Δp) is given by

$$\Delta p = 1/3(p^* - p^{**} - 50) \qquad (1)$$

where p^* is the pressure at surface, and p^{**} is the pressure at the level of tropopause. The vertical structure of the LFM model is shown in Figure 1.

In the horizontal domain, there are 3,021 (53 x 57) grid points, covering the North American continent and its adjacent oceans. (The grid spacing is 190.5 km at 60°N.) An approximation of the LFM grid is shown in Figure 2.

2. The Data

Radiosonde observations are the primary data used in the humidity analysis. Where radiosonde data are unavailable or non-existent, estimates of mean RH in all three layers are obtained from weather and cloud observations that are made by surface stations and ships. These estimates constitute a supplementary or secondary source of data. In addition, over certain oceanic areas satellite data are used to derive RH estimates in all three layers. These satellite bogus data are presently considered to be secondary data, which are of the same reliability and importance as the RH estimates derived from surface observations. The various types of data are now discussed below:

a. Radiosonde data

Radiosonde data are obtained from the disk file "ADPUPA" and are first subject to a series of checks commonly employed in an objective analysis program. Some of the checks are performed to make sure that:

- at each reporting station, the reports are arranged in proper order according to the observed pressure at each level, and that no level is duplicated;
- at each pressure level, both temperature and dew point temperature are available so that RH can be computed at that level; if either one is missing at any level, that level will be ignored;
- station elevation is of reasonable value and is within 300 meters of the standard height computed from the pressure at the first reporting level;

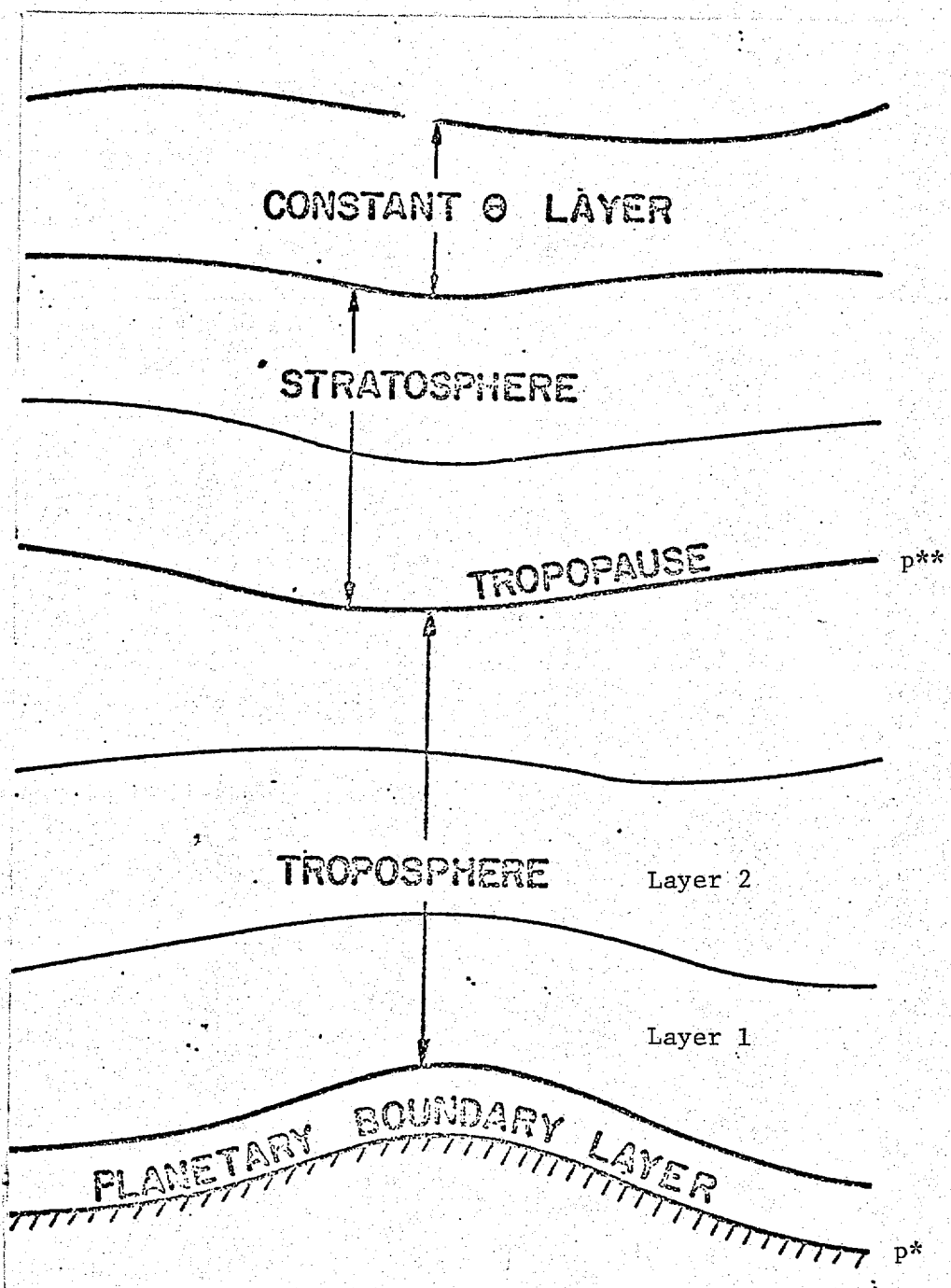


Figure 1. The vertical structure of LFM model at NMC.

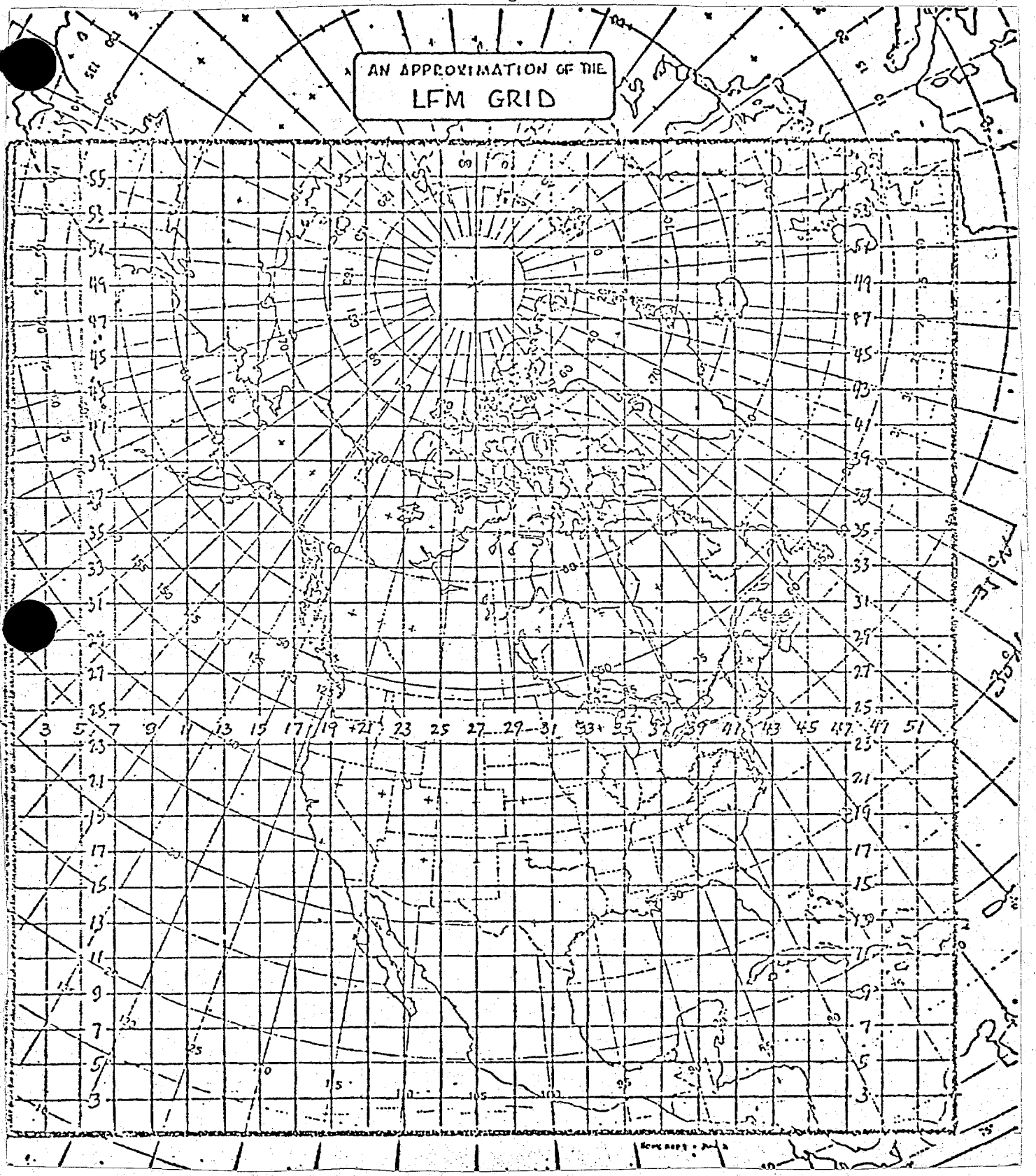


Figure 2. An approximation of the LFM grid.

. at each reporting station, the pressure at the first reporting level is of reasonable value, which is between 700 and 1080 mb;

. at each station, there are reports for at least one level.

For all radiosonde reports that have passed the checks, a series of computations are carried out to arrive at a mean value of relative humidity in each of the three layers. For each of the three layers, mean values of temperature and dew point temperature for each of the sub-layers are computed first:

$$\overline{T} = \frac{T_k + T_{k-1}}{2} \tag{2}$$

$$\overline{T}_d = \frac{T_{d_k} + T_{d_{k-1}}}{2} \tag{3}$$

where T_k and T_{k-1} are temperature at two adjacent levels with indices of k and k-1 which form a sublayer, and $\overline{T}$ is the mean value for the sublayer. The mean dew point temperature $\overline{T}_d$ is computed in a similar manner with formula (3).

From $\overline{T}$ and $\overline{T}_d$ mean values of vapor pressure and saturation vapor pressure, respectively, for each sublayer are computed:

$$\overline{e} = 10^{(8.4051 - 2353/\overline{T}_d} \tag{4}$$

$$\overline{e}_s = 10^{(8.4051 - 2353/\overline{T})} \tag{5}$$

where $\overline{e}$ and $\overline{e}_s$ are mean vapor pressure and saturation vapor pressure, respectively. The ratio of $\overline{e}$ to $\overline{e}_s$ gives the mean value of RH for the sub-layer:

$$\overline{RH} = \overline{e}/\overline{e}_s \tag{6}$$

where $\overline{RH}$ is the mean RH.

