EASTERN REGION TECHNICAL ATTACHMENT
No. 89-10
September 12, 1989

THE EFFECTS OF BAD AND/OR MISSING DATA ON OUTPUT FROM ADAP

JEFF S. WALDSTREICHER
NATIONAL WEATHER SERVICE FORECAST OFFICE
BOSTON MA

1. BACKGROUND

With the development of objective analysis programs such as the AFOS Data Analysis Programs or ADAP (Bothwell 1988), the operational forecaster now has the means to evaluate the state of the atmosphere in real-time on a smaller scale than ever before. Different analysis fields can be created by blending time and distance weighting of surface and upper-air observations (Doswell, 1977 and Bothwell et. al. 1985). Objective analysis programs can paint a picture of the atmosphere on a sub-synoptic scale, and track how that picture is changing from hour to hour. In the future, with increased computer power and advanced remote sensing technology, the forecaster will be able to examine atmospheric features on even smaller spatial and temporal scales. One of the greatest limiting factors on the reliability of these objective analysis programs is, and will continue to be, the availability and quality of the observations that support them.

The data coverage over the Eastern Region is, for the most part, very good. Forecasters in the east do not have to contend with the data void regions present over areas like west Texas. This more extensive data coverage makes programs like ADAP even more valuable to Eastern Region forecasters. It does not, however, mean that missing data is a problem that can be ignored. Additionally, erroneous data is a problem that can occur any place and at any time. Most objective analysis programs, including ADAP, do check for missing data (i.e. a missing observation or a missing element like dew point). ADAP does not, however, discard an existing value just because it is not "in-line" with the surrounding data. This is done primarily because what may initially appear to be anomalous or erroneous data, may actually be associated with a thunderstorm, or may be related to some other mesoscale feature. The only indication of this feature may be that one "strange looking" value. Doswell (1982) states that for these reasons, when a question of data validity arises, it is probably better to err on the side of data retention.

These are factors that should always be kept in mind when evaluating output from an objective analysis program. Forecasters should be especially alert for situations when the problems of missing and bad data occur simultaneously. The filtering aspects of time and distance weighting within objective analysis programs normally help smooth out erroneous data while still providing enhancement and continuity of true mesoscale features. These filters, however, become less effective when there is less "good" data available to smooth out any "bad" data. This was the case over New England at 1400 UTC on July 14, 1989.

2. THE PROBLEM

Figure 1 shows the analysis grid used for ADAP at WSFO Boston. The open circles are surface stations used while the closed circles indicate the sounding locations utilized. At 14Z on July 14th, ADAP was run for the first time that day. All of the 12Z U.S. soundings were available to the program, but as is sometimes the case at this time of the morning, some of the Canadian soundings were not available.

Below is a listing of the SAVOBS.DT file generated by ADAP at 14Z:

```
            SAO CHECK LIST FOR FILE   SA14Z.DT
CHECK FOLLOWING STATION FOR ERROR IN DATA
MRB  PP= 124 TT=  73 TD= -99 DD=  36 VV=  12 GG=  -9 AL= 990
STATION LEB MISSING
STATION YTR MISSING
STATION YVV MISSING
STATION YWA MISSING
STATION YXR MISSING
STATION YXU MISSING
STATION YYB MISSING
STATION YYZ MISSING
STATION YZR MISSING
END
```

SAVOBS.DT is an error message file that should be checked for every ADAP run. Notice that in addition to several missing Canadian observations (this happens occasionally, especially if the program is run too soon after the hour, and the effects are usually very obvious on the ADAP output), Lebanon, Hew Hampshire is also missing. Martinsburg, West Virginia was also flagged by the program as having a missing dew point.

Figures 2a-d are the ADAP generated surface-base lifted index (SSL), cap strength (SSC), mixing ratio (SMR), and moisture flux convergence (SMC) maps, respectively, for 14Z. Figure 3 is a regional surface plot also generated within our ADAP macro. At first glace, the output appears reasonable, with the effects of the missing Canadian data quite evident. The LI computed from the 12Z Albany, New York sounding was approximately -2, which fits in well with the LI analysis over eastern New York. This fact, however, might lead a forecaster to question the +5 LI center over Vermont. Since a cold front had moved through New England overnight, this might represent drier, more stable air moving in. The almost identical looking bulls-eyes over the same location in the cap strength and mixing ratio fields might also foster some suspicion, but they also seem to confirm the hypothesis of drier and more stable air.

A look at Figure 3, however, reveals that the suspicion was justified. Notice that Montpelier, Vermont (MPV - located in north central Vermont) has reported an erroneous dew point of 33 degrees. Referring back to Figure 1, notice that this area of the grid does not have the density of data that most other areas have. Since the LEB observation was missing, the effects of the bad data really become acute. For some of the grid points over northern and central Vermont and New Hampshire, the only stations available to help smooth out the bad data were Burlington, Vermont (BTV) and Sherbrooke, Quebec (YSC). Since ADAP can only utilize a maximum of 100 stations in its final analysis (this was done to reduce computer time), stations such as St. Johnsbury, Vermont (plotted on Figure 3 northeast of Montpelier), and Laconia, New Hampshire (in central New Hampshire) are not included in the final analysis.

3. CORRECTING THE BAD DATA

In an attempt to see just how much the one bad number corrupted the output, as well as to confirm that the atmosphere was still unstable, I edited the SA14Z.DT data file, changing only the Montpelier dew point from 33 to 60 degrees (MPV reported 59 the previous hour and 62 the following hour). Making this change took only a few seconds using the E:F/SA14Z.DT command at an ADM, and since the observations were already decoded, only the analysis portion of our ADAP macro needed to be run. The analysis portion of our ADAP macro takes only 2-3 minutes to run while the entire macro (including decoding the observations) takes about 10 minutes. The resulting ADAP output is shown in Figures 4a-d. Notice how the bulls-eyes over Vermont and Hew Hampshire have completely disappeared. The lack of difference in the moisture flux convergence field was due to the very weak wind field present (see Figure 3). The theta advection field (STA - not shown) displayed no differences between the two runs, which should be expected since the bad dew point should only affect the moisture and stability related fields.

