U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 81-5

THE EXPERIMENTAL CONVECTIVE OUTLOOK (AC) CHART: COMPARATIVE VERIFICATION AND PRELIMINARY EVALUATION

Ronald M. Reap, Donald S. Foster, and Steven J. Weiss

September 1981

THE EXPERIMENTAL CONVECTIVE OUTLOOK (AC) CHART: COMPARATIVE VERIFICATION AND PRELIMINARY EVALUATION

Ronald M. Reap, Donald S. Foster, and Steven J. Weiss[1]

1. INTRODUCTION

This office note gives a description and evaluation of the new automated convective outlook (AC) chart developed at the Techniques Development Laboratory for use as guidance by severe storm forecasters at the National Severe Storms Forecast Center (NSSFC). The automated AC is prepared once-daily from Model Output Statistics (MOS) probability forecasts of thunderstorms and severe local storms. A sample of the MOS probabilities (Reap and Foster, 1979) is shown in Fig. 1. The probabilities and the derived AC chart are both valid for the 12-36 h interval following 0000 GMT initial data time. The AC chart is constructed in the format of an Automation of Field Operations and Services (AFOS) product and is transmitted to NSSFC via high-speed line for display on the AFOS graphics system. To provide operational backup, the automated AC is also transmitted to NSSFC in the form of printout containing background geography with superimposed digits indicating the severe local storm risk area(s).

2. THRESHOLD PROCEDURE

In order to construct the AC chart, it is necessary to generate categorical forecasts from the probability forecasts of thunderstorms and severe local storms. As a first step, the product of the unconditional thunderstorm probability and the conditional severe local storm probability is computed for each of the manually-digitized radar data (MDR) grid blocks upon which the probabilities are based (Reap and Foster, 1979). The grid blocks are approximately 75-80 km on a side. Next, the product field of unconditional severe local storm probabilities obtained from the previous step is searched to obtain the maximum probability value (P_m). A threshold value is then assigned, based on the value of P_m, as shown in Table 1 and Fig. 2. To obtain the threshold values, we stratified probability forecasts of the spring and summer seasons for the 1977-79 period into 12 categories where $0 < P_m \leq 3$, $3 < P_m \leq 6$, $6 < P_m \leq 9$, ...,$33 < P_m \leq 36$. That is, all the probability forecasts for each day were placed into one of 12 categories depending on the value of P_m for that day. Each category was then independently verified to determine the threshold value that gave the maximum critical success index (CSI)[2]. The results of this verification, as shown in Fig. 2, indicate a marked variation of the threshold value and the associated probability of detection (POD)[3] with P_m. In our opinion, the pronounced increase of the threshold value and POD with increasing P_m reflects the tendency towards increased "clustering" of severe weather events on active storm days in response to specific and fairly localized features in the large-scale flow. In making this statement, we implicity assume that the severe local storm probabilities are reliable and reflect to a reasonable degree the amount and intensity of severe storm activity.

---

[1]Affiliated with the National Severe Storms Forecast Center

[2]The CSI and POD were computed according to the definitions established by Donaldson et al. (1975). They are defined in Section 4 below.

[3]Ibid.

Table 1. Threshold value as a function of maximum unconditional severe local storm probability (P_m).

Threshold Value (%)	P_m (%)
2.0	2.0
3.0	3.0
4.0	4.5
5.0	7.0
6.0	10.0
7.0	13.0
8.0	16.0
9.0	19.1
10.0	22.1
11.0	25.2
12.0	28.2
13.0	31.3
14.0	34.3

3. CONSTRUCTION OF THE AC CHART

The procedure for delineating outlook areas for both thunderstorms and severe local storms on the experimental AC chart is now described in detail.

A. Severe Storm Areas

To determine if multiple outlook areas exist, the unconditional severe local storm probability array is searched to locate areas of maximum probability surrounded by probabilities less than 2%. On the basis of this scan, up to five separate areas can be identified and contoured with as many as three distinct threshold values. As previously indicated, all threshold values are extracted from the data given by Table 1. From experience with the automated AC, we have found that outlook areas with threshold values of 4% or greater are of sufficient areal extent to contour and display. However, when the threshold value is 3% and there are less than five MDR grid blocks within the area, the threshold is lowered to 2.5% and the area redefined. If the new area still contains less than five blocks, it will not be contoured because of its small size. Also, if no areas are detected with a threshold of 2% or greater, the message "SVR LCL STM PROBABILITIES BELOW THRESHOLD" is printed across the AC chart.

The coverage for each severe local storm outlook area is computed by dividing the number of expected severe storm grid blocks by the total number of blocks within the outlook area. The number of expected severe storm blocks is obtained by multiplying the average probability for the blocks in the outlook area by the total number of blocks in the area. Therefore, the coverage is equivalent to the average probability for all grid blocks within the outlook area.

The risk codes and associated coverages for the severe local storm outlook areas, as given by Table 2, were defined by NSSFC (NWS, 1980) in terms of the high-resolution MDR grid in operational use since March 1, 1978. The grid blocks are approximately 20 km on a side, or one-fourth the size of the old MDR grid blocks. Therefore, it is necessary to convert the coverage estimates based on

the MOS probabilities, which were derived on the old MDR grid, to coverage estimates valid on the new grid. Based on our experience to date, we have assumed that each expected severe storm block on the old MDR grid is equivalent to two severe storm blocks on the new grid. While this assumption appears to work reasonably well, we plan to examine a fairly large sample of severe storm reports on both grids to determine the average number of new MDR grid blocks affected by each report on the old MDR grid.

Table 2. Risk codes and associated coverages within the severe local storm outlook area.

Coverage (%)	Code
1.5	Approaching (APCG)
2.0	Slight risk (SLGT)
6.0	Moderate risk (MOD)
10.0	High risk (HIGH)

The risk codes in Table 2 are plotted within the outlook area at the location of the maximum probability value. If the coverage is less than 1.5%, the area will not be contoured.

Samples of the automated AC chart prepared by the NOAA FR-80 system are shown in Figs. 3 and 4. The cases were selected to illustrate a variety of multiple areas and coverages. A sample automated AC as displayed on AFOS is shown in Fig. 5. The AC chart is also transmitted in the form of printout with background geography as backup to the AFOS display.

B. Thunderstorm Areas

The thunderstorm probabilities are scanned with a constant threshold value of 35%, as determined from previous studies by Reap and Foster (1979). Areas with less than nine blocks on the old MDR grid are not contoured because of their small size. If no areas are identified by the 35% value, the message "THUNDER-STORM PROBABILITIES BELOW THRESHOLD" is printed across the AC chart. If either the thunderstorm or severe local storm probabilities are missing for any reason, the message "AC DATA NOT AVAILABLE" is printed across the chart.

4. VERIFICATION SCHEME

Severe local storm forecasts from the experimental AC chart were verified at NSSFC for the period April 10, 1980 to September 15, 1980. The NSSFC operational verification system for severe local storm forecast products (Weiss et al., 1980a) was used to compile objective statistics. Verification is based, in part, upon the critical success index (CSI) (Donaldson et al., 1975). The CSI is computed by dividing all weather events into four groups: x - severe storms correctly predicted, y - severe storms not predicted, z - non-severe weather predicted to be severe, and w - non-severe weather correctly predicted.

The probability of detection (POD) is the proportion of severe events correctly forecast and is given by

$$POD = x/(x+y). \qquad (1)$$

A perfect forecast has a POD of unity. The false alarm ratio (FAR) is the proportion of predictions which fail to verify, or

$$FAR = z/(x+z). \qquad (2)$$

A FAR of zero indicates a perfect forecast. The critical success index is the ratio of successful predictions to the number of severe events and false alarms as given by

$$CSI = x/(x+y+z) = [(POD)^{-1} + (1-FAR)^{-1}-1]^{-1}. \qquad (3)$$

The CSI ranges from zero to unity, with higher numbers indicating better forecasts.

The CSI, also known as the threat score, was originally devised to objectively verify point forecasts. Obviously, severe local storm outlooks cannot be treated as point forecasts because they generally cover thousands of square miles. The discrepancy between the original formulation of the CSI and the large areal extent of severe weather outlooks does not affect the computation of the POD, which is the percentage of all severe events that occur within a valid outlook. However, several serious problems are encountered when attempting to define the FAR for an area forecast. It can be shown that an inappropriate FAR (and resultant CSI) will often be computed unless both the density and the areal distribution of severe events are incorporated into the FAR calculation (Weiss et al., 1980a; Weiss et al., 1980b). Accordingly, modification of the CSI is necessary before it can be used in a meaningful fashion to verify rather isolated events over a continuous forecast region.

To compute the FAR, the outlook area is divided into MDR blocks. A report in a block is considered to be a severe occurrence for the entire block. The spatial distribution of occurrence blocks within the outlook is then computed. For verification purposes, forecast densities (Table 2) are approximated by

slight risk = 2.78% = 1/36,
moderate risk = 6.25% = 1/16, and
high risk = 11.11% =1/9.

Accordingly, an occurrence block in a slight risk forecast area verifies a 6x6 grid array of MDR blocks surrounding the report, in a moderate risk forecast area it verifies a 4x4 grid array, and in a high risk forecast area it verifies a 3x3 grid array.