The computed $\overline{RH}$ for all sublayers are then used to compute the mean value of the boundary layer, and the two tropospheric layers:

$$\overline{RH}_\sigma = \frac{1}{\ell n\ p_1 - \ell n\ p_n} \sum_{k-1}^{n-1} \overline{RH}_k\ \ell n\left(\frac{p_k}{p_{k+1}}\right) \tag{7}$$

where k is the index for sublayer, $\overline{RH}_k$ is the mean RH for each sublayer, n is total number of sublayers, and $\overline{RH}_\sigma$ is the mean value for each of the three sigma layers.

It is noted that data at both mandatory and significant levels can be used in the computations of mean RH in the sigma layers. However, significant-level data are usually not available at analysis time.

b. Surface weather and cloud observations

As mentioned earlier, estimates of mean RH in all three sigma layers are derived from surface observations, which are obtained from the data file "ADPSFC." The input data used to derive the estimates are these observations taken at surface stations: present weather, fractions of sky cover by low, middle, and high clouds. From the reported present weather codes, relative humidity values (RHWW) are assigned according to Table 1.

```
      DATA (WWRH(I), I=1,100) /
C                    PRESENT WEATHER TABLE (RH)
C        0     1     2     3     4     5     6     7     8     9
C
     0  65.,  65.,  65.,  65.,  65.,  65.,  50.,  50.,  50.,  55.,
     1  90.,  90.,  90.,  85.,  90.,  95.,  95.,  95.,  90.,  90.,
     2  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,
     3  50.,  50.,  50.,  50.,  50.,  50.,  60.,  60.,  60.,  60.,
     4  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,  90.,
     5  95.,  99.,  99.,  99.,  99.,  99.,  95.,  99.,  95.,  99.,
     6  95.,  99.,  99.,  99.,  99.,  99.,  95.,  99.,  95.,  99.,
     7  95.,  99.,  99.,  99.,  99.,  99.,  90.,  90.,  90.,  90.,
     8  95.,  99.,  99.,  95.,  99.,  95.,  99.,  95.,  99.,  99.,
     9  99.,  95.,  99.,  95.,  99.,  95.,  99.,  99.,  95.,  99./
```

Table 1. Relative humidity assignments for various present weather codes used in synoptic surface reports.

Relative humidity values are assigned according to fractions of sky cover by the following formula:

$$\overline{RH}(\%) = M - A \cos\left(\frac{\pi}{8} OKTAS\right) \qquad (8)$$

where OKTAS is sky cover by low, middle, or high clouds in eighths; M and A are constants which have different values for low, middle, and high clouds. Figure 3 presents the graphical expression for formula (8). For $\overline{RH}$ in Layer 2 (RH 2) inferred from middle cloud, M has a value of 60 and A of 15.

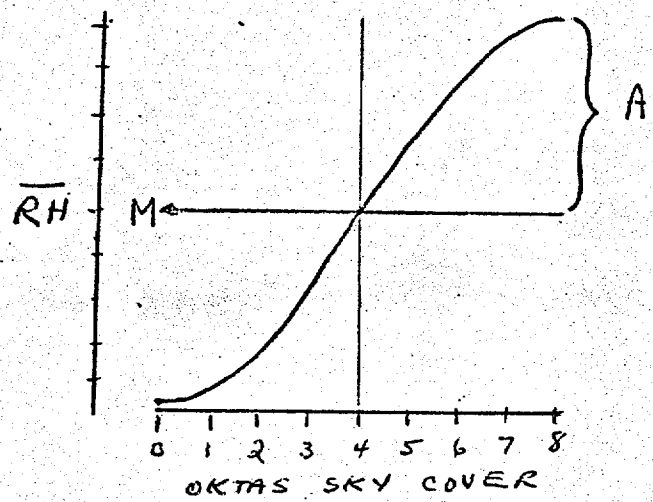


Figure 3. Graphical expression for $\overline{RH} = M - A \cos(\frac{\pi}{8} \text{OKTAS})$

When the existence of low cloud is reported, the additional information of height of cloud base is important. If the cloud base height is 2,000 ft. or higher, only $\overline{RH}$ in Layer 1 (RH 1) is inferred from the low cloud amount with M = 75, A = 15. However, when the cloud base height is less than 2,000 ft, both $\overline{RH}$ in Layer 1 (RH 1) and that in the boundary layer (RH BL) are derived from the low cloud amount through Formula (8). In this case, different values of M and A are used; for RH1, M = 70, A = 10; and for RHBL, M = 79, A = 19.

The temperature, dewpoint temperature and pressure observed at a surface station are used to calculate a surface relative humidity (RHG). The final inferred $\overline{RH}$ in the boundary layer is then the average of RHG, RHWW (RH inferred from the observed present weather) and RHBL (RH inferred from the amount of low cloud with its base height less than 2,000 ft). The final inferred $\overline{RH}$ in Layer 1 is the average of RHWW and RH1 (RH inferred from the amount of low cloud, the value of which depends on whether the cloud base is above or below 2,000 ft). No averaging is carried out in Layer 2, where the $\overline{RH}$ is just that inferred from middle cloud amount (RH2).

c. Satellite data

Satellite imagery and mapped VTPR channel 7 data are used to derive estimated mean values of relative humidity (RH) for three sigma layers over areas of the eastern North Pacific Ocean (170°E West Coast) and the Gulf of Mexico.

A numerical code from one through ten is used to represent eight different vertical distributions of RH, depending upon the interpretation of satellite data. This numerical code and its location (i, j points) is fed into a computer program which converts the code into a standard synoptic ship code. The ship code for each bogus point becomes the bogus data of "SFCBOG" file.

The bogus ship codes of "SFCBOG" file are read by the LFM moisture analysis program as if they were regular ship codes or data. The analysis program then generates bogus vertical RH profiles accordingly. In Table 2 the eight different vertical distributions of RH and their corresponding bogus surface reports, which are generated by the numerical codes from one to ten, are illustrated. It is noted that numerical codes 8, 9, and 10 generate one identical vertical distribution of RH, since the corresponding bogus surface reports have the same weather code and same observation on low and middle clouds.

		Sigma layer		
	Nddff VVwwW $N_hC_LhC_MC_H$	Boundary	Tropospheric 1	Tropospheric 2
(1)	8XXXX XX62X 7747X	85%	89%	75%
(2)	8XXXX XX80X 7747X	84%	87%	75%
(3)	6XXXX XX80X 3347X	76%	80%	71%
(4)	4XXXX XX02X 4092X	62%	49%	60%
(5)	8XXXX XX02X 8092X	62%	49%	75%
(6)	8XXXX XX60X 8740X	84%	87%	38%
(7)	4XXXX XX02X 4740X	68%	67%	38%
(8)	0XXXX XX00X 00900	62%	49%	38%
(9)	8XXXX XX00X 00902	62%	49%	38%
(10)	0XXXX XX00X 00901	62%	49%	38%

Table 2. RH estimates from satellite bogus surface reports in file=SFCBOG for use in the LFM analysis program.

It is also noted that the RH values in the boundary layer have a limited range of 62% to 85%. This range of values does not cover the highly moist conditions reflected in bogus surface reports under codes 1, 2, and 6, or the rather dry conditions intended by bogus surface reports under codes 4, 5, 8, 9, and 10. Changes in the analysis program are required, if the estimated RH values in the boundary layer are to be in closer agreement with the interpretation from satellite observations.

d. Special treatment of all grid points south of 30°N.

The "observed" mean RH at all grid points south of 30°N are adjusted according to the formula

$$RH = \hat{RH} \times (2 \sin \phi) \qquad (9)$$

where $\hat{RH}$ is the original value, ϕ is latitude, and RH is the final assigned value.

3. The First-guess (or background) Field

The first-guess (or background) field (G) for LFM moisture analysis is constructed from the 12-hr forecast field (F), climatology (C) according to the following algorithm:

(1) North of 35°N latitude $G = F$

(2) Between 35°N and 20°N $G = W_1 F + (1 - W_1)C$
where $W_1 = (\sin\phi - \sin 20^\circ)/(\sin 35^\circ - \sin 20^\circ)$
ϕ is latitude

(3) Between 20°N and 15°N $G = W_2 C + (1 - W_2)D$
where $W_2 = (\sin\phi - \sin 15^\circ)/(\sin 20^\circ - \sin 15^\circ)$
$D = 40\%$

(4) South of 15°N $G = D = 40\%$

The resulting background field is then reduced by a factor SATRH, which varies seasonably between 0.90 to 0.96 in the following fashion:

(1) Nov. 21-Mar. 20, SATRH = 0.96;

(2) Mar. 21-May 20, SATRH decreases from 0.96 to 0.90 linearly;

(3) May 21-Sept.20, SATRH = 0.90;

(4) Sept. 21-Nov. 20, SATRH increases from 0.90 to 0.96 linearly.

4. Procedure

The major steps employed in LFM humidity analysis are illustrated in a simple flow chart shown in Figure 4. First, the various data (RH computed from radiosonde observations, and that estimated from surface and satellite observations), and first-guess (or background) values of RH are obtained, and put on the 53 x 57 grid points in all three sigma layers. A rather simple and straight-forward scheme is used to transfer both radiosonde data and RH estimates from observation points to grid points. The principle is that any one observation is allowed to influence its four adjacent grid points only. Furthermore, if there is only one observation within a grid box of four points, all four points would take on the same value as that of the observation. However, if more than one observation are located in a grid box a distance-weighted average would be assigned to the four grid points in that box.

As mentioned earlier, radiosonde observations are the primary data which take precedence over RH estimates from surface and satellite bogus data. Therefore, RH estimates from surface and satellite bogus data are used to fill only those grid points that have not been affected by computed RH values from radiosonde data. After data of all types have been placed on grid points, those at all points south of 30°N are subject to the reduction treatment of formula (9).

After the utilization of data of all types, if there are still grid points unaffected, first-guess (or background) RH values described in Section 3 are then used to fill those points. The RH fields in all three sigma layers are then smoothed and desmoothed to produce the final "analysis."

5. Some Sample Analyzed Fields

Some sample analyzed fields of RH are presented in Figures 5 through 7 for illustration. In the figures, RH fields in the boundary layer during various stages of analysis for the case of 00 GMT 12 January 1977 are shown. The analyzed area shown in each figure is that of a rectangle over the conterminous United States, and the isopleths are those of constant RH labeled in three (30%), five (50%), seven (70%), and nine (90%) intervals.