4. CONCLUSIONS

It is apparent from this example that one bad number can severely impact objective analysis programs like ADAP, especially over data sparse regions where there is less "good" data to smooth out the problem. In addition, this problem was enhanced by the fact that ADAP had not been run earlier in the day. ADAP has the capability to look back at the previous two hours of data and incorporate these values into the analysis via "time weighting" much the same way observations surrounding a grid point are "distance-weighted". While ADAP does not require the previous two hours of data be present, it does help in the smoothing process by providing continuity. Since, the program did not have the surface data from 12Z or 13Z), it was unable to "time-filter" the bad data. This indicates that frequent ADAP runs help the reliability of the output, as well as helping to maintain the continuity of mesoscale features.

Had the initial output been accepted as valid, with the bulls-eyes determined to represent drier and more stable air moving into the region, a forecaster might be lulled into a false sense of security. Later ADAP runs indicated that rather than more stable air entering the region, the atmosphere continued to be unstable, with surface-based LI's remaining -3 to -5 throughout the day. The weak cap present was broken with the help of surface heating, and with some weak PVA which moved across the region during the afternoon, thunderstorms developed rapidly around 17Z and continued throughout the afternoon. While no reports of severe weather were received over land, there were numerous reports of small (pea-sized) hail, heavy downpours and frequent lightning throughout central and southern New England that afternoon and evening. Several Special Marine Warnings were needed for the coastal waters south of New England as the thunderstorms moved offshore. A waterspout was also reported over Great Peconic Bay between the Twin Forks of Long Island around 22Z.

Objective analysis programs like ADAP are exceptionally valuable forecasting tools. It is important, however, that like any other forecasting tool (e.g. MOS etc...) they be used with a full understanding of their capabilities and limitations. Blindly accepting any meteorological information can lead even the best forecaster astray.

5. REFERENCES

Bothwell, P.D. 1988: Forecasting Convection With the AFOS Data Analysis Programs (ADAP-Version 2.0). NOAA Technical Memorandum SR-122. Fort Worth, Texas. 92pp.

Bothwell, P.D., R.A. Maddox. C. A. Doswell III, and K.C. Crawford, 1985: Operational Methods for Increasing the Reliability of Information Derived from Conventional Surface and Upper-Air Data. Preprints, 14th Conf. on Severe Local Storms (Indianapolis, IN), AMS, Boston MA, 402-405.

Doswell, C. A. III, 1977: Obtaining Meteorlogically Significant Surface Divergence Fields Through the Filtering Property of Objective Analysis. Mon. Wea. Rev., 105, 885-892.

Doswell, C. A. III, 1982: The Operational Meteorology of Convective Weather, Volume I: Operational Mesoanalysis, NOAA Technical Memorandum NWS-NSSFC-5, NOAA/NWS, Kansas City, Mo.,

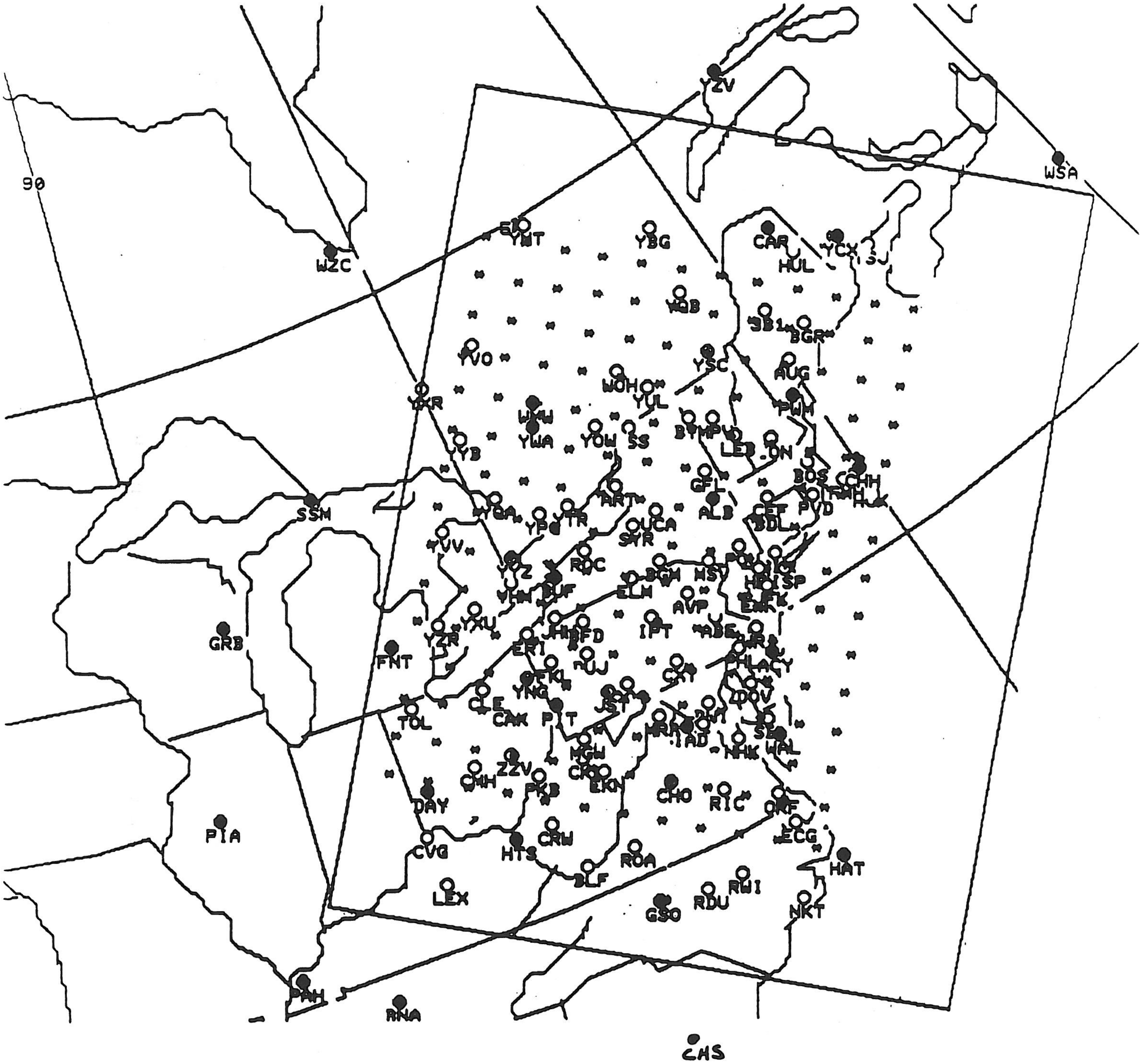

FIGURE 1: Analysis grid used in the ADAP program at WSFO Boston.
Open Circles indicate surface observations used.
Closed Circles indicate soundings (and surface obs.) used.