Each severe report within the outlook area is considered individually, and the blocks it verifies are noted. Storm clustering is accounted for since <u>each block can be verified only once</u>. The total number of MDR blocks verified within the outlook are counted and the good area percentage (that proportion of the outlook area affected by severe storms) is computed and denoted by A. The FAR is now defined as

$$FAR = 1-A = 1-(\text{affected area/area of outlook}).$$

If the good area is nearly equal to the area of the outlook, the FAR is quite small. This occurs when severe reports have a quasi-uniform distribution across

an outlook area that had a correctly forecast density (coverage). Conversely, if very few reports occur within the outlook area, or if most reports are clustered in a relatively small portion of the outlook, only a small fraction of MDR grid blocks is verified. The good area percentage is small, resulting in a high FAR. The primary function of the areal distribution factor is to determine quantitatively the amount of overforecasting that results from the outlook areas being too large or from an insufficient density of reports.

If the good area is equal to the total area, the entire outlook is affected by severe storms. When this occurs, a further check is needed to determine if the density of severe local storms is underforecast. This consists of redefining the FAR via an empirical function which approaches unity as the degree of underforecast increases. Because of the limitations in our forecast ability and the low climatic frequency of dense severe storm occurrences, this empirical redefinition is seldom necessary.

Individual outlooks may indicate severe thunderstorms in several unconnected sections of the nation. When this occurs, separate FAR's are calculated for each area, and an area-weighted FAR is computed for the entire outlook. The averaged FAR is then used with the POD, or total percentage of all severe events within the multiple forecast areas, to calculate a single CSI via Eq. (3).

5. COMPARATIVE VERIFICATION

To determine the relative skill level of the severe local storm forecasts, a verification procedure was established to compare the experimental AC with both the conventional facimile (FAX) thunderstorm and conditional severe storm probability guidance and the operational SELS[4] convective outlook. To ensure compatibility in the comparison, some uniform standards were established.

(1) The forecast valid period must be the same. Thus, only the early morning SELS outlook (issued at 0800Z) was verified, since its 24-h valid period beginning at 1200Z is identical to the valid period of the guidance.

(2) The forecast area under consideration must be the same. Thus, verification was done only on the coarse-mesh (≈75 km on a side) MDR grid blocks for which probability guidance forecasts are produced. SELS outlook areas (or portions of outlook areas) which fall outside of the grid were omitted from the computations. Consequently, the original SELS outlook was altered 49 times (nearly one-third of all SELS outlooks) during the verification period to satisfy this criterion.

(3) The forecasts were verified for the same days. If data were missing for either guidance forecast, that day was eliminated from the sample. Guidance data were missing on 7 days during the 158-day verification period from April 10, 1980 to September 15, 1980.

(4) The forecasts were verified on the same severe local storm reports. These included reports of tornadoes, hail $\geq$2 cm (3/4 in), convective wind gusts $\geq$93 km h^{-1} (50 kt) and/or wind damage, and extreme turbulence. Verification was based on the "smooth" SELS log of severe local storms. Although the

[4]Severe local storm forecast unit at NSSFC

identifiable sources of error have been removed from the data base, some errors and biases likely remain (e.g., Galway, 1977; Kelly et al., 1978). These biases are mainly related to the influence of population density. Only reports which fall within the guidance forecast grid were included in the data sample; reports outside the grid were omitted. A total of 435 severe reports (or 9% of all reports during the verification period) were eliminated from the data base to satisfy this criterion. The large majority of the omitted reports occurred over portions of North Dakota, the western parts of South Dakota and Nebraska, and eastern Wyoming.

(5) All forecasts were categorical. Thus, the FAX probability values were transformed into categorical forecasts, similar to the SELS outlook and the experimental AC. As suggested by Foster and Reap (1978), the 6% severe local storm conditional probability contour along with the 35% thunderstorm probability contour were used by SELS to outline a severe local storm forecast area. To define the density/risk category of the forecast area, the following severe storm conditional probability thresholds were used: slight risk $\geq$6%, moderate $\geq$15%, high $\geq$30%.

It should be noted that the FAX probability guidance chart was available to the SELS forecaster prior to the issuance of the SELS outlook. However, the experimental AC chart was not available to the forecaster during the 1980 verification period and, therefore, was not used as input guidance for the SELS product.

6. VERIFICATION RESULTS[5]

During the period from April 10, 1980 to September 15, 1980, both the TDL products and the SELS outlook indicated areas of potential severe thunderstorm development (e.g., slight risk or greater) virtually the same number of times. Figure 6 shows the number of severe forecasts for all three products during each month of the verification period. Severe thunderstorms were forecast nearly every day from May through August. The high frequency of severe storm outlooks from late spring through the summer correlates well with the monthly totals of severe storm reports occurring within the TDL forecast grid, as shown in Fig. 7. Parts of May and June were unusually active tornado producers (Ostby and Wilson, 1981) and, in general, the warm season of 1980 accounted for an unusually high number of severe storm occurrences.

The monthly POD for each forecast product in shown in Fig. 8. The largest POD's occurred in June and July during the peak of the severe storm season. The FAX outlooks contained nearly 20% more reports than either the experimental AC chart (EXP) or the SELS outlook, which exhibited similar skill. The relatively high POD displayed by the FAX chart is directly related to the average area of the outlooks (Fig. 9). All three forecasts had the largest area during June, the most active severe storm month. For the entire verification period, the average SELS outlook area was slightly larger than the average EXP outlook area. However, the average FAX outlook area of $\approx$ 1,100,000 km^2 ($\approx$ 322,000 n mi^2) was more than twice the size of the EXP area.

[5]A detailed tabular listing of monthly, seasonal, etc., statistics can be found in the appendix.

When the average good area is computed (Fig. 10), the FAX outlooks again display the largest area. The SELS and FAX outlooks show similar monthly trends, reaching a peak in June. However, the EXP outlooks continue to show an increase in the average good area through the late summer.

That portion of an outlook not affected by severe storms is called the bad area, and is defined as the total area minus the good area. The average bad area in Fig. 11 indicates that the SELS and EXP outlooks were similar during most of the period, with the SELS bad area slightly larger. The FAX bad area was generally much larger, averaging nearly 686,000 km^2 (225,000 n mi^2).

A useful statistic can be obtained by forming the ratio of the bad area to the good area, called the area ratio. When the area ratio equals unity, the good area equals the bad area, implying that one-half the outlook area was affected by severe local storms. As the bad area becomes larger relative to the good area, the area ratio increases proportionately. Average area ratios (Fig. 12) show that the FAX outlooks had a consistently large area ratio during the active storm season. Overall, the FAX bad area was more than twice as large as its good area. The EXP area ratio was significantly smaller, and became even better after May. The SELS area ratio exhibited the least fluctuation (except for September), and generally was better than that for the FAX chart.

In addition to computing the previous statistical measures, we also examined the forecast coverage. Coverage is defined as the ratio of the number of high resolution (20 km on a side) MDR blocks containing severe storms within the outlook area to the total number of MDR blocks in the outlook. The average coverage shown in Fig. 13 reveals only minor monthly fluctuations, with the overall statistics appearing roughly inverse to the area statistics (Fig. 9). The coverage bias (CBIAS) is given by the ratio of the actual coverage to the forecast coverage. However, when the actual coverage falls within the categorical extremes of the forecast (e.g., moderate risk should have 6-10% coverage), then CBIAS is set to unity. If CBIAS is less than unity, the coverage was overforecast, and vice versa. The EXP chart (Fig. 14) exhibits the greatest CBIAS skill each month. SELS averages slightly better than the FAX chart over the long term, but illustrates a gradual decrease in skill after the late spring.

Finally, the average FAR and CSI are given in Fig. 15. Both TDL products show similar FAR trends, with the highest occurring during the mid-summer. The SELS FAR is comparable to the EXP chart during May and June, but is somewhat worse during the other months. Overall, the FAR of the EXP chart is significantly superior to the FAX version, with the SELS outlook demonstrating an intermediate level of skill. The CSI curves in Fig. 15 reveal only small differences through the early summer, with the EXP chart showing an increase in skill during July and August. Overall, the EXP chart had the highest average CSI.

7. DISCUSSION AND PRELIMINARY EVALUATION

A cursory examination of the CSI statistics indicates that the experimental AC chart (EXP) provides a higher degree of forecast skill than either the facsimile (FAX) guidance chart or the operational 0800Z SELS outlook. Further, the CSI results show the FAX guidance to be slightly superior to the SELS outlook during most of the verification period. While a single verification score may be desirable to evaluate different prognostic techniques, other factors must also be

investigated to insure that a realistic evaluation is made. For example, case studies allow a comparison of subjective and objective appraisals and may serve to identify possible weaknesses inherent in the verification algorithm. Further, only by examining specific attributes of the forecast techniques, such as the POD and FAR, can a detailed analysis of strengths and weaknesses be attained.

Several key attributes of the forecasts were compared on a daily basis to help determine the reliability of the prognoses. The attributes included outlook area, POD, FAR, and CSI. The comparisons determined which outlook had the largest area, largest POD, smallest FAR, and largest CSI for each valid day during the verification period. Three comparisons were made to match each forecast with the two remaining forecasts; i.e., EXP versus FAX, SELS versus FAX, and SELS versus EXP. In most cases, both outlooks in the two-way comparison either forecast severe thunderstorms or both outlooks did not. In the case where only one of the outlooks forecast severe storms, the tabulation of the FAR and CSI "winners" was based on the following criteria:

(1) The "no severe" outlook was deemed superior only when no severe thunderstorms actually occurred.

(2) The positive severe thunderstorm outlook was deemed superior only when the POD $\geq .50$ and the FAR $\leq .50$.

(3) When neither of the above criteria were satisfied, the CSI and FAR attributes were not tabulated for that day.