The analyzed field with RH values computed from radiosonde data, which are plotted at all reporting stations, is shown in Figure 5. The shaded areas are those within which grid points have not been influenced by data. Among these, there are areas where radiosonde stations do not exist: the Gulf of Mexico; areas in Mexico bordering Texas, Arizona, and Southern California; a part of Southern Ontario in Canada; parts of Utah and Arizona; and Eastern Iowa. There are other areas which are not covered by data due

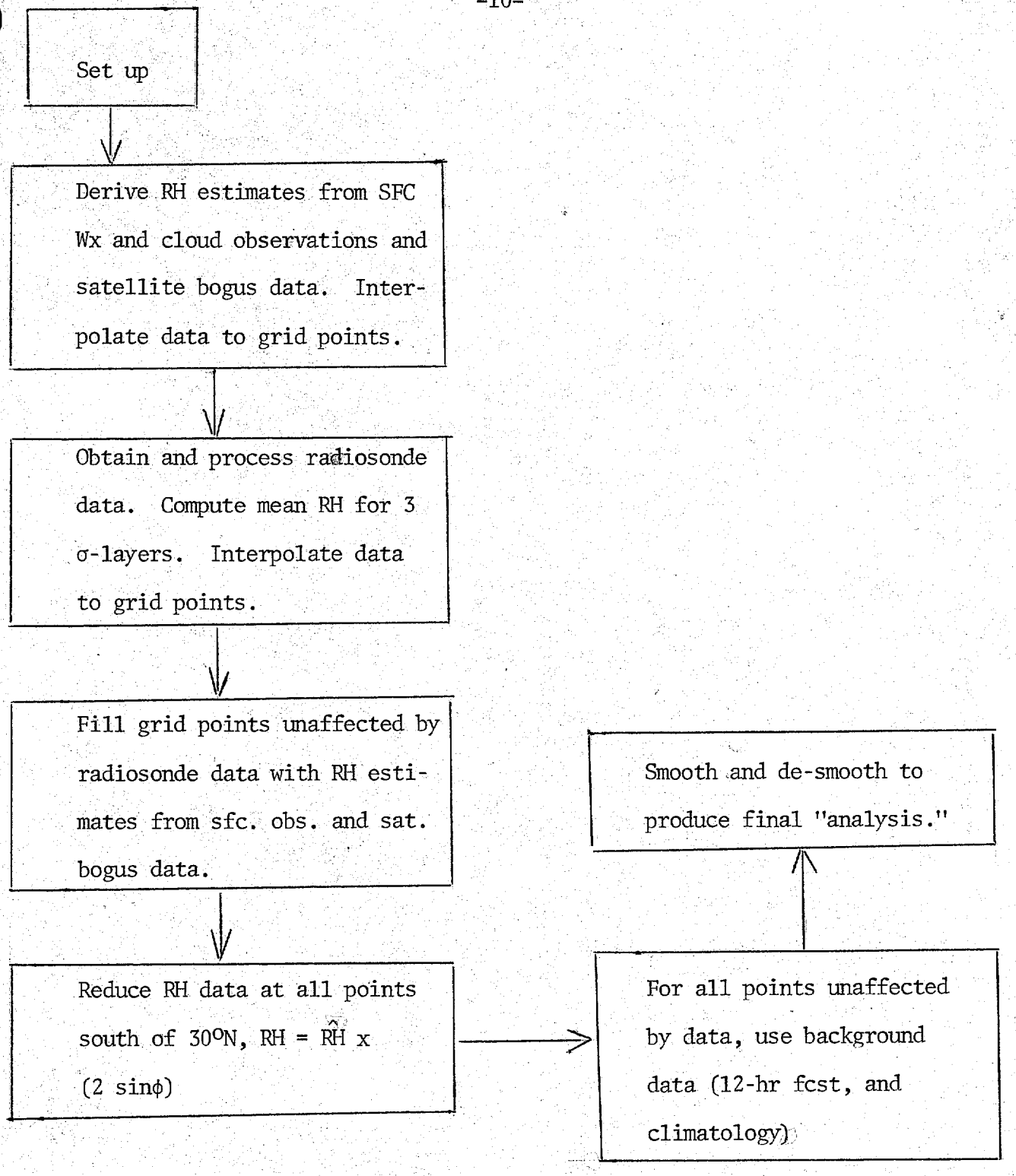


Figure 4. Major steps in LFM humidity analysis

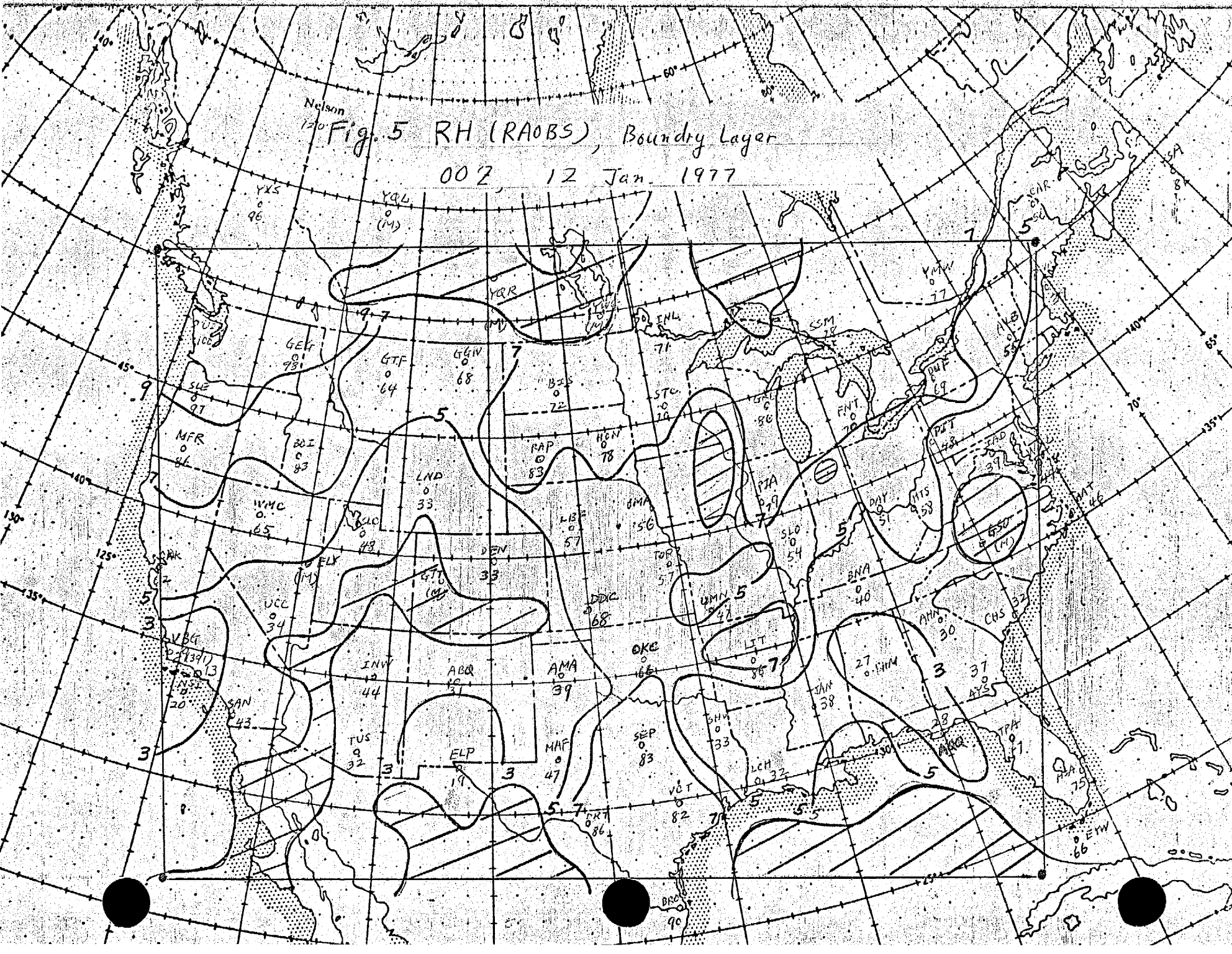
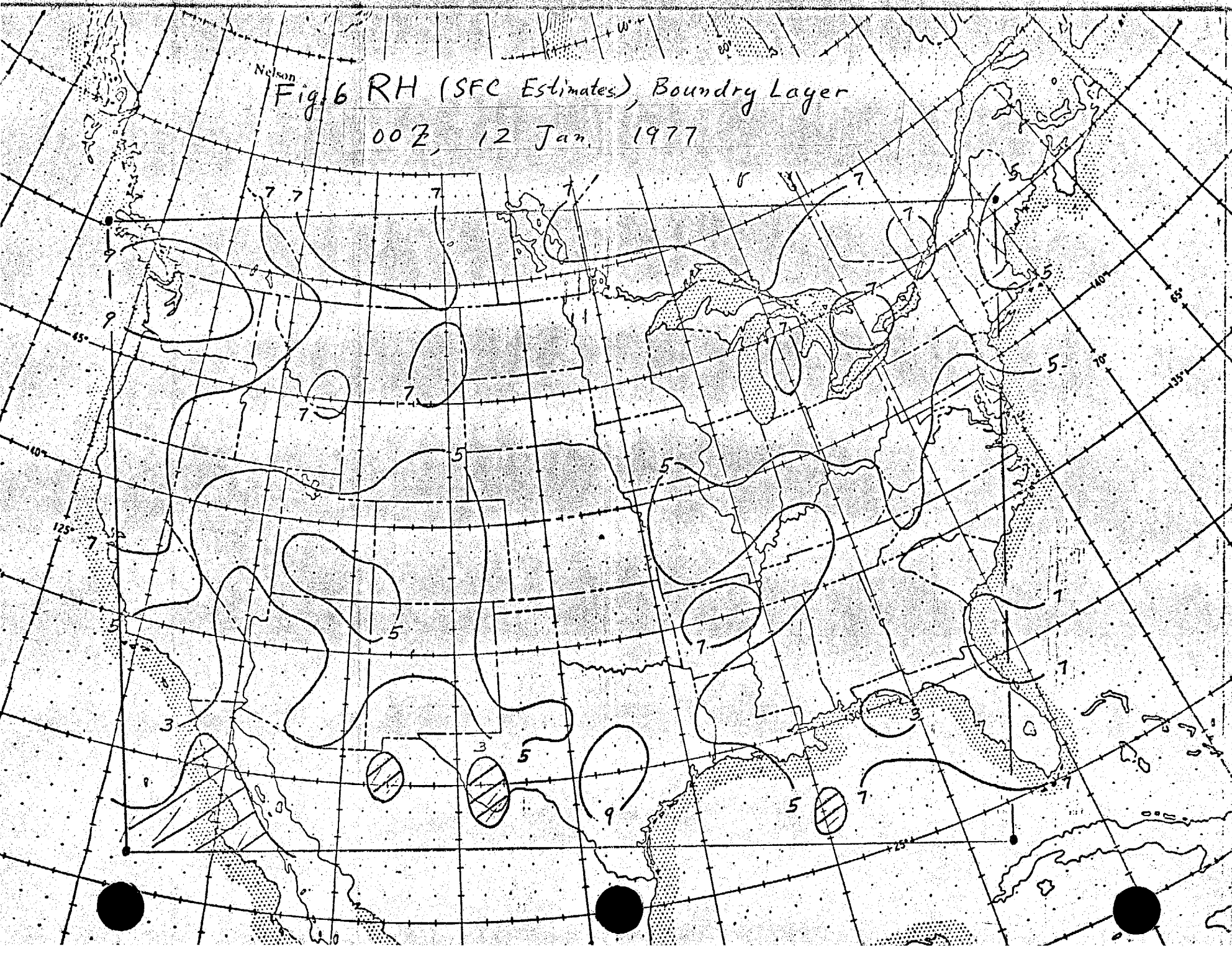
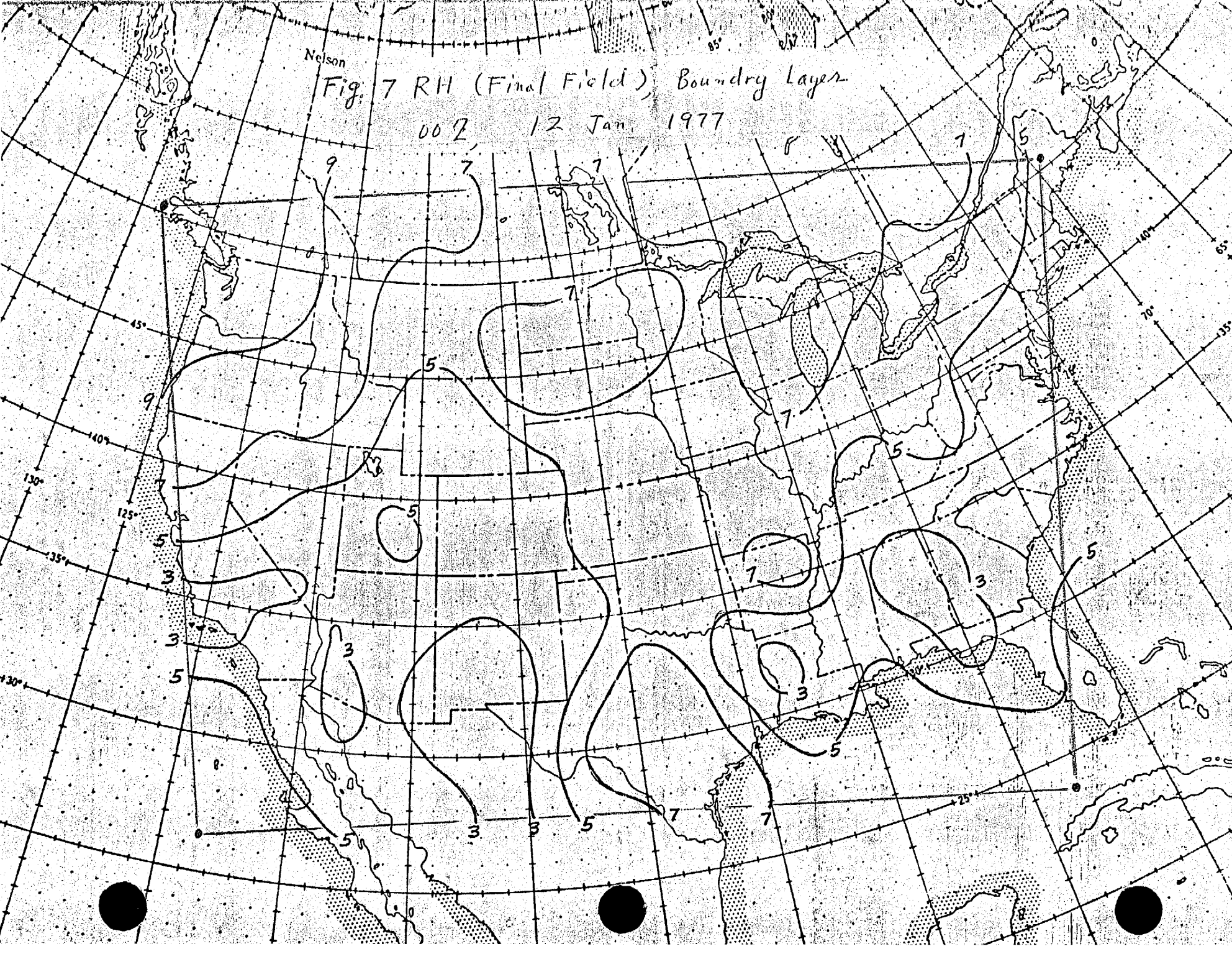