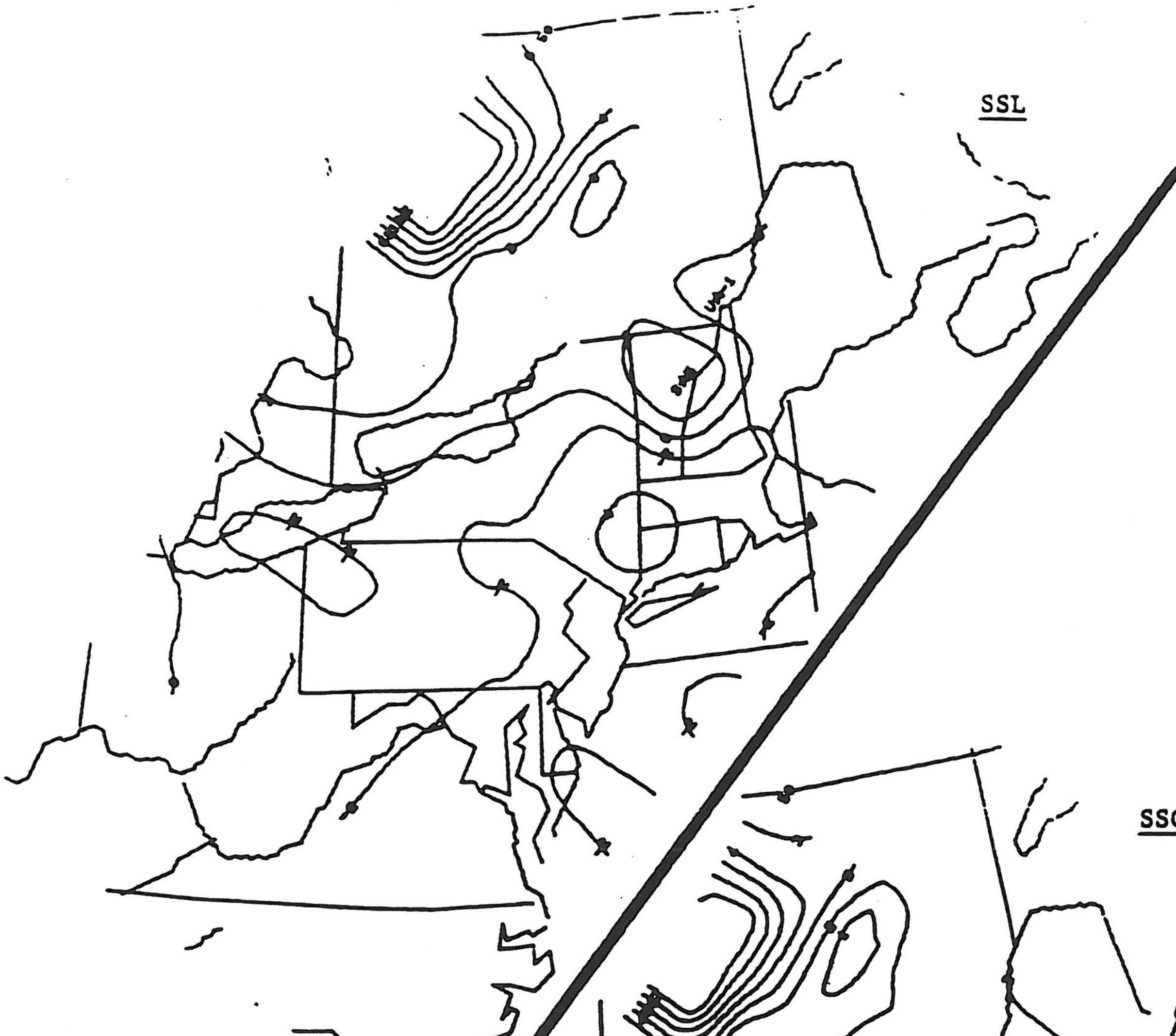

FIGURE 2a: Surface-Based Lifted Index (SSL) for 14Z July 14th.