The results of the EXP-FAX comparison are presented in Table 3. Both products called for the chance of severe local storms nearly the same number of times during the period, although not always on the same days. The FAX outlook area and POD were both larger on a very high percentage of the comparison days. However, according to the FAR and CSI, the EXP chart demonstrated greater skill on many more days than the FAX chart. It appears that the EXP outlook has succeeded in decreasing the excessive outlook area depicted by the FAX guidance, while maintaining a reasonably high percentage of severe storm occurrences within the reduced area.

Table 3. Daily comparison of experimental (EXP) and facsimile (FAX) outlooks. Numbers in parentheses refer to percentage of days during the 151-day verification period.

	EXP	FAX
Number of severe outlooks	126 (83%)	129 (85%)
Largest area	14 (9%)	122 (81%)
Largest probability of detection	10 (7%)	94 (62%)
Smallest false alarm ratio	78 (52%)	34 (23%)
Largest critical success index	63 (42%)	50 (35%)

The results of the SELS-FAX comparison, given by Table 4, are similar to the EXP-FAX results. Again, the FAX outlook had a larger area and larger POD on more

than one-half the days. The SELS outlook had the largest area on less than one-quarter of the days. However, the FAR and CSI tabulations indicate that the SELS outlooks were superior according to these attributes more often than the FAX guidance.

Table 4. Same as Table 3, except for SELS and FAX outlooks.

	SELS	FAX
Number of severe outlooks	130 (86%)	129 (85%)
Largest area	33 (22%)	107 (71%)
Largest probability of detection	25 (17%)	84 (56%)
Smallest false alarm ratio	72 (48%)	40 (26%)
Largest critical success index	64 (42%)	50 (33%)

It is important to expand the discussion of the FAR computation in light of the large areas encompassed by a significant percentage of the outlooks. The FAR is equivalent to the ratio of the bad area divided by the total area. Given a constant FAR, the bad area is then proportional to the total area of the outlook. As the total area increases, the bad area also increases, resulting in more counties and states (and more population) being unnecessarily alerted to the threat of severe local storms. In this way, the FAR applied in the CSI formula Eq. (3) does not completely reflect the increase in the area of false alarms as the forecast area becomes larger and larger.

Recalling Fig. 15, the difference between the three forecasts does not appear that large where the FAR is examined alone. However, when translated into the bad area (Fig. 11), it is seen that the typical FAX outlook falsely alarmed an area larger than the state of Texas. The bad area associated with typical SELS and EXP outlooks was significantly less, reducing the falsely alarmed area by nearly 50% or more. A further illustration of the large degree of overforecasting by the FAX chart is given by the area ratios in Fig. 12. Even when the good area is taken into account, the area ratios indicate that the FAX outlook still falsely alarmed a much larger area than either the SELS or EXP products.

Thus, when utilizing the FAR for large area forecasts, a transformation into actual areas falsely alarmed is highly desirable, rather than only specifying a percentage of the total area. The difference in relative skills can be more easily highlighted, allowing a more definitive evaluation to be made. In this case, it is seen that the EXP chart and SELS outlook are generally superior to the FAX guidance and are more able to provide an acceptable level of guidance at the NWS field forecast level.

Table 5 presents results from the SELS-EXP comparison. On a daily basis, the SELS forecast had a larger outlook area more than one-half the time, while the EXP chart produced a larger area about one-third the time. The POD tabulation was quite similar, with the SELS outlook containing the most reports more often than the EXP chart. According to both the FAR and CSI, the EXP chart was superior on over 40% of the days, with the SELS product better on nearly one-third the days. This daily breakdown indicates the new EXP chart comparable to, and occasionally exceeding, the skill level of the 0800Z SELS outlook.

Table 5. Same as Table 3, except for SELS and EXP outlooks.

	SELS	EXP
Number of severe outlooks	130 (86%)	126 (83%)
Largest area	86 (57%)	54 (36%)
Largest probability of detection	56 (37%)	49 (32%)
Smallest false alarm ratio	46 (30%)	64 (42%)
Largest critical success index	48 (32%)	65 (43%)

One last attribute was examined to aid in the evaluation. The highest risk category assigned to each outlook was determined for each day during the verification period, as shown in Table 6. During this time, forecasts of no severe thunderstorms occurred approximately the same percentage of days for all outlooks, with the EXP chart failing to reach the severe threshold slightly more often than the FAX or SELS outlooks. The SELS and FAX forecasts also called for a slight risk and a moderate risk of severe thunderstorms in about the same proportion. Note, however, the anomaly in the EXP forecasts for the slight and moderate risk categories. The EXP chart forecast moderate risk relatively few times, opting instead to call for a slight risk on most outlooks. Unlike the SELS and FAX outlooks, the EXP chart failed to reach the moderate risk threshold after June 29, 1980.

Table 6. Highest daily severe local storm risk category selected by SELS, FAX, and EXP outlooks.

	SELS	FAX	EXP
No severe	21	22	25
Slight risk	85	87	109
Moderate risk	45	40	16
High risk	0	1	2

An examination of active severe storm days reveals an apparent low bias in the risk category selection of the EXP chart. By arbitrarily defining an active severe storm day as one having 50 or more severe reports within the TDL forecast grid, we identified 31 days meeting this criterion. On these active days, the EXP chart forecast the slight risk category 20 times (nearly two-thirds of the active days), compared to 11 and 12 times for the FAX and SELS outlooks, respectively. Seventeen of the active days occurred during the period June 15-September 15 when the summer severe storm probability equations were run. During this period, the EXP chart did not reach the moderate risk threshold on any of these 17 active storm days. Conversely, both the SELS and FAX outlooks called for a moderate risk of severe thunderstorms on about one-half the active summer days.

A comparison of the highest risk category selected for each outlook and the actual coverage obtained within the outlook was performed for each of the 31 active days. It was then determined whether the coverage of severe storms was underforcast, overforecast, or correctly forecast. Results in Table 7 indicate

that all three products correctly forecast the proper risk category on nearly one-half the active days. On the remaining days, the FAX chart greatly overforecast the coverage, while the EXP chart generally underforecast the coverage. No consistent bias was found in the coverage forecast for the SELS outlooks. Examination of the coverage bias (CBIAS), previously described, confirms the preceding analysis. The average CBIAS during the 31 active days was FAX=0.85, EXP=1.21, and SELS=1.06.

Table 7. Coverage forecast skill on active severe storm days by SELS, FAX, and EXP outlooks.

Coverage forecast	SELS	FAX	EXP
Correct	15	14	12
Underforecast	7	1	15
Overforecast	9	16	4

The overforecast bias in coverage for the FAX chart that is evident on the active storm days is undoubtedly related to the very large outlook areas. The cause of the underforecast bias in coverage for the EXP chart, however, may be related to an underestimation of coverage that occurs during the generation of the automated categorical severe storm forecasts. (Recall that an empirical assumption is used when converting the coverage from the original coarse MDR grid to the current high-resolution grid.)

Finally, it should be noted that a bias in the risk category selection may have a direct effect on the verification statistics. The good area algorithm is based upon proportionately smaller MDR grid arrays being affected by severe storms as the risk category increases. For a typical outlook, the larger grid array associated with the slight risk category will result in a larger good area computation when compared with a moderate (or high) risk outlook outlining the same severe storm occurrences. Consequently, a smaller FAR and larger CSI will result. This effect may not occur on individual days when dense severe storm activity occurs over a wide area. However, for time periods longer than several weeks, long-term statistics will likely indicate greater prognostic skill when the slight risk category is used exclusively when severe storms are expected. This may partially explain the apparent increase in skill demonstrated by the EXP chart in July and August, as given by the good area in Fig. 10 and the CSI in Fig. 15.

To examine this possibility in more detail, the SELS verification was re-run during the period July 1 through September 15 with all severe outlooks calling for a slight risk of severe thunderstorms. This was comparable to the EXP outlook which forecast only the slight risk category during that period. The recalculation resulted in a moderate improvement in the overall level of skill. The SELS good area increased 15%, the CBIAS increased 14%, the FAR and area ratio decreased 6% and 20% respectively, and the CSI increased 12%. Although the EXP statistics still indicated greater skill, the magnitude of the difference was narrowed significantly. It appears that the higher skill level demonstrated by the EXP chart, especially during the summer months when an underforecast of the

coverage is prevalent, is partially a result of a bias in the verification algorithm.

In summary, the experimental AC chart was found to provide a significant level of prognostic skill in medium-range forecasts of severe local storms. Based on its performance during the 1980 warm season, the new guidance can be expected to reduce the overly large outlook areas exhibited by the FAX probability guidance while retaining a high number of severe local storm reports, thus maintaining a respectable probability of detection (POD) while substantially reducing the false alarm ratio (FAR). Verification statistics indicate that the skill level of the experimental AC often exceeds that of both the conventional FAX outlook guidance and the operational 0800Z SELS outlook. Except for an underforecast bias in selecting the risk category, especially during the summer months, the experimental AC guidance appears to be reliable. Consistent with previous studies (e.g., Weiss, 1977; Pearson and Weiss, 1979), a greater degree of skill was exhibited by all three forecast products on active days with numerous tornadoes (see appendix).

Beginning April 15, 1981, NSSFC has been receiving additional daily guidance pertaining to the experimental AC. This additional package includes linearized fields of the five leading predictors in the conditional severe storm probability equation. Accordingly, SELS forecasters should be able to understand more explicitly the basis for the severe thunderstorm guidance and be able to use the guidance more effectively.

The recent history of numerical weather prediction and the associated MOS products reveals a link between improved objective guidance and more skillful man-made forecasts (e.g., Shuman, 1978; Charba and Klein, 1980). Owing to the availablility of the new medium-range severe storm guidance for operational use during the 1981 warm season, the prospects for improving the 0800Z SELS outlook appear favorable.