Fig. 5 RH (RAOBS), Boundry Layer
00Z, 12 Jan. 1977
Nelson
YXS 96
YQL (M)
YXR (M)
YMW 77
YCR 50
INL 71
SSM
GEG 98
GTF 64
GGW 68
BIS 72
STC 70
GFL 86
FNT 70
PIT 48
DUP 69
SLE 97
MFR 86
BOI 93
RAP 83
HON 78
LND 33
PIA 79
DAY 51
MTS 58
GSO (M)
MAT 46
VMC 65
BLU 49
LBF 57
OMA 56
SLO 54
DEN 33
TOP 57
ELK (M)
GJT (M)
VCC 34
DDC 68
ENA 40
AHN 30
CHS 32
UMN 46
OKE 66
LIT 86
VBG (391) 13
INW 44
ABQ 16 31
AMA 39
FAM 27
JAN 38
AYS 37
NSI 20
SAN 43
TUS 32
ELP 19
MAF 47
SEP 83
SHV 33
LCH 32
TPA 71
AMQ 28
VCT 82
ORT 86
HAT 75
ETW 66
BRO 90


Nelson
Fig. 6 RH (SFC Estimates), Boundry Layer
00Z, 12 Jan. 1977


Nelson
Fig. 7 RH (Final Field), Boundry Layer
00 Z 12 Jan 1977

to missing reports from stations in the areas: area over parts of North Carolina and Virginia due to missing reports from Greensboro, N.C. (GSO); area over southwestern Colorado and eastern Utah because of missing reports from Grand Junction, Colorado (GJT); and area of southern Saskatchewan and southern Manitoba in Canada, where reports from Regina, Sask. (YQ) and Winnipeg, Man. (YWG) were missing.

The analyzed RH field in the boundary layer with RH estimates from surface observations and satellite bogus data is presented in Figure 6. It is noted that except for a small area in the Gulf of Mexico, and three small areas in northern Mexico, the analyzed area is now covered with data.

The final field of RH analysis in the boundary layer is shown in Figure 7. Recall that radiosonde data (Fig. 5) are the primary data used, RH estimates from surface and satellite data (Fig. 6) fill those grid points where radiosonde data are lacking, and then first-guess (or background) RH values are used to fill the points still uncovered by any data. Also recall that the combined data field is finally smoothed and desmoothed to produce the final field in Fig. 7. The first-guess (or background) field is not shown here, since the area of analysis is well covered by data of various types, and the first-guess field makes very little contribution in this case.

6. Reference

Desmarais, A. J., "Notes on mean relative humidity analyses in sigma layers for the PE model," unpublished manuscript, NMC, NOAA, July 1971.

PART II

A description of the Flattery humidity analysis for the six-layer primitive equation (PE) model

David Parrish

Development Division

Part II. A description of the Flattery humidity analysis

Although the Flattery analysis system was designed primarily around height and wind analyses, it has also been adapted for analysis of relative humidity. The procedure is almost identical with the height-wind analysis, differing only in that it is univariate. Since the details of the height wind analysis have been described elsewhere (Technical Procedures Bulletin #105), we will give only a brief description of the basic method. More attention will be focused on the nature of the data that is used, the treatment in the vertical, and the quality of the resulting analyzed fields.

The relative humidity is represented by a series expansion, such as

$$RH(\lambda,\phi,p) = \sum_{n=1}^{3} \sum_{\ell=0}^{24} \sum_{m=1}^{24} \left(a_{mn}^{\ell} \cos\ell\lambda + b_{mn}^{\ell} \sin\ell\lambda\right) H_m^{\ell}(\phi) V_n(p) \qquad (1)$$

where λ,ϕ,p are east longitude, north latitude, and pressure respectively, $H_m^{\ell}(\phi)$ are the Hough height functions, $V_n(p)$ are empirical orthogonal functions, and a_{mn}^{ℓ}, b_{mn}^{ℓ} are the coefficients of the series. The coefficients are obtained by minimizing, in a rough sense, the root-mean-square difference between $RH(\lambda,\phi,p)$ and the available data. As in the height-wind system, this is accomplished by making adjustments to a forecast first guess set of a's and b's valid at the analysis time; long waves (small m, ℓ, and n in equation (1)) are altered first, and progressively shorter scales are corrected with each iteration. This technique is analogous to the variation of scan radii used in the Cressman analysis. It is an attempt to compensate for irregular data distribution.

Once the coefficients have been obtained, the series (1) is evaluated on a 2.5^{o} latitude-longitude grid at the six mandatory levels 1000, 850, 700, 500, 400, and 300 mb. These are the principal output fields, from which others are derived. The 2.5^{o} fields go directly to the global forecast model. To obtain input RH fields for the hemispheric polar stereographic model, these latitude-longitude fields are interpolated bi-quadratically to the 65 x 65 polar stereographic grid. Tropopause and surface pressures are obtained from Flattery temperature and height analyses also interpolated to the 65 x 65 grid. The tropopause surface is determined using a routine written by Gustafson and later modified by McDonell (see NMC Tech. Memo #33). Sigma layer humidities are obtained by linear interpolation in pressure from the mandatory levels to the centers (in pressure) of the three lowest sigma layers (one boundary and two troposphere). Finally, these are smoothed-desmoothed twice.

The primary data sources for this analysis are radiosonde, surface, and bogus (from satellite cloud photographs). Only mandatory level radiosonde data are used, and surface data are inserted at their reported surface pressure values. Satellite bogus data enter as present weather code groups, which are converted by a table lookup to one of ten pre-determined relative humidity profiles. Present weather groups available in surface reports are ignored, so surface reports are not extended up explicitly, as they are in the LFM analysis. However, there is an extension in the vertical, implied by the vertical functions $V_n(p)$, which we will consider further on.

The vertical functions are the first three eigenvectors of a 6 x 6 covariance matrix obtained from all available radiosonde data at the six mandatory levels, 1000-300 mb. When the analysis was first put into routine operation, the functions were recomputed once every 12 hours from current data. However, we found it necessary to fix the functions because every time they were recomputed, large random changes would result in the guess field. This caused problems with the satellite moisture bogus operation, which relies heavily on the first guess moisture field.