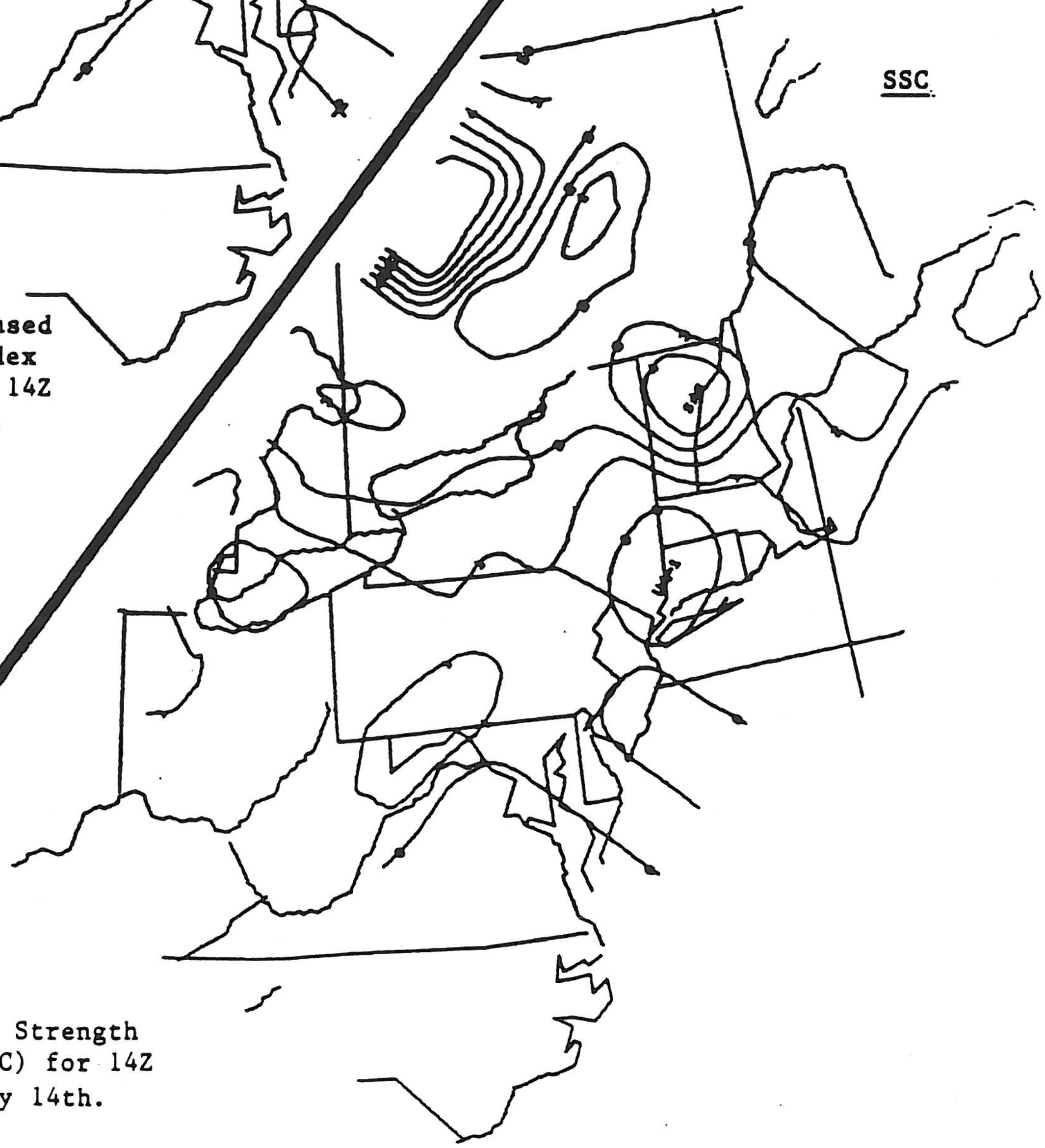

FIGURE 2b: Cap Strength (SSC) for 14Z July 14th.

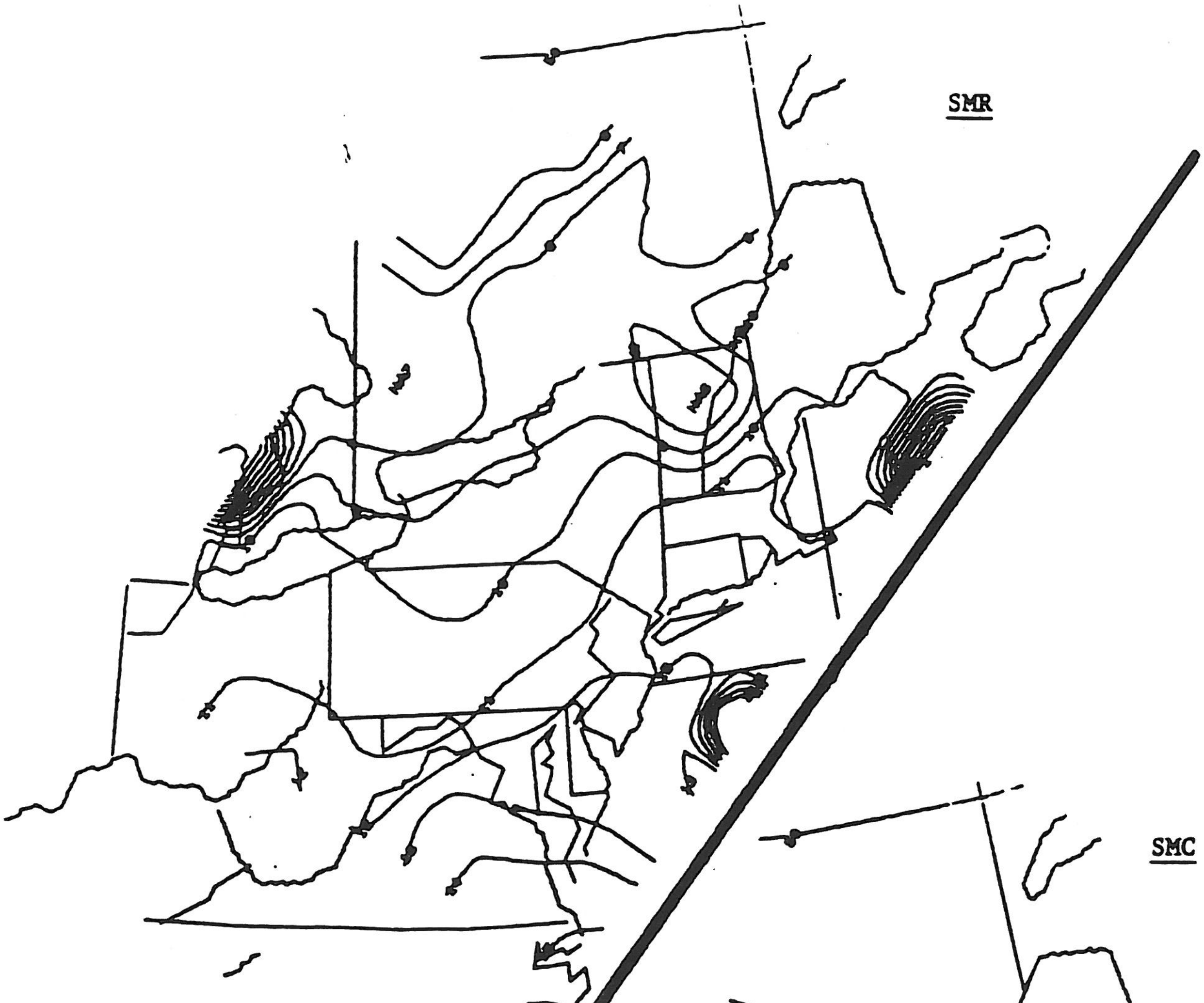

FIGURE 2c: Surface Mixing Ratio (SMR) for 14Z July 14th.