REFERENCES

Charba, J. P., and W. H. Klein, 1980: Skill in precipitation forecasting in the National Weather Service. Bull. Amer. Meteor. Soc., 61, 1546-1555.

Donaldson, R. J., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. Preprints Ninth Conference on Severe Local Storms, Norman, Amer. Meteor. Soc., 321-326.

Foster, D. S., and R. M. Reap, 1978: Comparative verification of the operational 24-h convective outlooks with the objective severe local storm guidance based on model output statistics. TDL Office Note 78-7, National Weather Service, NOAA, U.S. Department of Commerce, 17 pp.

Galway, J. G., 1977: Some climatological aspects of tornado outbreaks. Mon. Wea. Rev., 105, 447-484.

Kelly, D. L., J. T. Schaefer, R. P. NcNulty, C. A. Doswell III, and R. F. Abbey, 1978: An augmented tornado climatology. Mon. Wea. Rev., 106, 1172-1183.

National Weather Service, 1980: SELS Severe weather outlook (AC). NWS Technical Procedures Bulletin No. 279, NOAA, U.S. Department of Commerce, 5 pp.

Ostby, F. P., and L. F. Wilson, 1981: The tornado season of 1980. Weatherwise, 34, 26-32.

Pearson, A., and S. J. Weiss, 1979: Some trends in forecast skill at the National Severe Storms Forecast Center. Bull. Amer. Meteor. Soc., 60, 319-326.

Reap, R. M., and D. S. Foster, 1979: Automated 12-36 hour probability forecasts of thunderstorms and severe local storms. J. Appl. Meteor., 18, 1304-1315.

Shuman, F. G., 1978: Numerical weather prediction. Bull. Amer. Meteor. Soc., 59, 5-17.

Weiss, S. J., 1977: Objective verification of the severe weather outlook at the National Severe Storms Forecast Center. Preprints Tenth Conference on Severe Local Storms, Omaha, Amer. Meteor. Soc., 395-402.

_____, D. L. Kelly, and J. T. Schaefer, 1980a: New objective verification techniques at the National Severe Storms Forecast Center. Preprints Eighth Conference on Weather Forecasting and Analysis, Denver, Amer. Meteor. Soc., 412-419.

_____, C. A. Doswell III, and F. P. Ostby, 1980b: Comments on "Automated 12-36 hour probability forecasts of thunderstorms and severe local storms." J. Appl. Meteor., 19, 1328-1333.

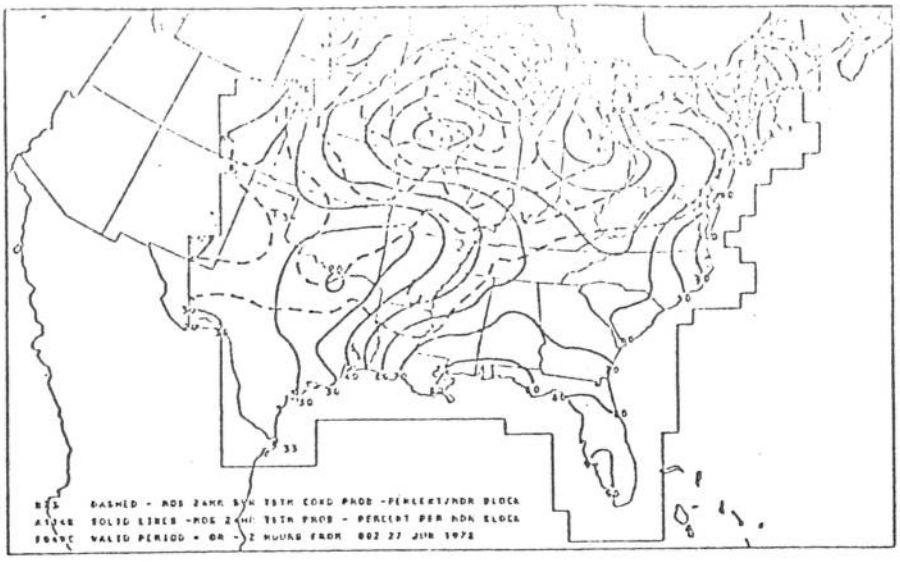

Figure 1. Computer-drawn map of thunderstorm probability (solid) and conditional probability of severe local storms (dashed). The probabilities are valid during the 12-36 h interval following 0000 GMT initial time.

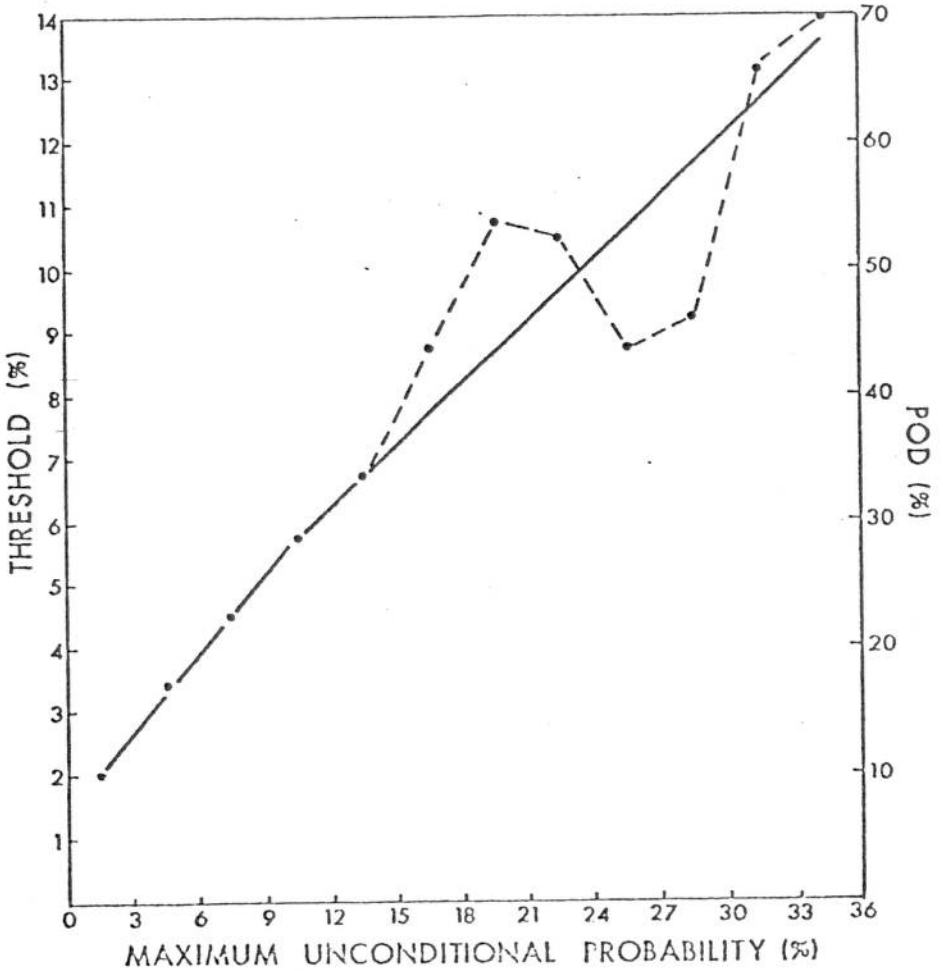

Figure 2. Threshold values that give the maximum CSI for each of 12 categories of severe local storm probability. Actual data (dashed) are shown along with linear fit (solid). Variability of dashed curve reflects the small number of cases in the higher probability categories.

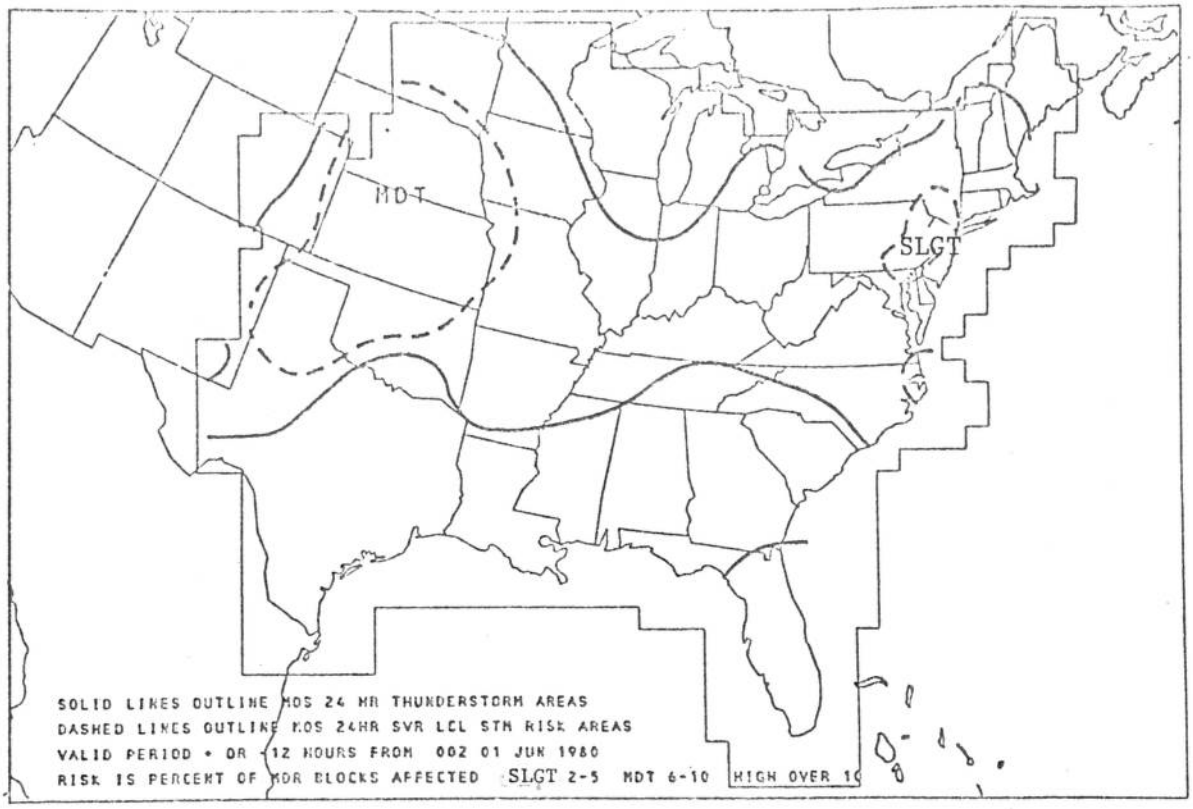

Figure 3. Sample automated AC chart showing severe local storm risk areas (dashed) and expected boundaries of general thunderstorm activity (solid).