The functions currently in use are illustrated in Figure 1. Even under the best circumstances (e.g., no horizontal variation to contend with), these three functions are limited in their ability to represent moisture soundings. Figure 2 gives three examples of exact least squares fits to radiosonde soundings. We note that the overall character of the soundings is preserved, but details are averaged out, with errors of up to 20% occurring. In actual analyses, the rms error in RH varies from 10-15%.

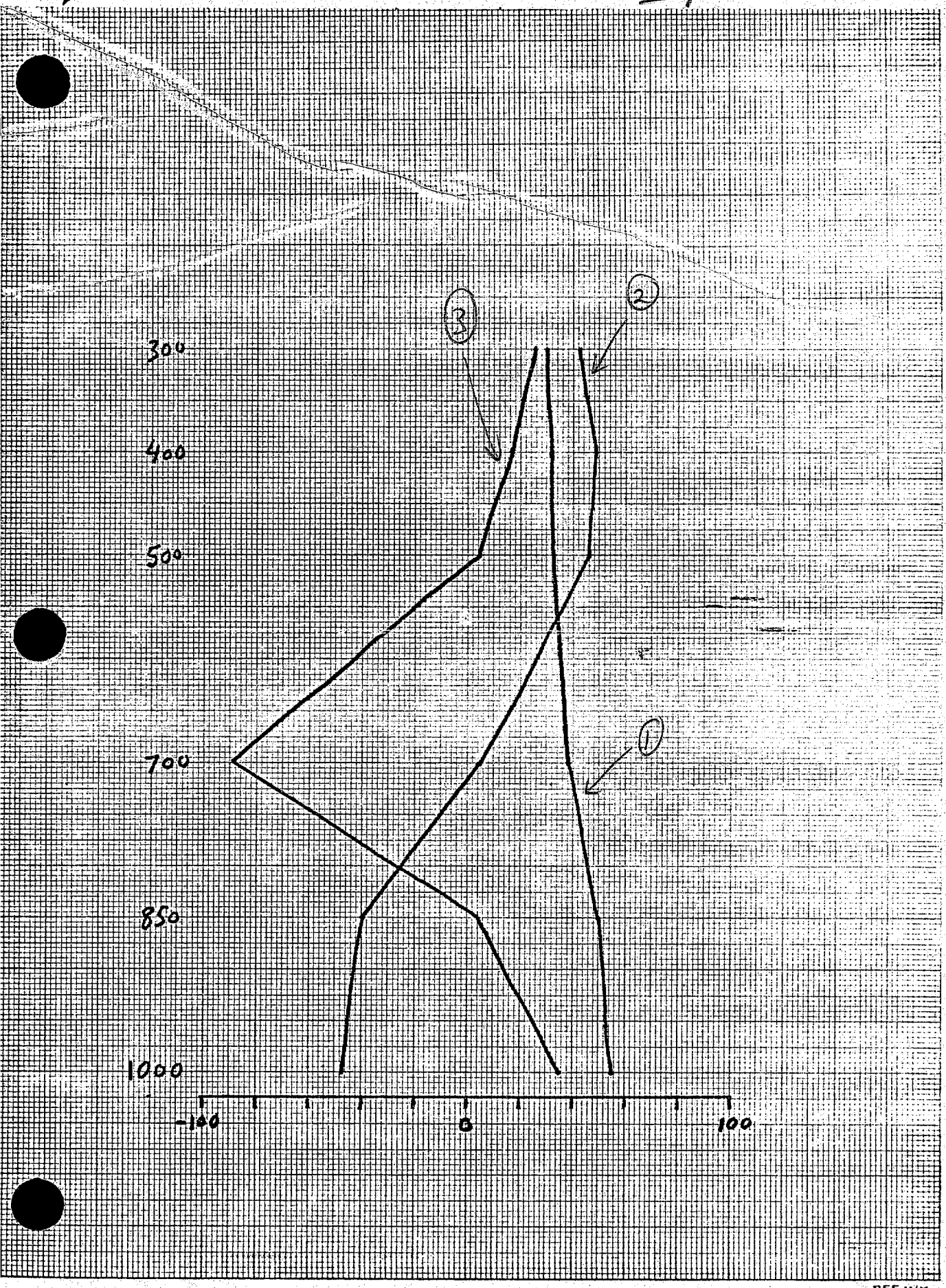
To see how the Flattery analysis compares with the only other one currently available, the LFM RH analysis, Figures 3(a) and 3(b) show Flattery and LFM analyses side-by-side for the boundary layer over the eastern United States. Some radiosonde data (as seen by the LFM analysis code) are plotted on Figure 3(a) for comparison. We note that, as in the vertical, the Flattery analysis only picks up relatively large scale features in the horizontal, ignoring the fine details (the shortest allowed horizontal wavelength is about 15°). The LFM analysis, however, draws to virtually every point, resulting in a much more detailed field.

Because surface data are available at one level and are not extended explicitly, as in the LFM analysis, the response of the vertical function fit to single level data determines the extension. This is illustrated in Figure 4, where a single data point at 1000 mb with value 100 is used to obtain the vertical profile. The fit was obtained by starting with a zero guess and iterating several times, first to alter the coefficient at the first function, then the second and third as in the actual analysis program. Note that the response falls off rapidly to zero, with small overshoots at 700 and 300 mb. For actual situations, the guess field will never be very close to zero, but typically around 50%. Then when the analysis responds to large amounts of moisture near the surface, through surface data only, there will be a slight drying tendency at 700 mb and vice versa for dry low level air.

The satellite bogus humidity, while it receives less weight, has much the same effect as radiosonde data, since data is provided at all mandatory levels. The ten profiles of bogus data are shown in Figure 5, along with the corresponding best fits to the vertical functions (solid lines). Here, as with the surface data, there is a difference in treatment of the data which results in inconsistencies between the Flattery and LFM analyses. The solid vertical lines in Figure 5 are estimates of the humidity values obtained for the satellite bogus data by the LFM code, which treats this data exactly like surface data. In some instances, the LFM numbers differ significantly from those seen by the Flattery code. This variation may explain persistent differences in the LFM and PE precipitation forecasts

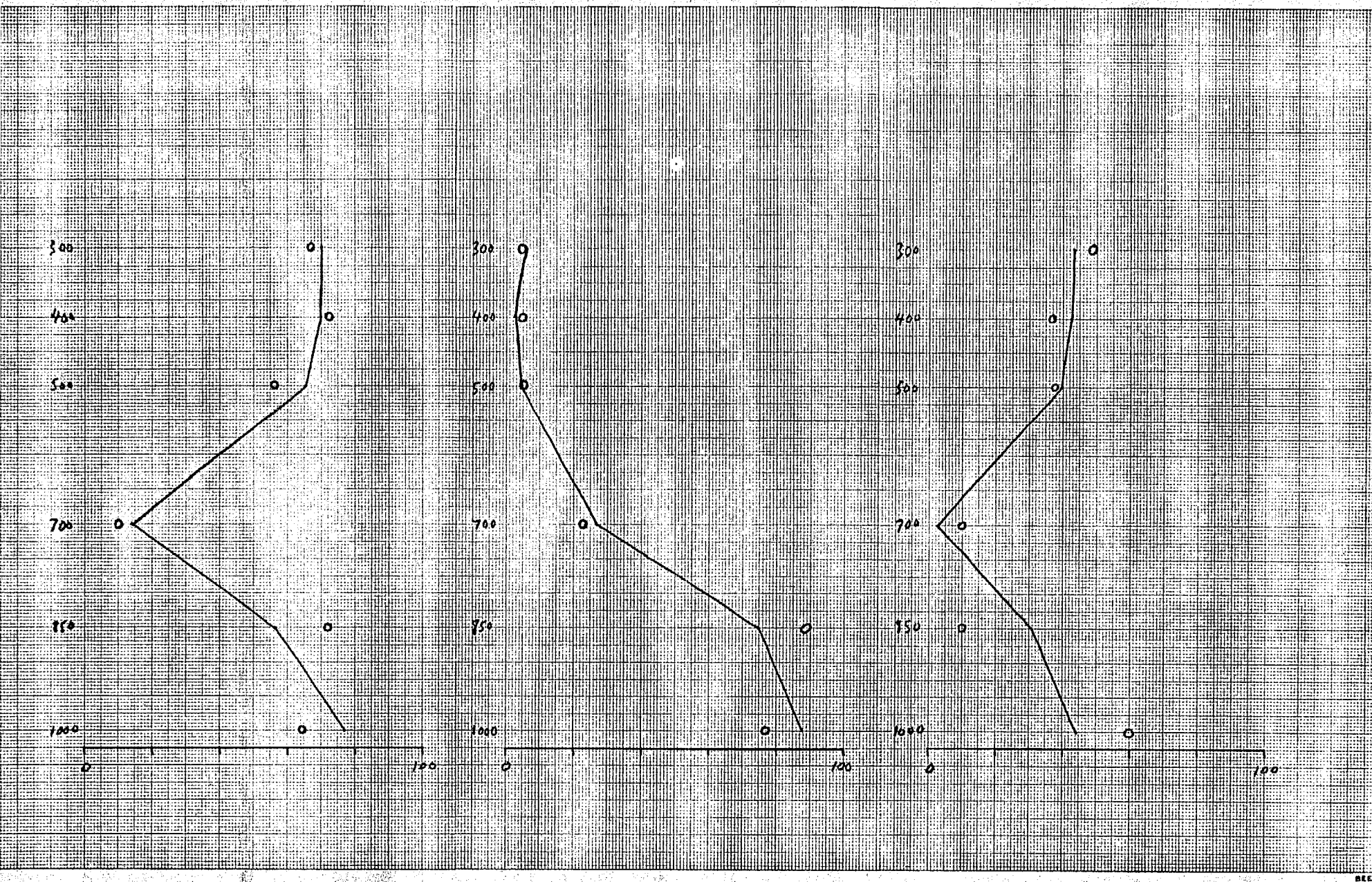
on the West Coast and near the Gulf of Mexico. As an example, we show in Figure 6 a detailed comparison between the Flattery and LFM analyses in the boundary layer around the Gulf of Mexico. The circles off the coast of Texas represent category 7 satellite bogus, which have a value of approximately 85% as seen by the Flattery analysis and 68% as seen by the LFM analysis. Neighboring coastal radiosonde stations are also plotted. Note how the LFM picks up the high humidity inland, but dries out rapidly as we move east into the Gulf, whereas the Flattery field is very wet over the interior of the Gulf. The 90% isopleth extends across the Gulf to the southeast, probably because of the first guess provided to the Flattery analysis which also is considerably different from that used by the LFM. Also, the final output from the LFM code has been reduced everywhere south of 30°N, making the wet-dry contrast even more pronounced.

Fig 1.