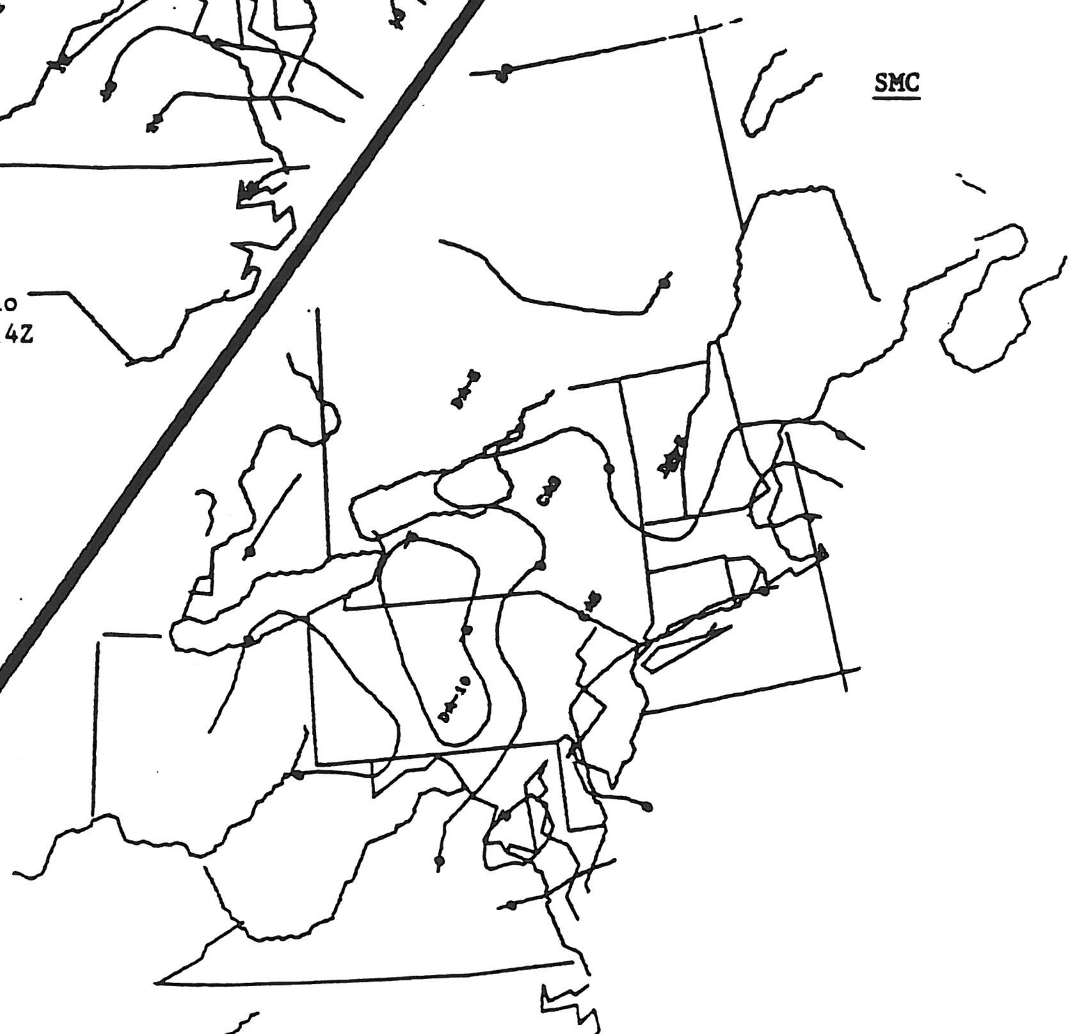

FIGURE 2d: Surface Moisture Flux Convergence (SMC) for 14Z July 14th.