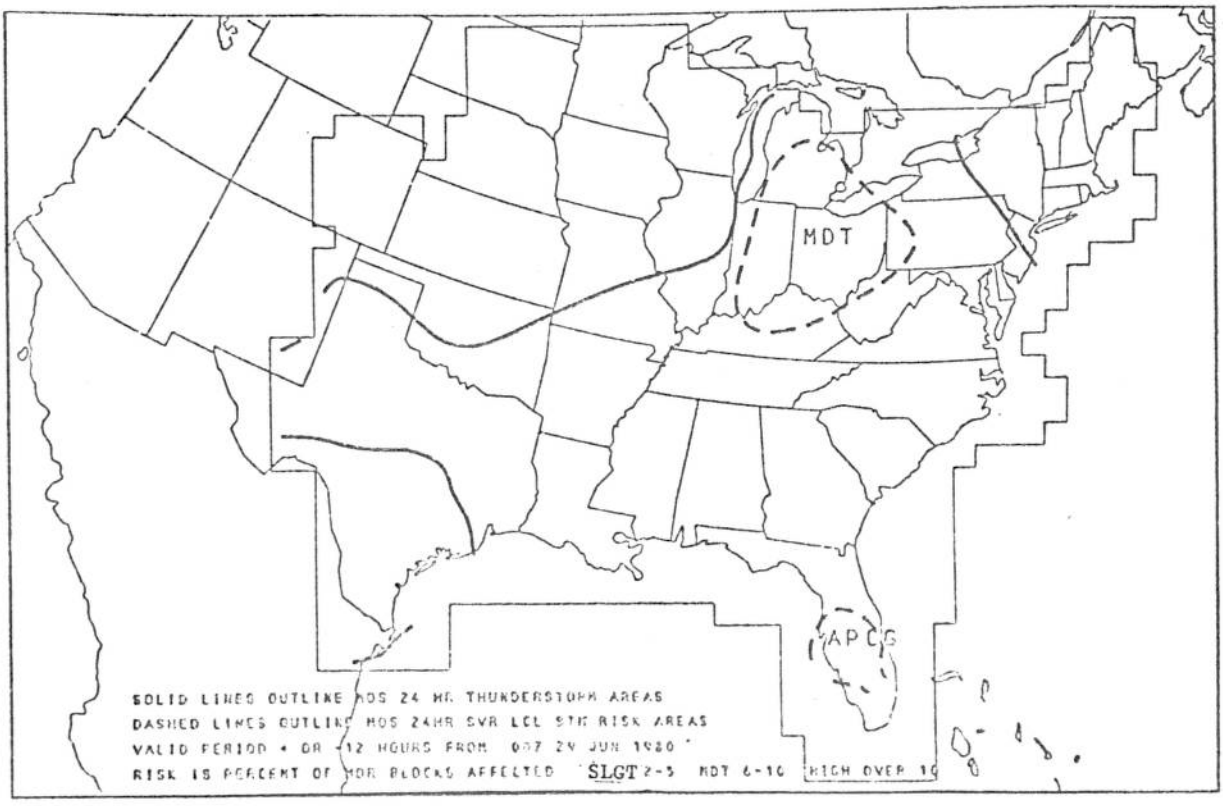

Figure 4. Same as Fig. 3, except for a different day.

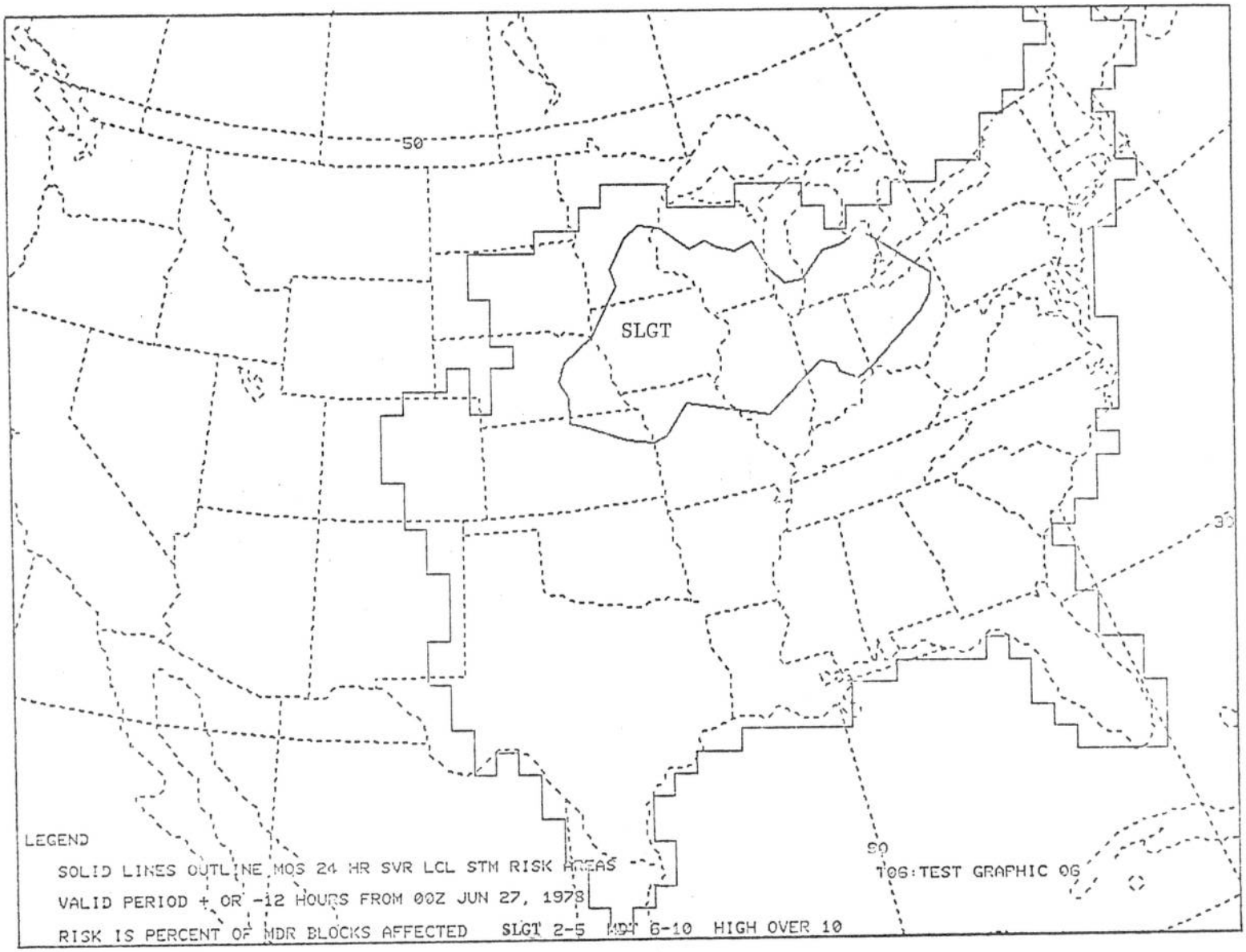

Figure 5. Sample automated AC chart as displayed on AFOS. Severe local storm risk area is given by solid line. Thunderstorm boundaries are displayed on a separate chart, not shown.

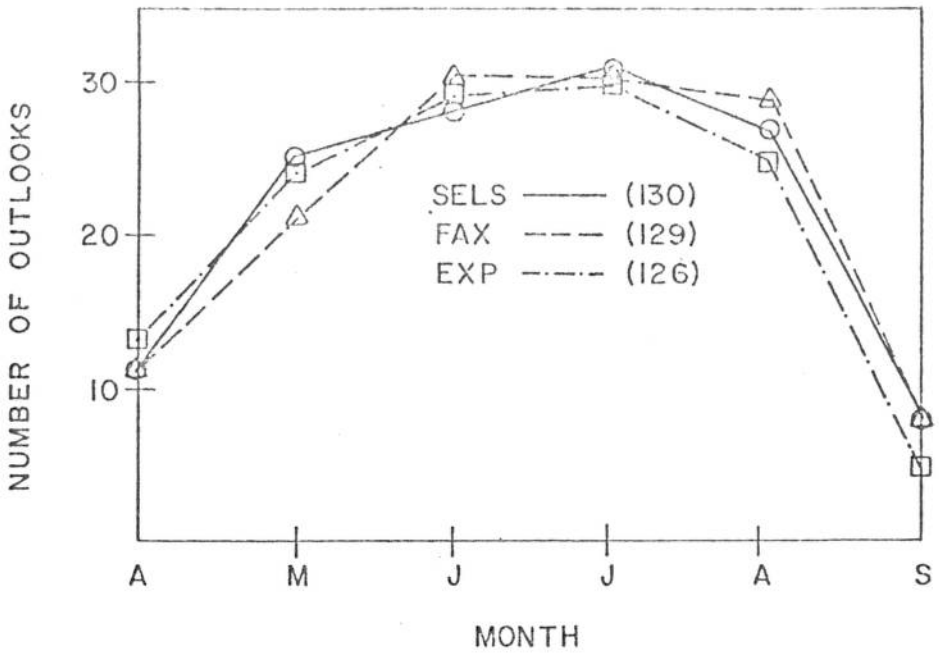

Figure 6. Number of severe storm outlooks of slight risk or greater issued by SELS (solid), facsimile (FAX - dashed) and experimental (EXP - dot-dashed). Numbers in parentheses refer to total number of severe outlooks issued during verification period.