BEE M/M

Fig. 2


300
400
500
700
850
1000
0
100
300
400
500
700
850
1000
0
100
300
400
500
700
850
1000
0
100

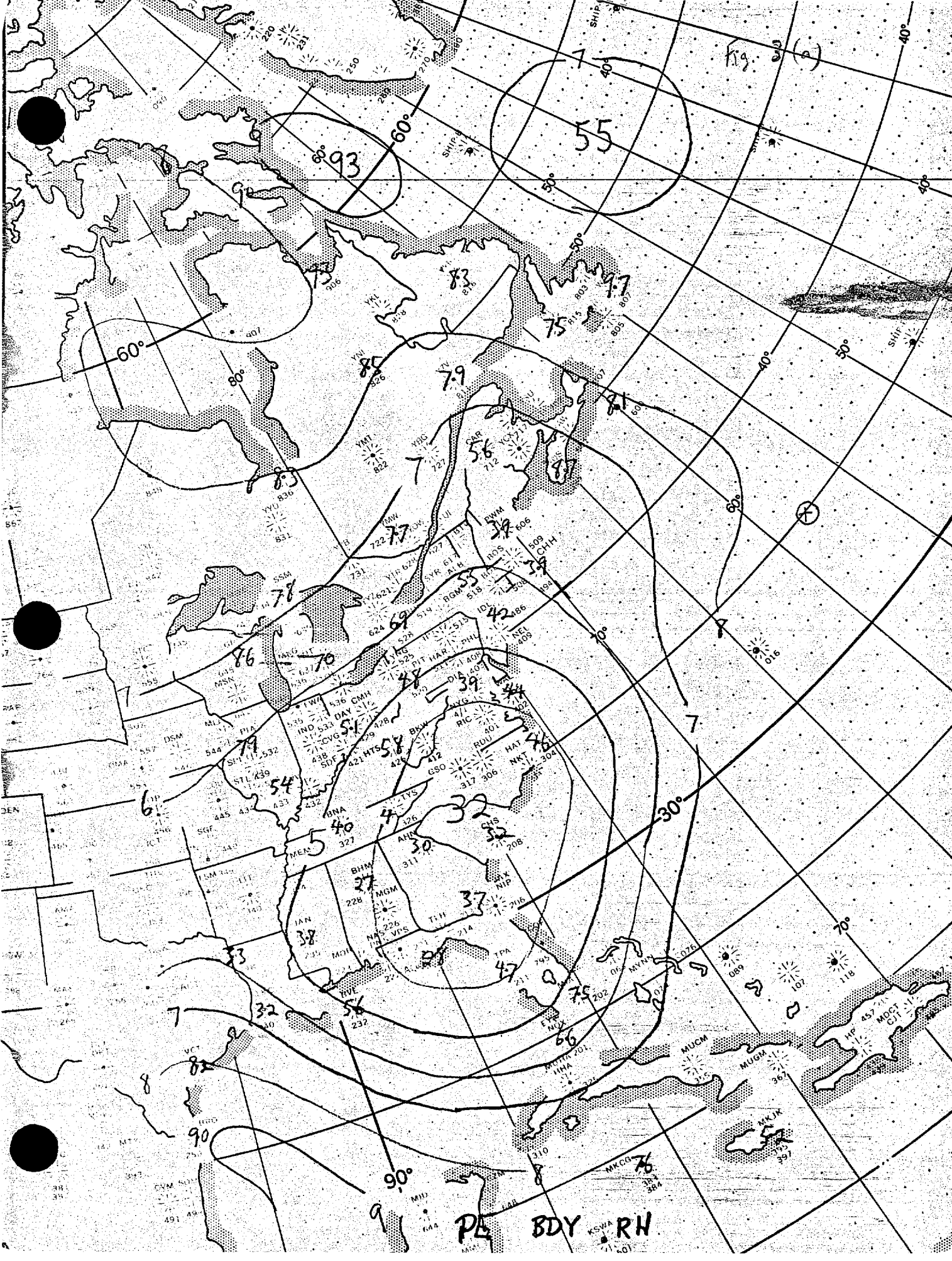


Fig. 3(a)
PE BDY RH

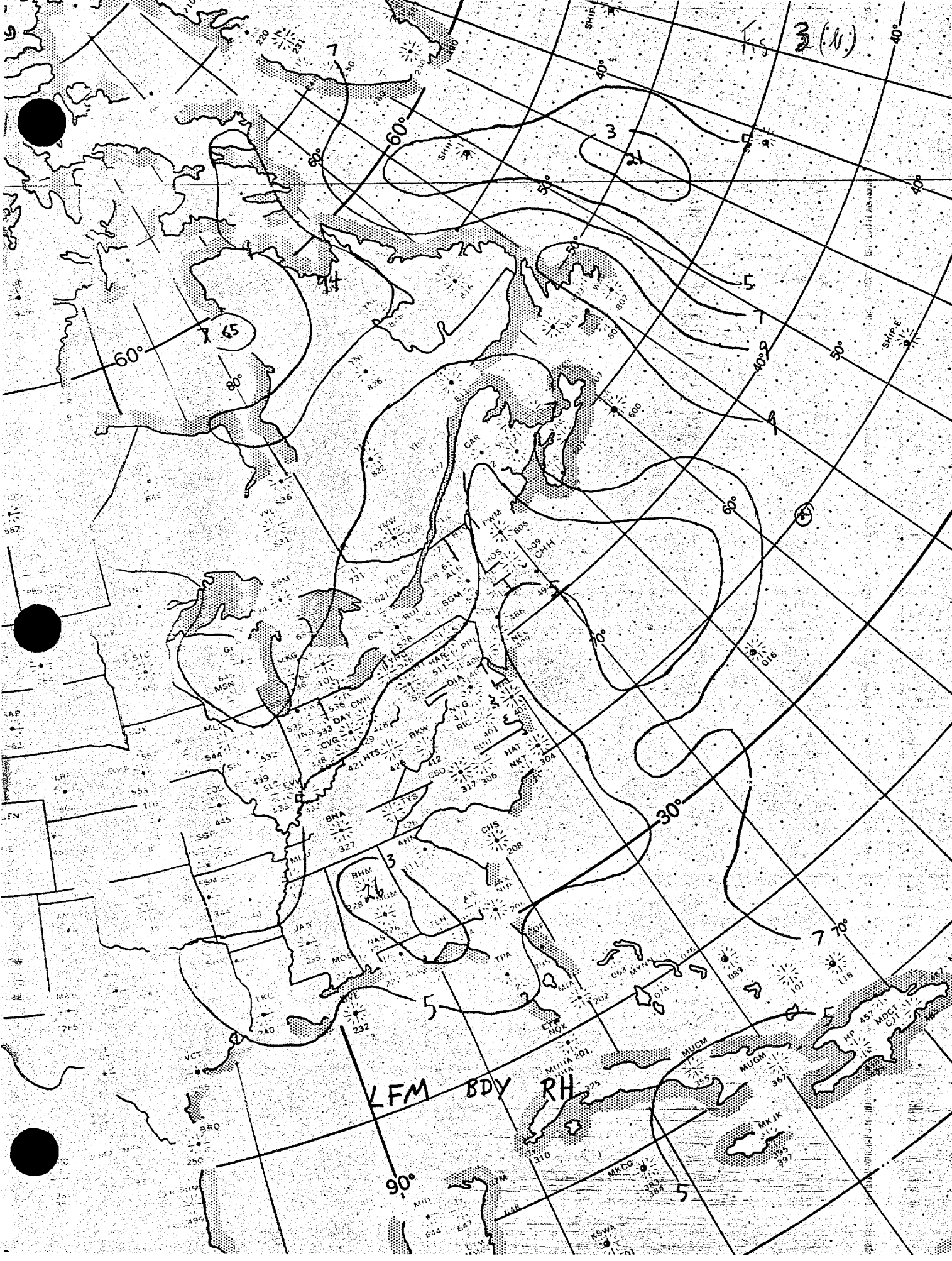
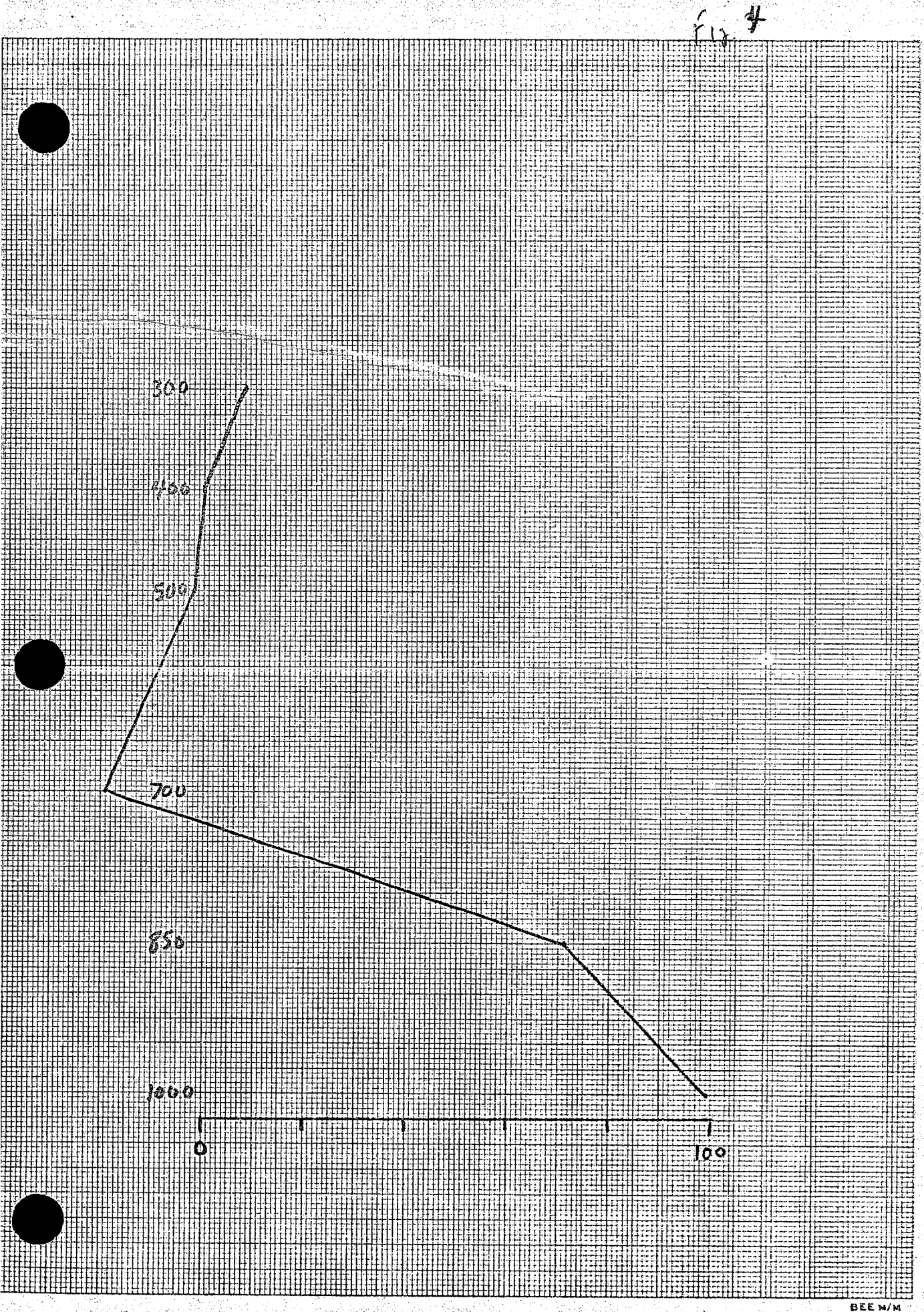