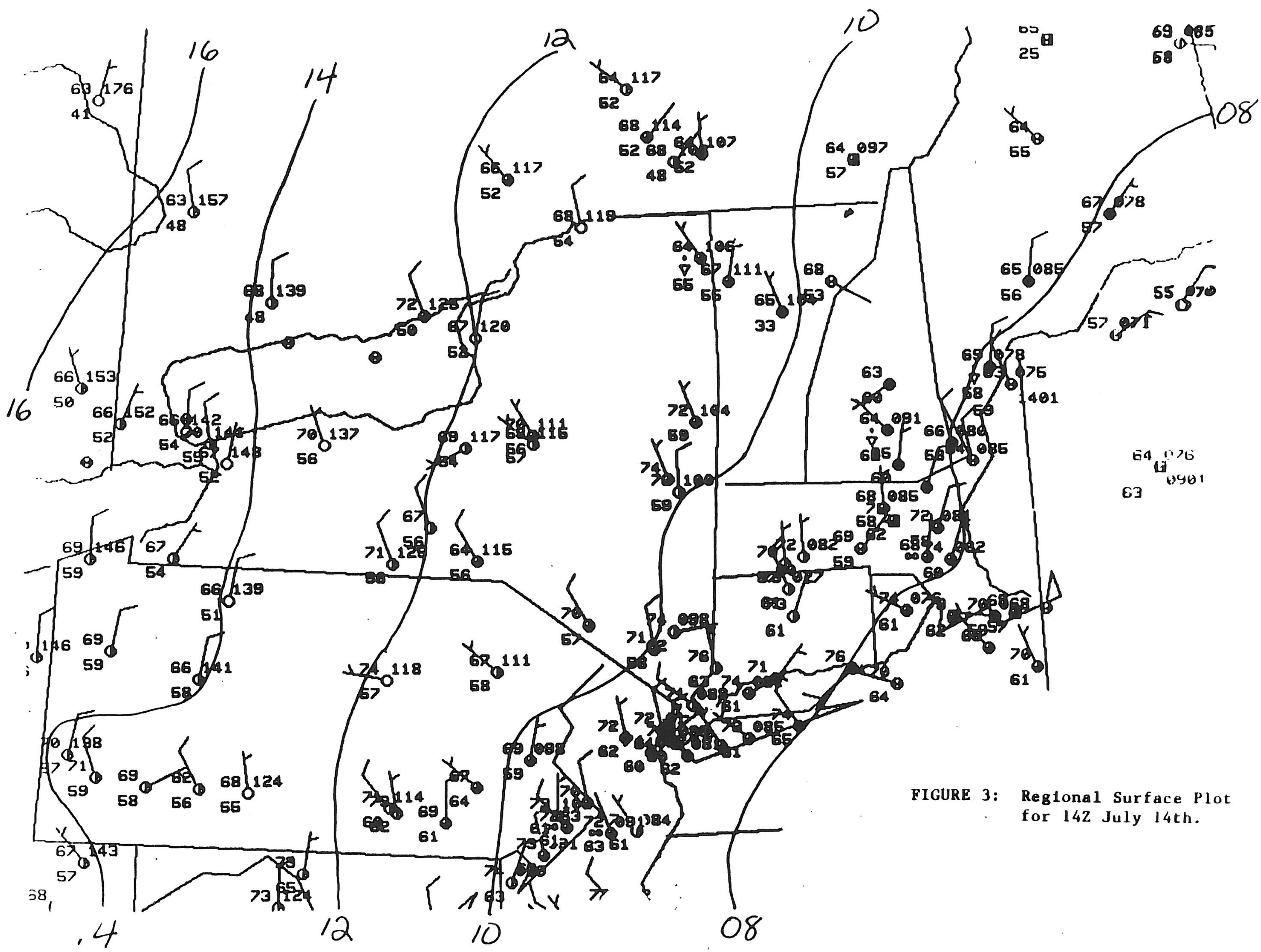

FIGURE 3: Regional Surface Plot for 14Z July 14th.

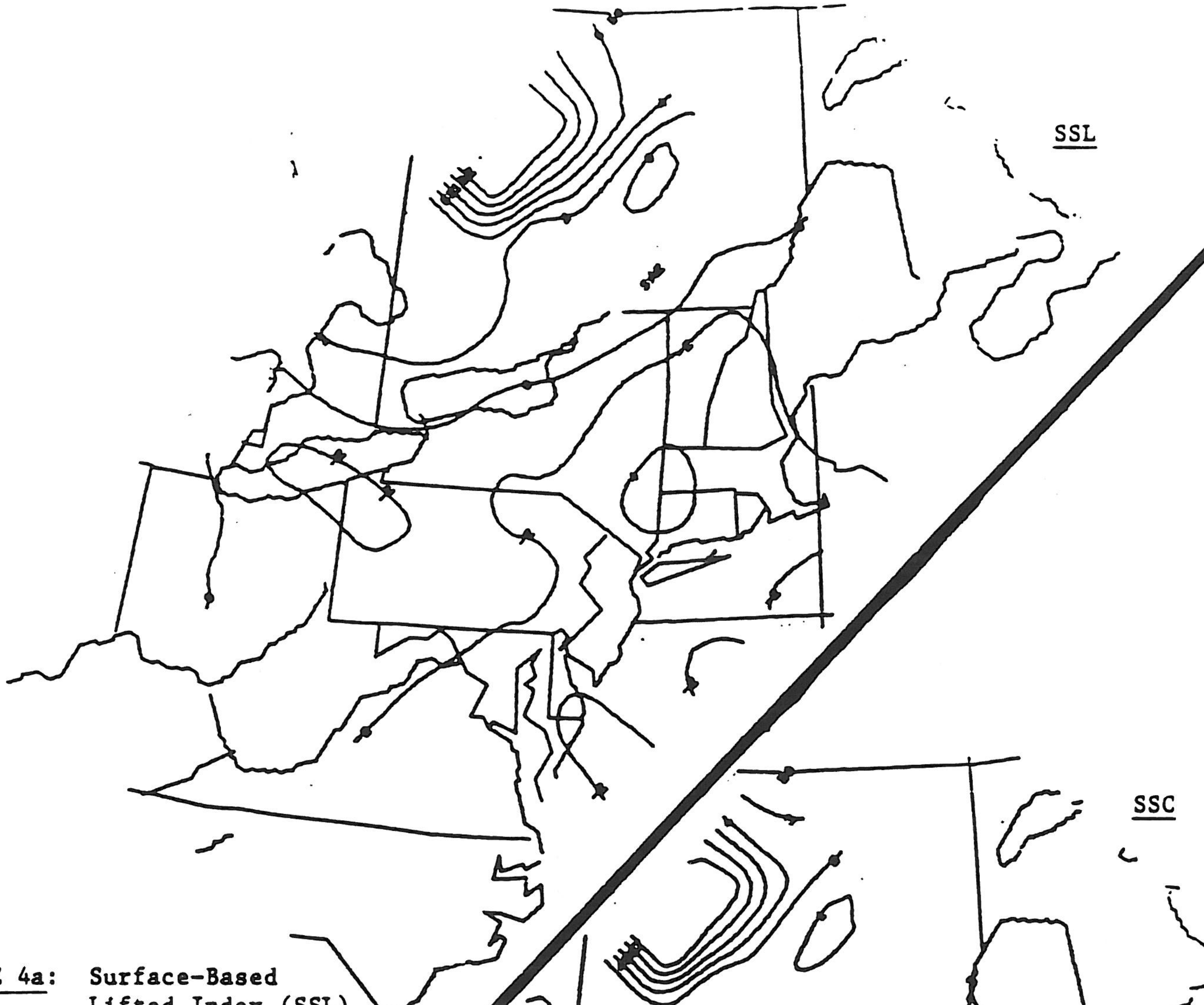

FIGURE 4a: Surface-Based Lifted Index (SSL) Recomputed for 14Z July 14th.