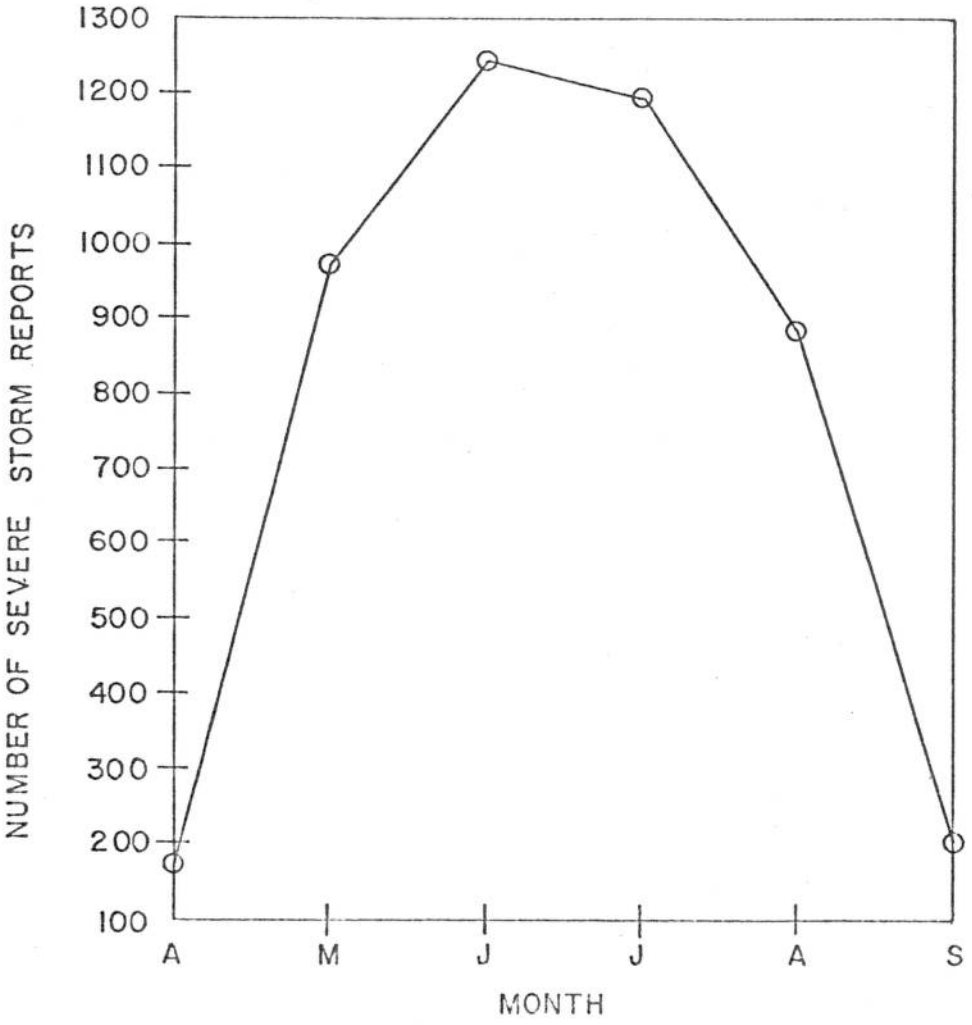

Figure 7. Number of severe local storm reports occurring within TDL forecast grid.

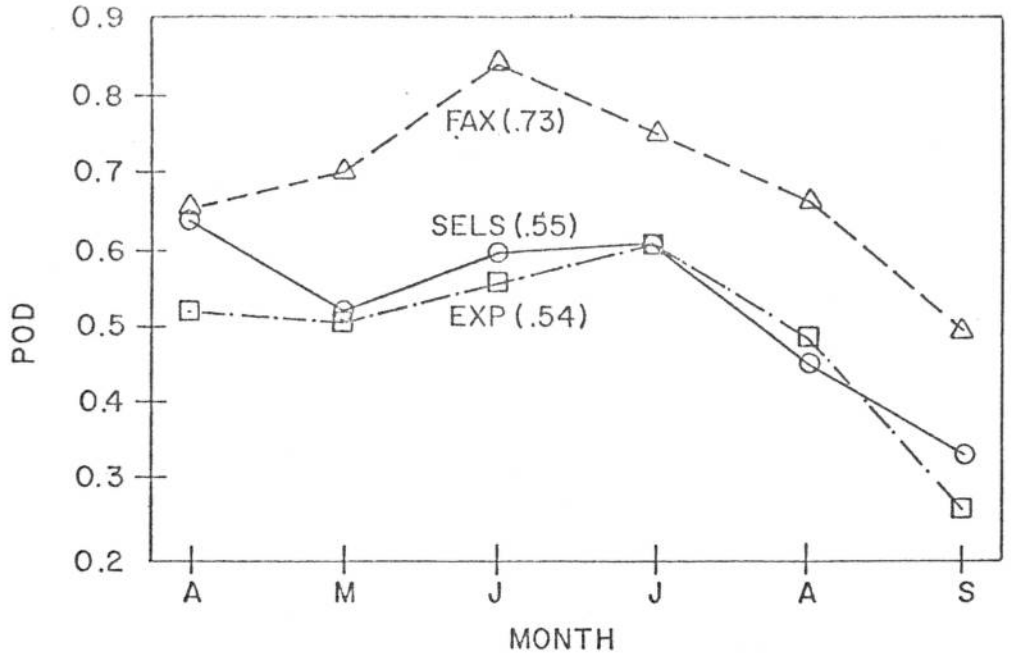

Figure 8. Average probability of detection for SELS (solid), FAX (dashed), and EXP (dot-dashed) outlooks. Numbers in parentheses refer to overall POD during verification period.

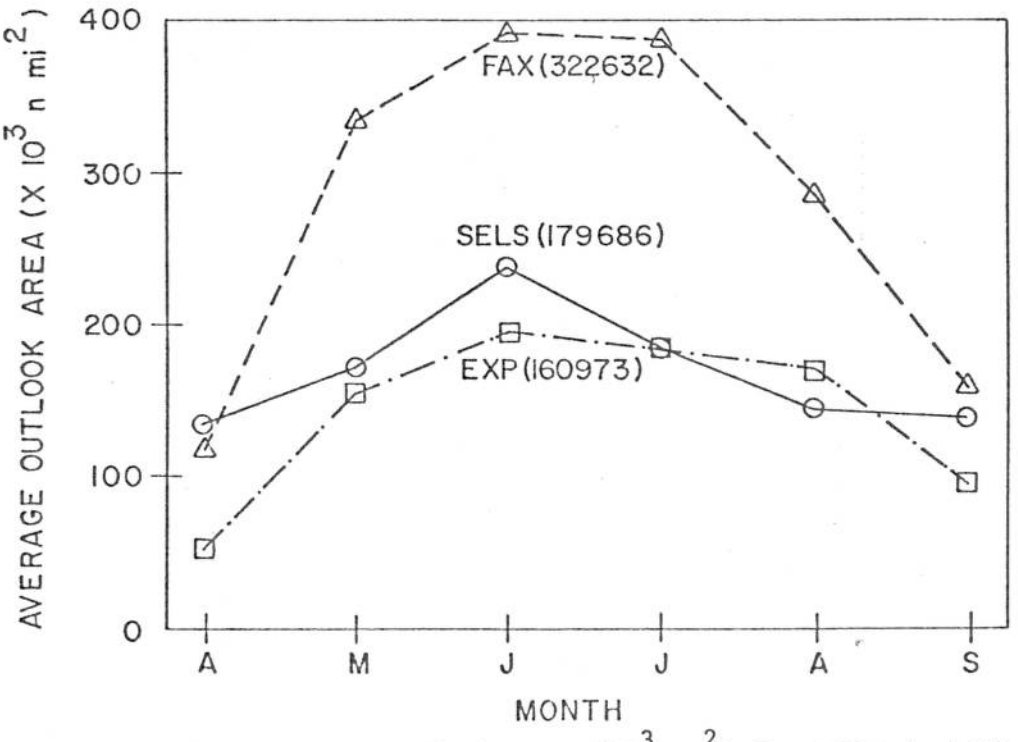

Figure 9. Average outlook area (10^3 nm^2) for SELS (solid), FAX (dashed), and EXP (dot-dashed) outlooks. Numbers in parentheses refer to overall average area during verification period.