Fig 3 (N.)
LFM BDY RH
60°
50°
40°
30°
90°
70°
80°

Fig. 4



BEE M/M

Figure 5

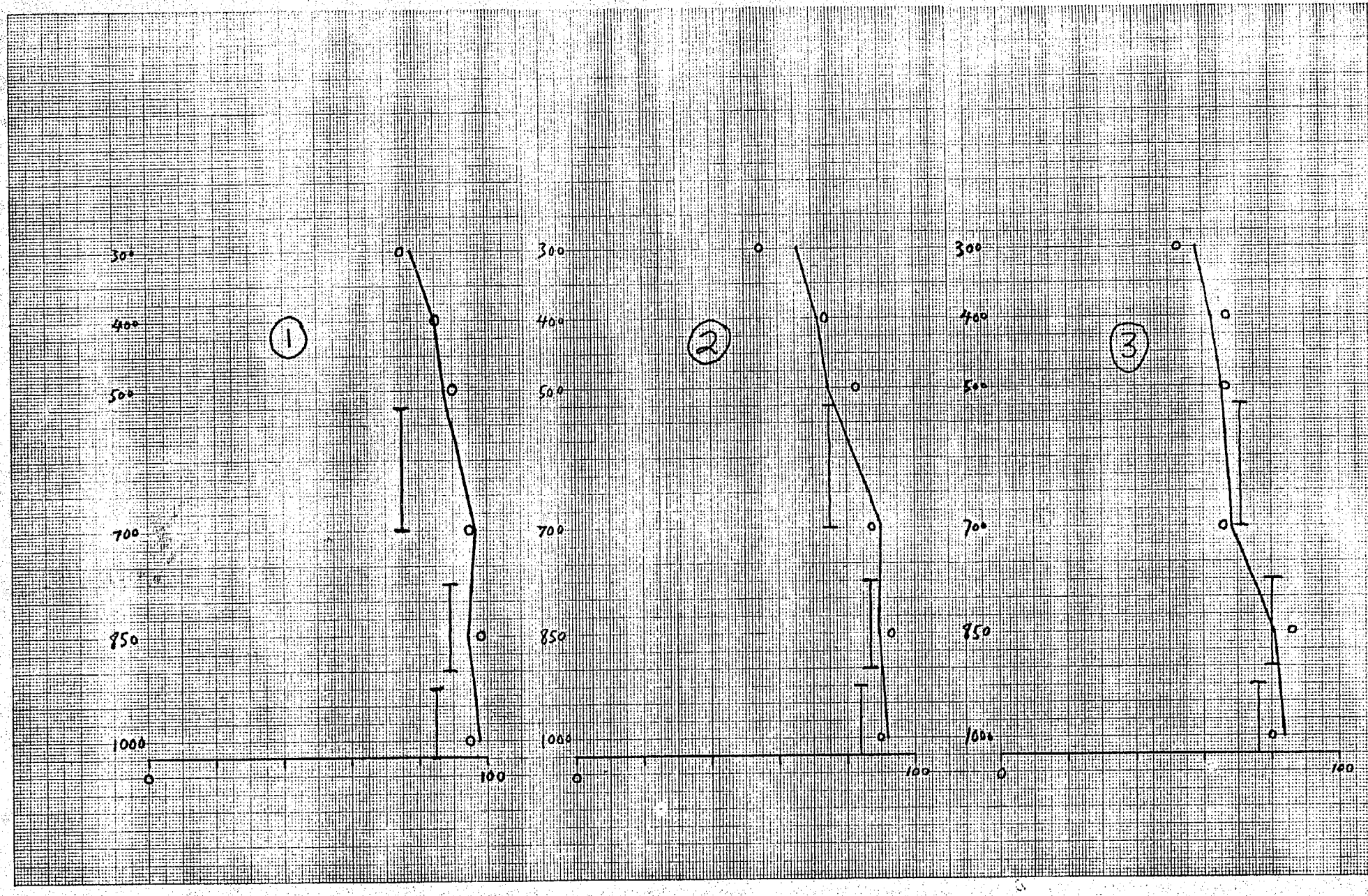


Figure 5

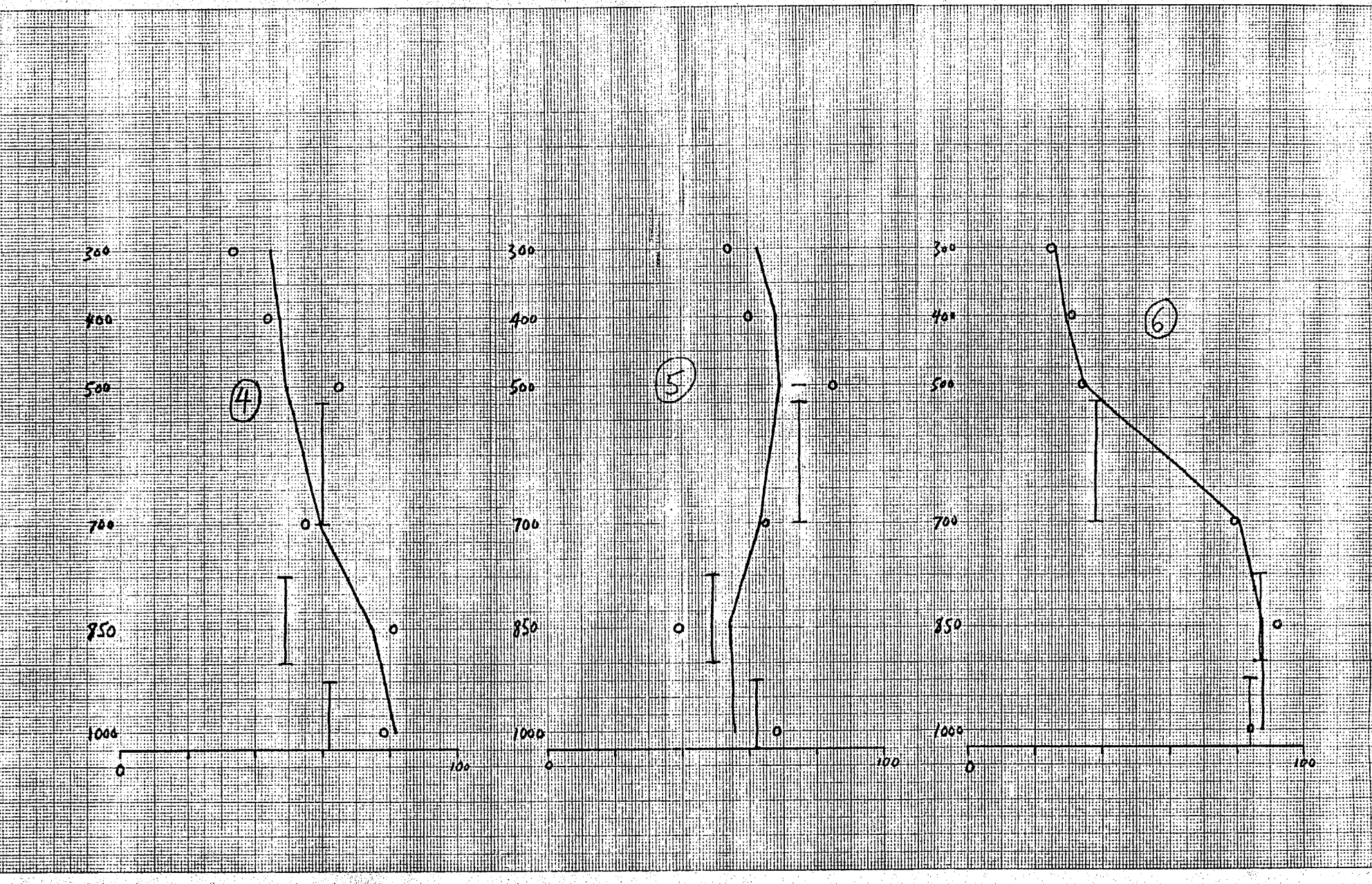
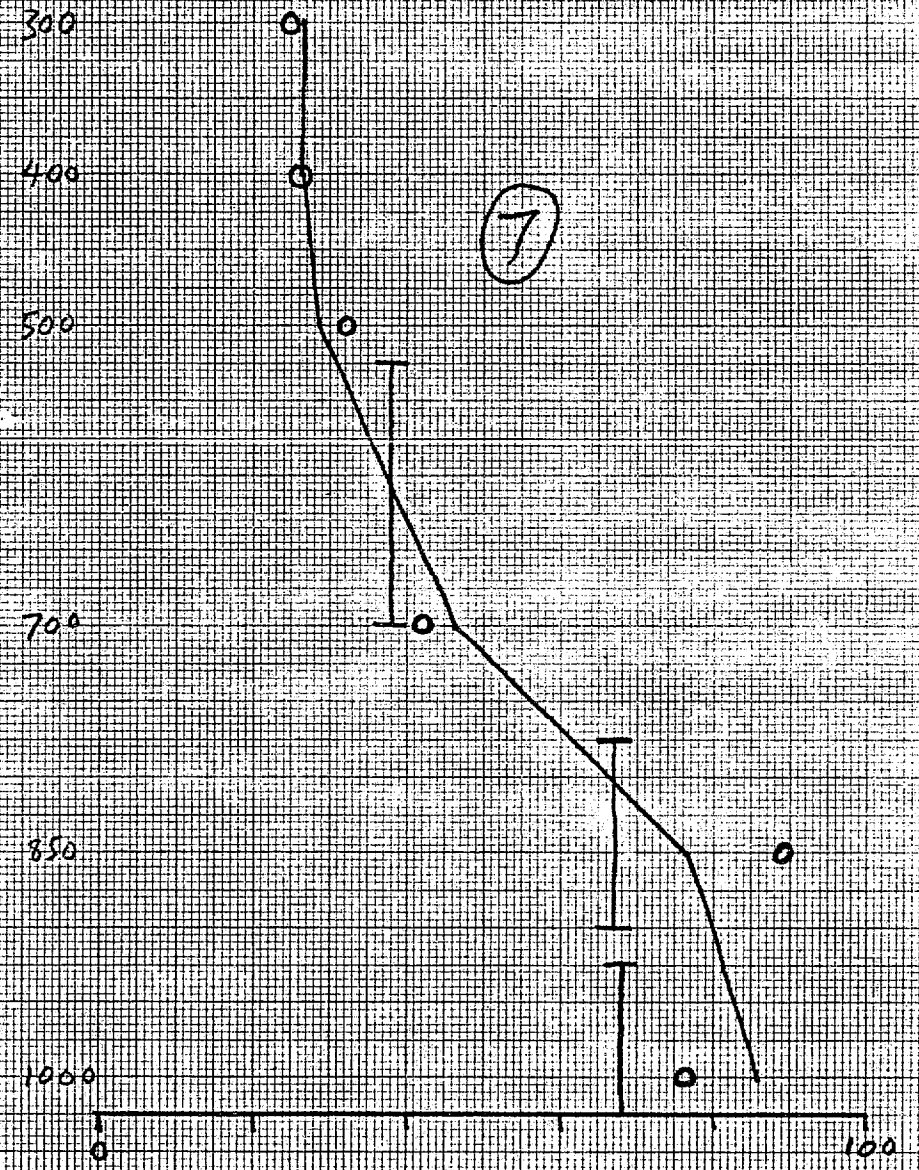
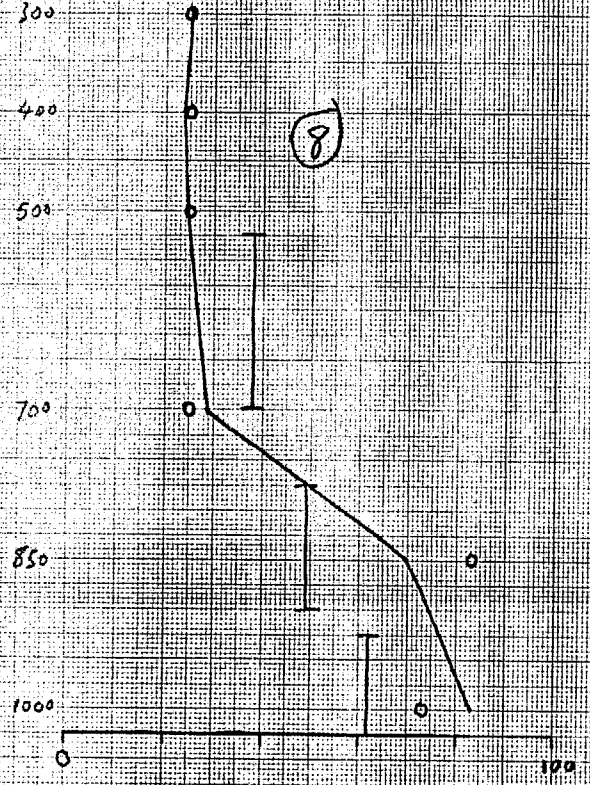