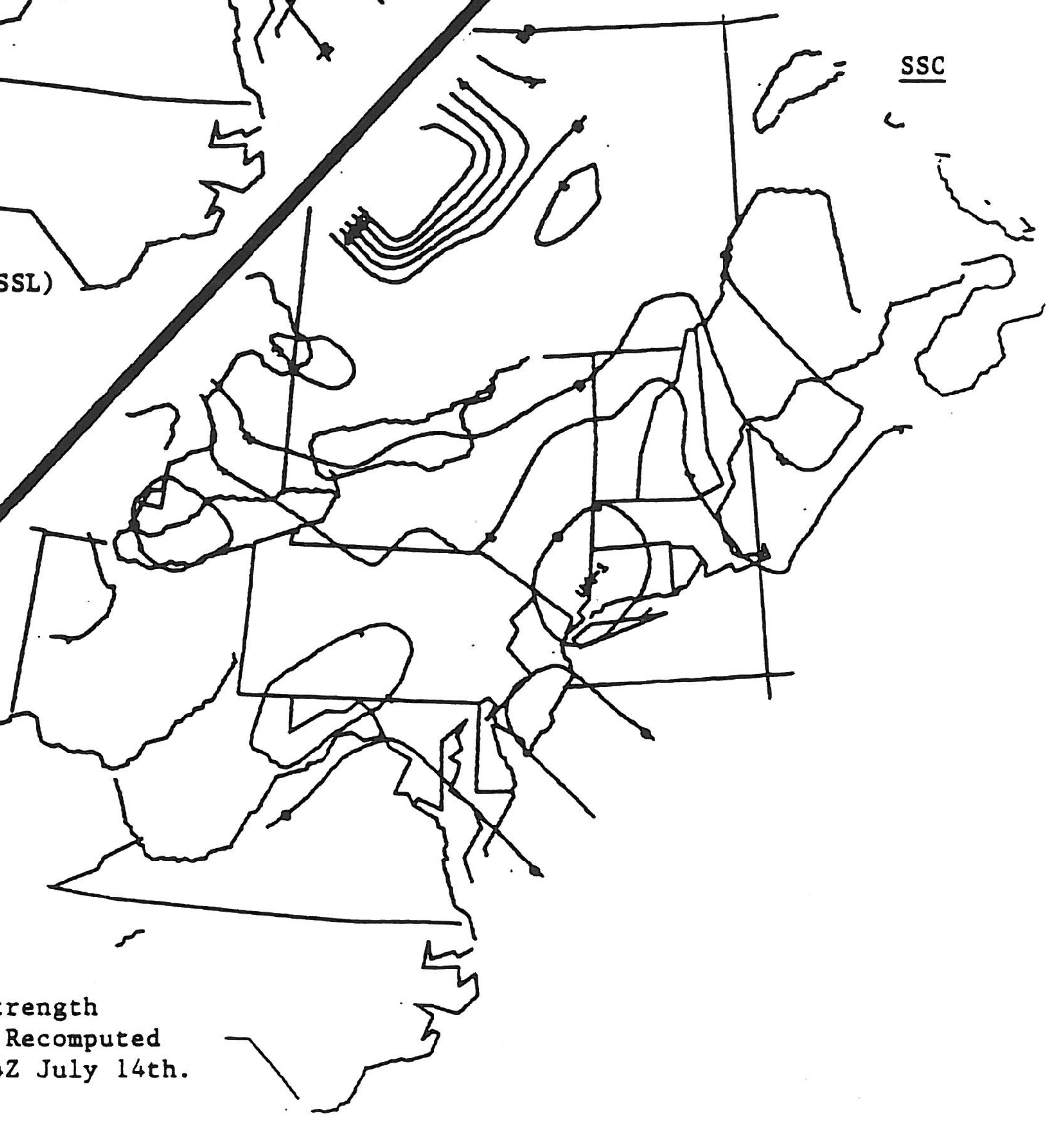

FIGURE 4b: Cap Strength (SSC) Recomputed for 14Z July 14th.