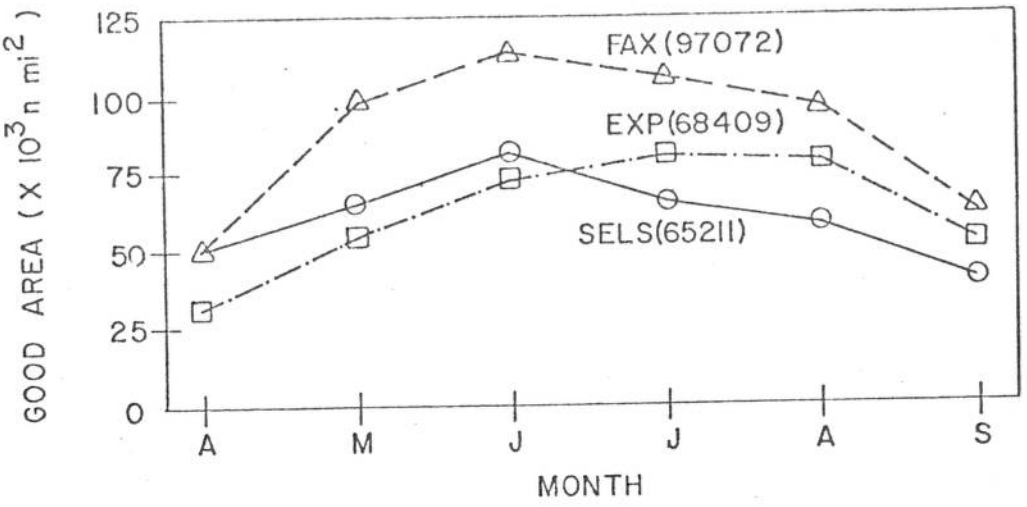

Figure 10. Same as Fig. 9, except for average good area.

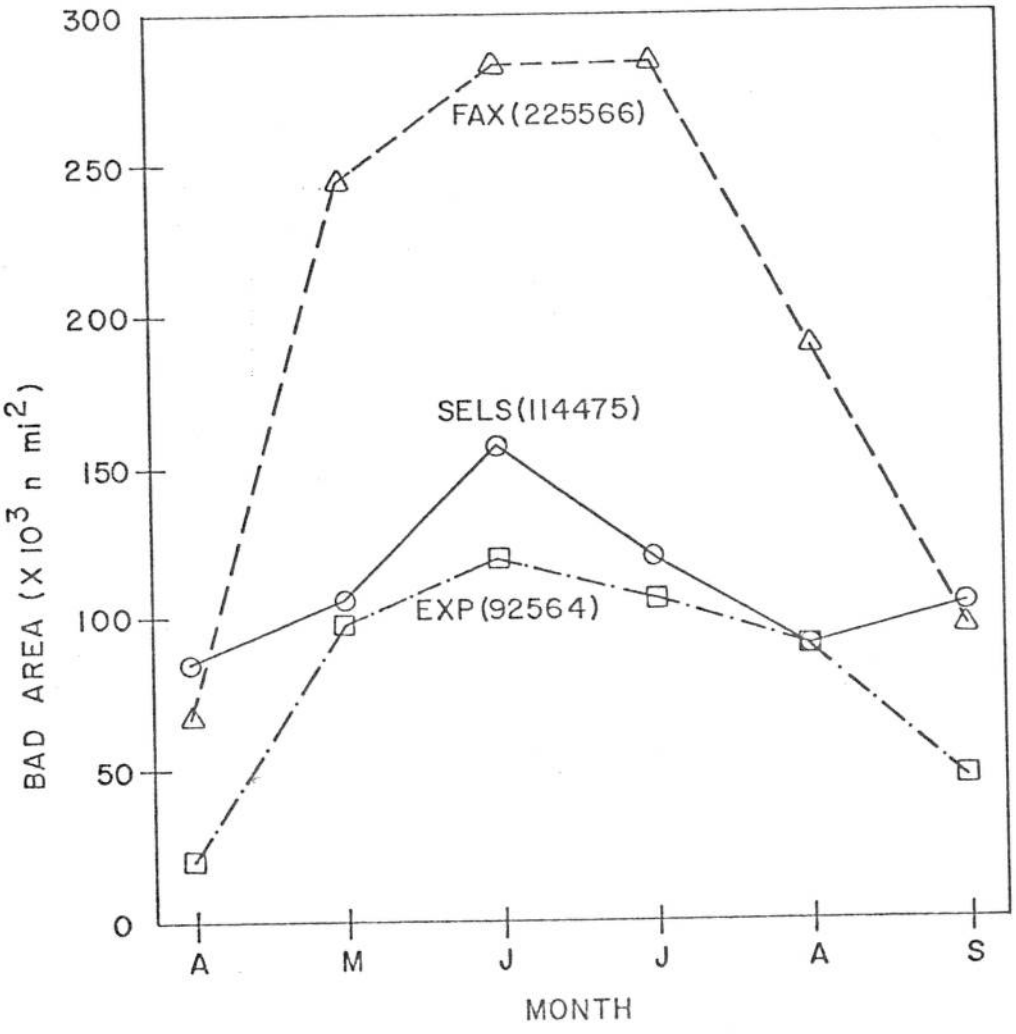

Figure 11. Same as Fig. 9, except for average bad area.

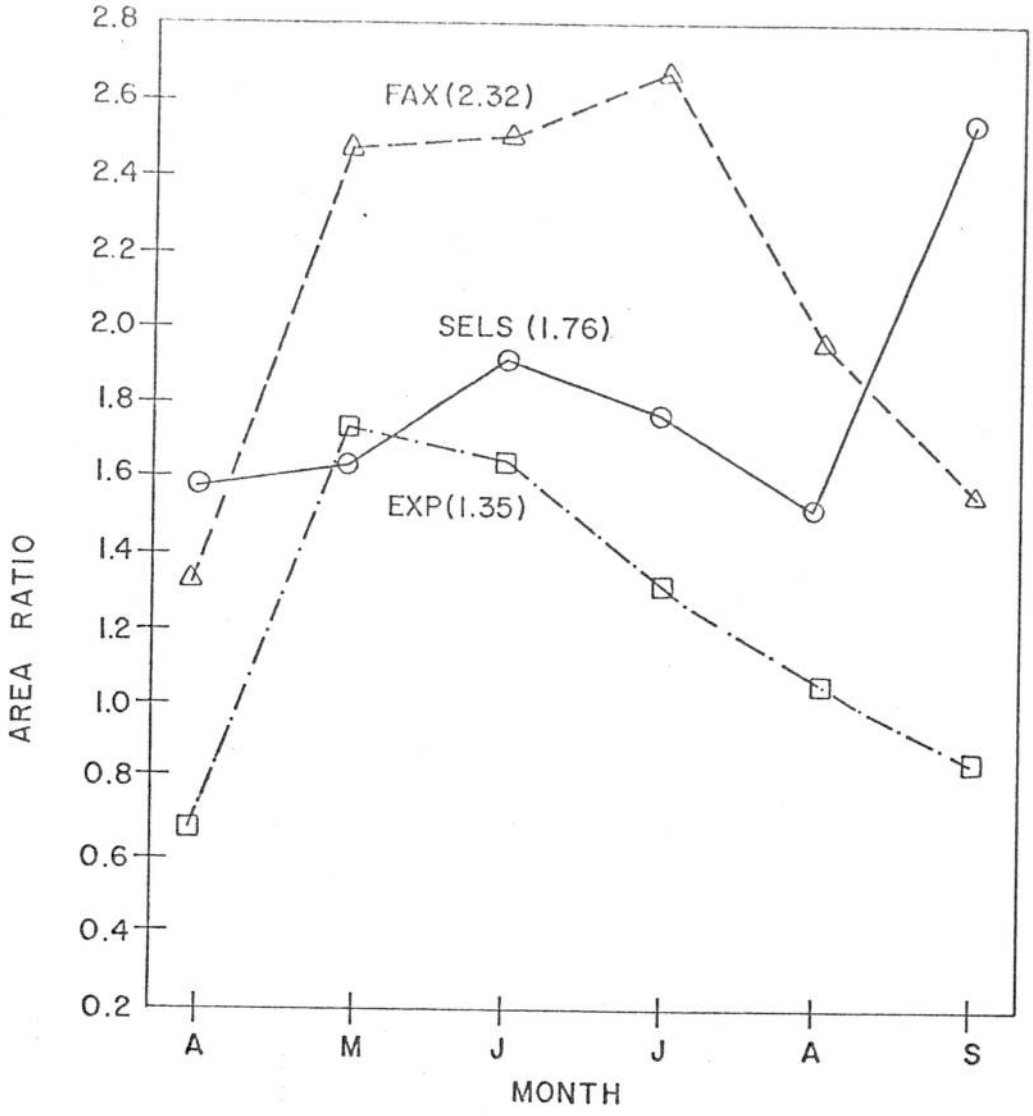

Figure 12. Same as Fig. 9, except for average area ratio.

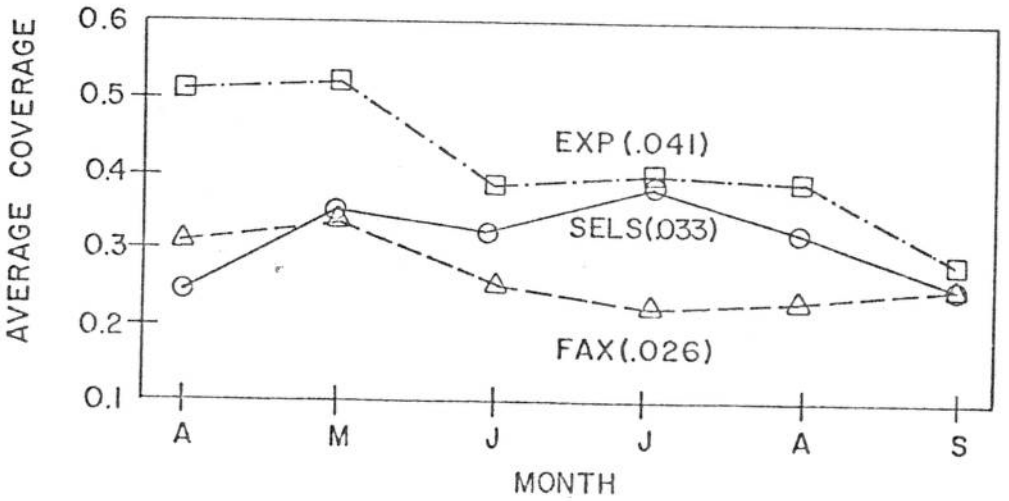

Figure 13. Same as Fig. 9, except for average coverage.