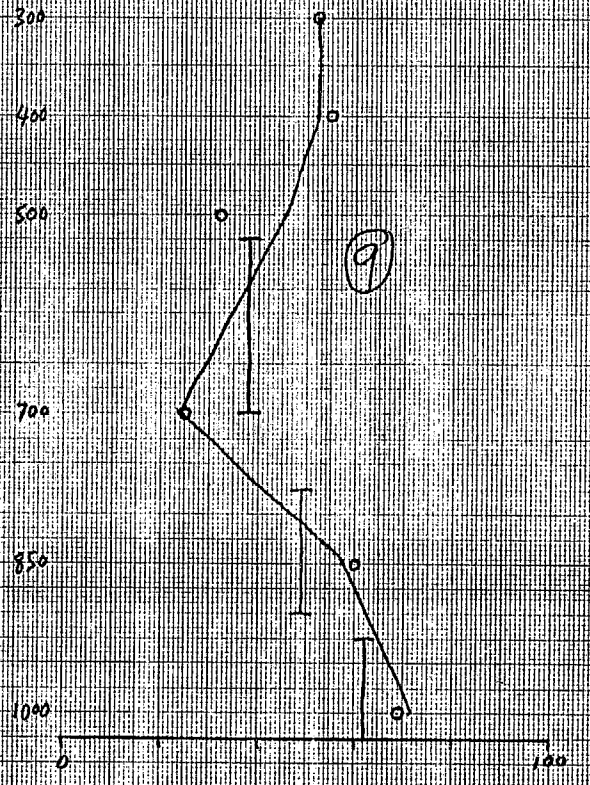
Figure 5

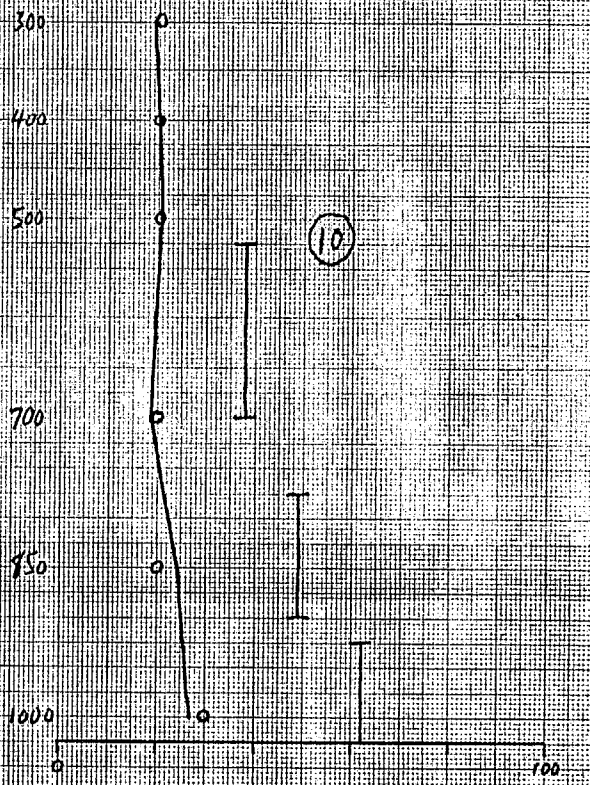


BEE M/M

Figure 5


300
400
500
700
850
1000
0
100
8


300
400
500
700
850
1000
0
100
9


300
400
500
700
850
1000
0
100
10

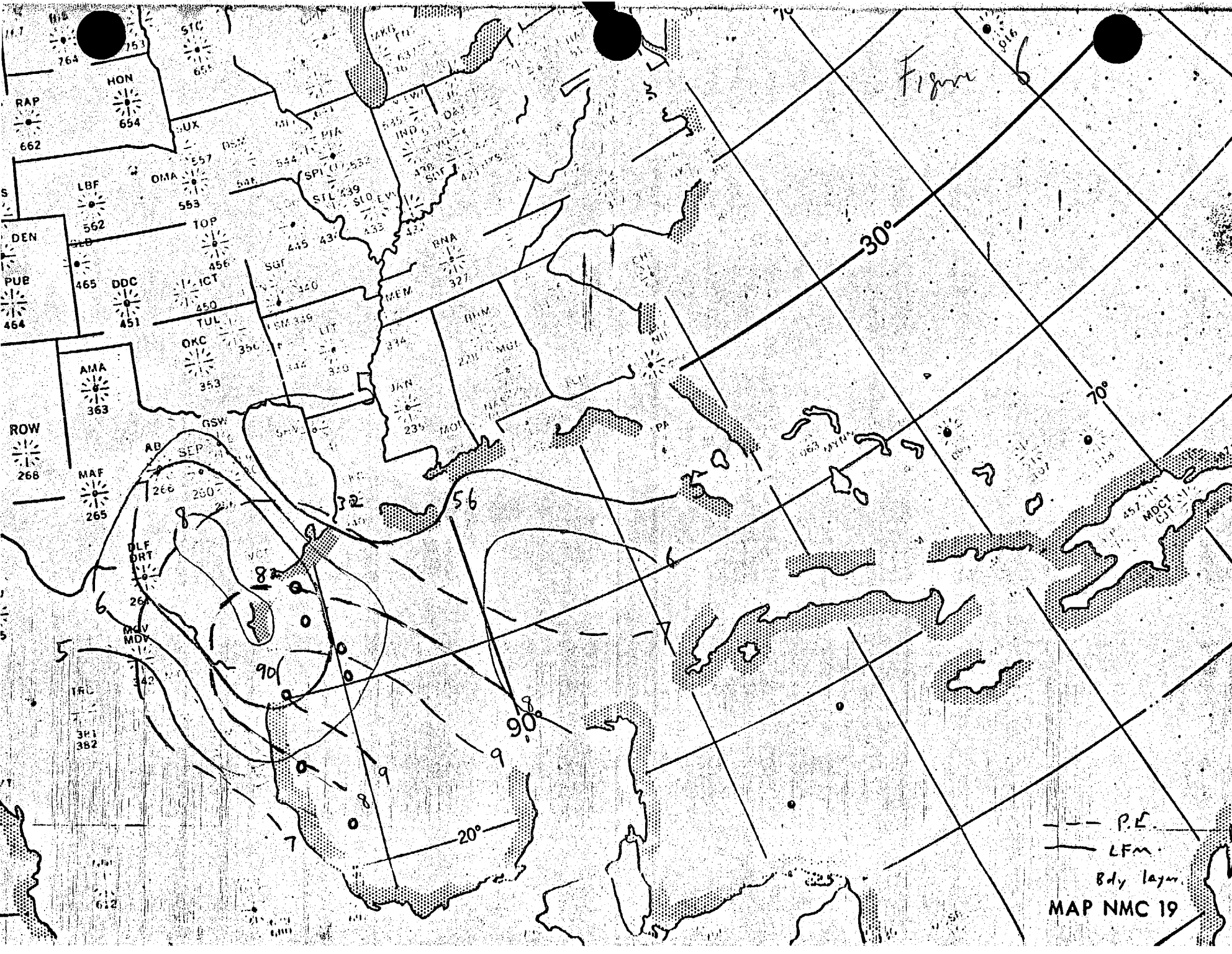


Figure 6
RAP
662
HON
654
STC
LBF
562
OMA
553
DEN
PUB
464
DDC
451
TOP
456
ICT
450
TUL
OKC
353
AMA
363
ROW
268
MAF
265
SEP
AB
GSW
DLF
DRT
MDV
MDV
342
TRC
382
PIA
SPI
STL
SLO
BNA
327
MEM
JAN
235
BHM
MGL
30°
70°
90°
20°
P.E.
LFM
Bdy layer
MAP NMC 19