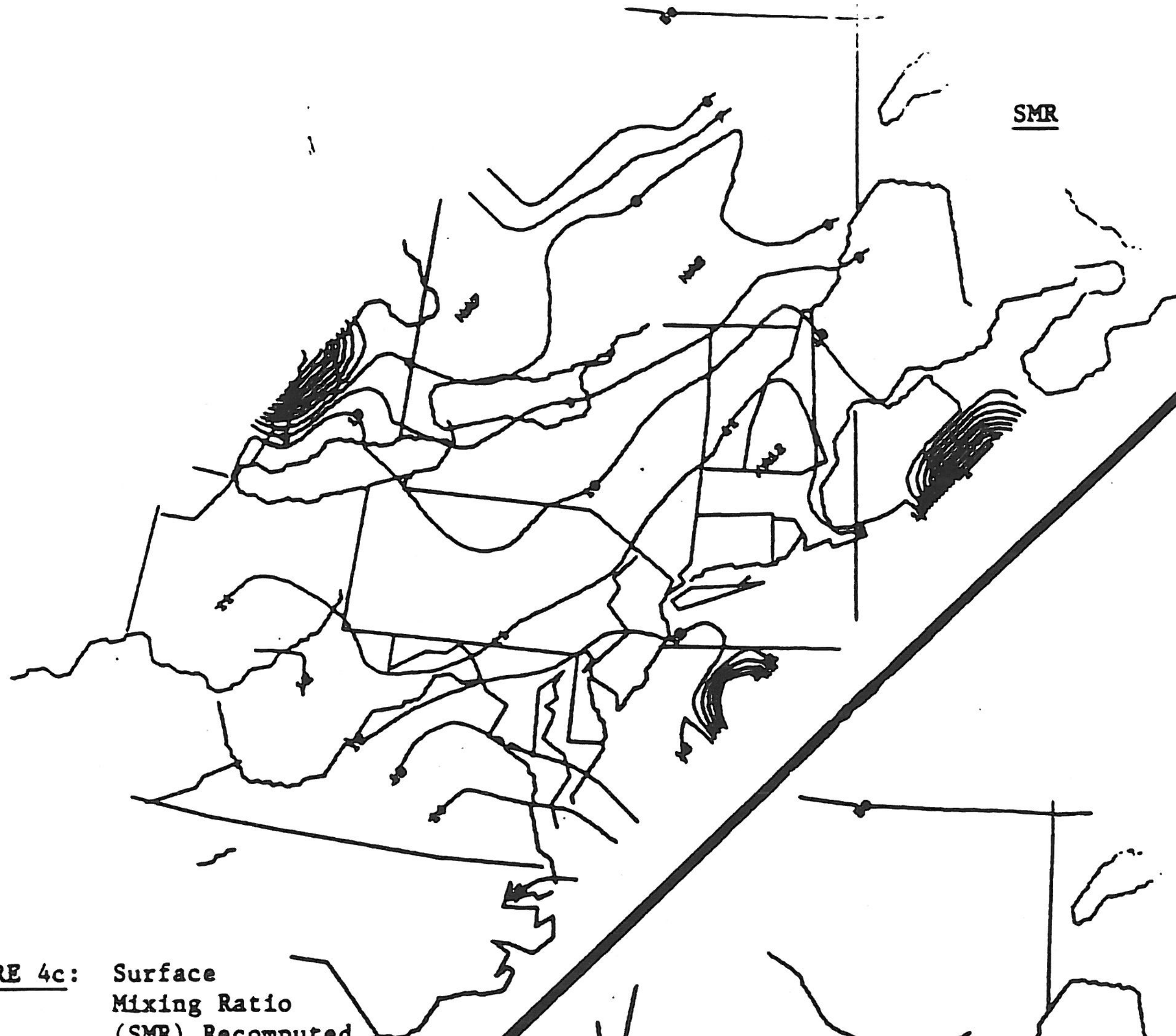

FIGURE 4c: Surface Mixing Ratio (SMR) Recomputed for 14Z July 14th.

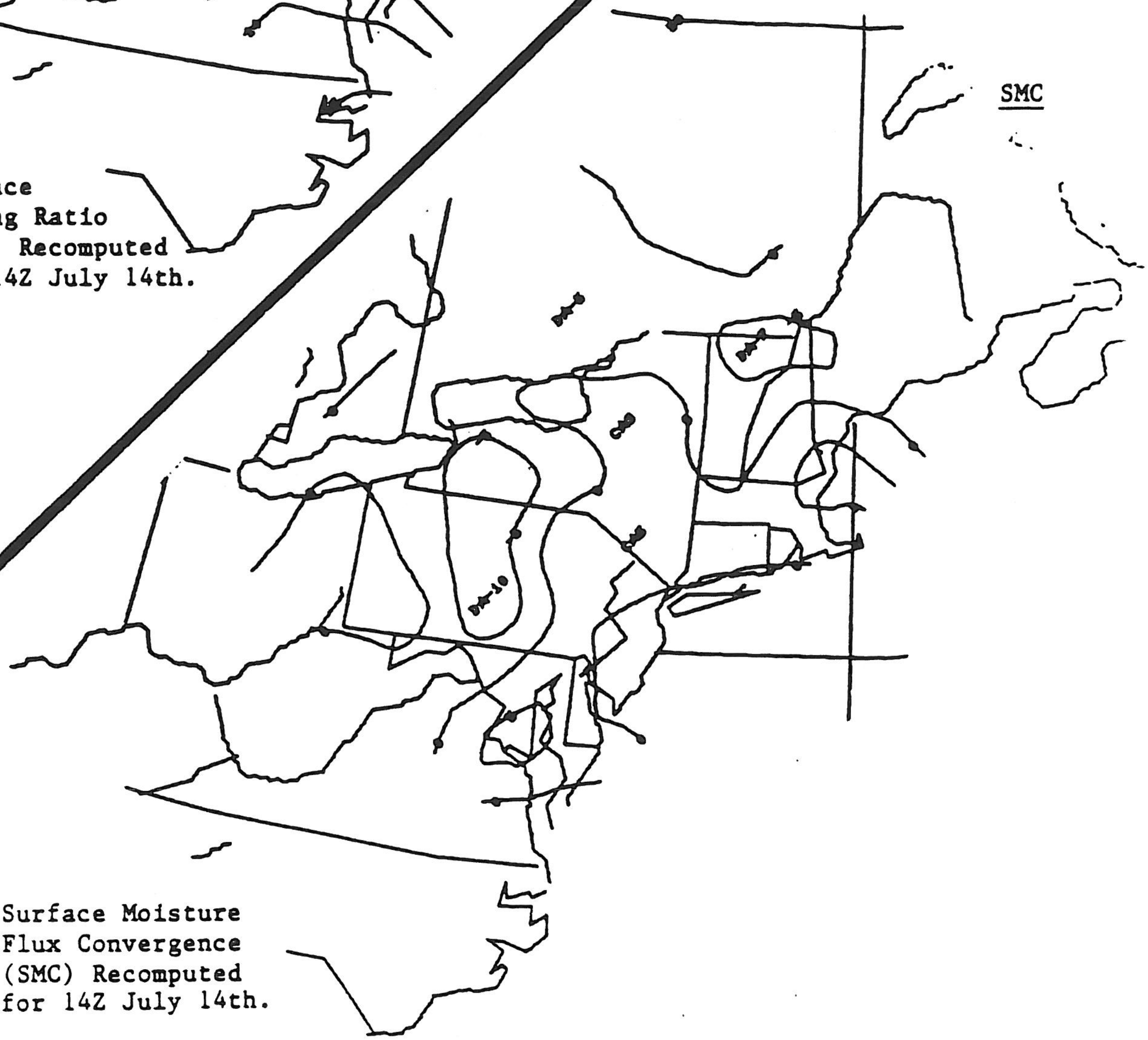

FIGURE 4d: Surface Moisture Flux Convergence (SMC) Recomputed for 14Z July 14th.