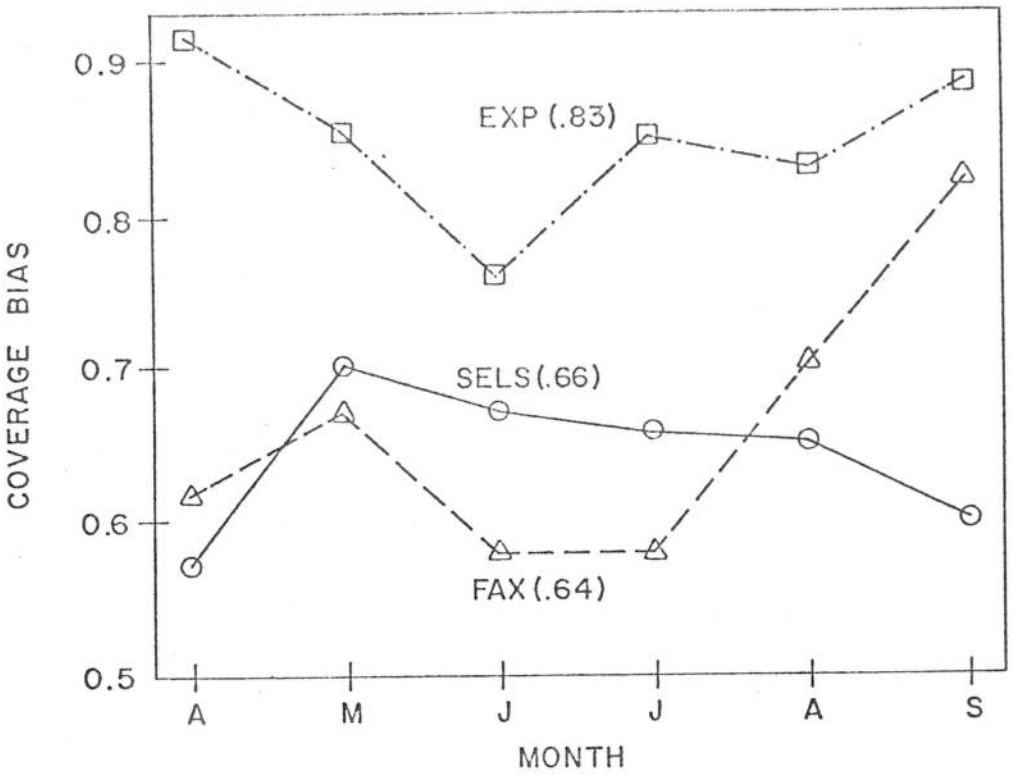

Figure 14. Same as Fig. 9, except for average coverage bias.

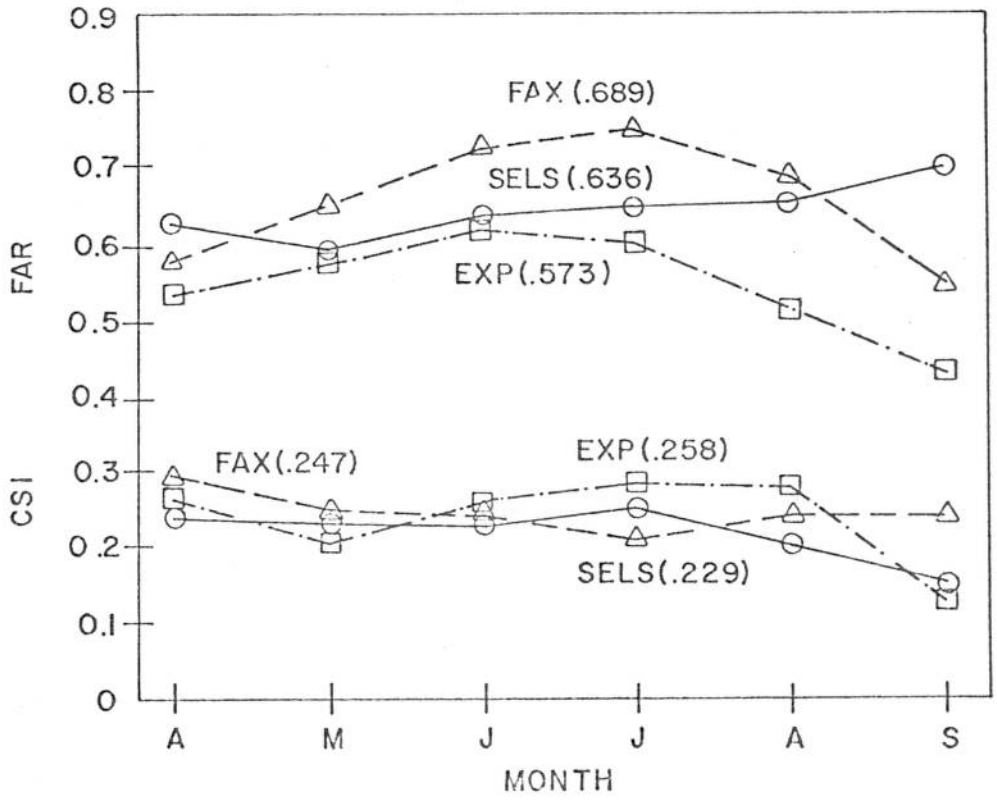

Figure 15. Same as Fig. 9, except for average false alarm ratio (top) and average critical success index (bottom).

APPENDIX

Monthly and seasonal verification statistics for SELS, FAX, and EXP convective outlooks prepared for period April 10, 1980 to September 15, 1980. Statistics also given for active severe storm days with 10 or more tornadoes.

	NBR OTLKS	AVG AREA	HITS	RPTS	POD	CVRG	CBIAS	FAR	GOOD AREA	BAD AREA	AREA RATIO	CSI
APR												
SELS	11	134752	111	173	.64	.024	.573	.632	51964	82788	1.59	.241
FAX	11	120372	113	173	.65	.031	.616	.583	51750	68621	1.33	.295
EXP	13	53478	91	173	.52	.051	.917	.535	31817	21661	0.68	.272
MAY												
SELS	25	173252	511	979	.52	.035	.697	.594	65957	107295	1.63	.230
FAX	21	341624	695	979	.70	.034	.670	.653	98508	243115	2.47	.245
EXP	24	154421	501	979	.51	.052	.850	.579	56276	98146	1.74	.214
JUN												
SELS	28	238408	749	1239	.60	.032	.668	.632	81450	156958	1.93	.234
FAX	30	396247	1043	1239	.84	.026	.579	.726	113069	283178	2.50	.244
EXP	29	192727	703	1239	.56	.038	.759	.620	73115	119612	1.64	.258
JUL												
SELS	31	184807	726	1187	.61	.039	.657	.646	66618	118189	1.77	.250
FAX	30	389121	902	1187	.75	.022	.579	.755	105944	283177	2.67	.212
EXP	30	185014	735	1187	.61	.040	.851	.606	80102	104912	1.31	.287
AUG												
SELS	27	148247	412	880	.46	.032	.650	.652	58789	89458	1.52	.206
FAX	29	285714	590	880	.67	.023	.702	.687	96934	188781	1.95	.247
EXP	25	170217	423	880	.49	.033	.831	.520	82764	87453	1.06	.279
SEP												
SELS	8	142350	68	197	.34	.025	.597	.698	40475	101875	2.52	.144
FAX	8	159403	99	197	.50	.031	.823	.551	62860	96544	1.54	.241
EXP	5	97311	55	197	.27	.033	.885	.438	52563	44747	0.85	.137
TOTALS												
SELS	130	179686	2577	4634	.55	.033	.655	.636	65211	114475	1.76	.229
FAX	129	322632	3442	4634	.73	.026	.640	.689	97072	225560	2.32	.247
EXP	126	160973	2508	4634	.54	.041	.834	.573	68409	92564	1.35	.258
SPRING												
SELS	48	191755	1192	1924	.61	.036	.717	.597	72926	118829	1.63	.257
FAX	46	303359	1520	1924	.79	.033	.633	.665	90494	212865	2.35	.267
EXP	50	136038	1103	1924	.57	.052	.875	.573	55177	80861	1.47	.255
SUMMER												
SELS	82	172625	1385	2731	.50	.031	.618	.659	60694	111930	1.84	.212
FAX	83	333317	1922	2731	.70	.023	.643	.702	100717	232600	2.31	.236
EXP	76	177380	1405	2731	.51	.034	.806	.573	77114	100266	1.30	.260
DAYS WITH ≥ 10 TORNADOES												
SELS	18	267821	1032	1439	.71	.065	1.000	.500	123929	143892	1.16	.373
FAX	16	472583	1223	1439	.84	.052	0.876	.581	174710	297873	1.70	.361
EXP	17	211426	915	1439	.63	.083	1.246	.440	107704	103722	0.96	.378