AD-A045 133 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 4/2
A STATISTICAL ANALYSIS OF WESTERN PACIFIC TROPICAL CYCLONE FORE--ETC(U)
JUN 77 D S NICKLIN

UNCLASSIFIED NL

1 OF 1
AD
A045133

END
DATE
FILMED
11 -77
DDC

NAVAL POSTGRADUATE SCHOOL

Monterey, California

(9) Master's **THESIS**

(6) A STATISTICAL ANALYSIS OF WESTERN PACIFIC TROPICAL CYCLONE FORECAST ERRORS

by

(10) Donald Stevens/Nicklin

(11) June 1977

(12) 51p.

Thesis Advisor: J. D. Jarrell
Thesis Advisor: W. van der Bijl

251 450

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Statistical Analysis of Western Pacific Tropical Cyclone Forecast Errors		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; June 1977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Donald Stevens Nicklin		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE June 1977
		13. NUMBER OF PAGES 51
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Western Pacific tropical cyclone position forecast errors for 10 years (1966-1975) are statistically analyzed. Variations of errors versus a dozen parameters are examined and the trends over the 10 years are discussed. Discriminant analysis techniques were used to isolate categories where forecasts were likely to be above and below the median in East-West and North-South error components. The discriminant analysis was tested on 1976 data and the results are presented. It was confirmed that a small number of next page

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

 SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) readily available parameters, such as location, maximum winds, and speed of movement, can, with reasonable effectiveness, classify a tropical cyclone forecast as representing a group either markedly above or below average errors.

ACCESSION for		
NTIS	White Section	X
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL. and/or SPECIAL	
A		

Approved for public release; distribution unlimited.

A Statistical Analysis of Western Pacific Tropical Cyclone Forecast Errors

by

Donald Stevens Nicklin
First Lieutenant, United States Air Force
B.S., Saint Louis University, 1973

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
June 1977

Author Donald S. Nicklin

Approved by: Jerry D. Jarrell
Thesis Advisor

Wm van der Bijl
Thesis Advisor

G. J. Haltiner
Chairman, Department of Meteorology

R. R. Fossum
Dean of Science and Engineering

ABSTRACT

Western Pacific tropical cyclone position forecast errors for 10 years (1966-1975) are statistically analyzed. Variations of errors versus a dozen parameters are examined and the trends over the 10 years are discussed. Discriminant analysis techniques were used to isolate categories where forecasts were likely to be above and below the median in East-West and North-South error components. The discriminant analysis was tested on 1976 data and the results are presented. It was confirmed that a small number of readily available parameters, such as location, maximum winds, and speed of movement, can, with reasonable effectiveness, classify a tropical cyclone forecast as representing a group either markedly above or below average errors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author would like to thank the many people from the various departments of the Naval Postgraduate School, and the Naval Environmental Prediction Research Facility, who contributed to the completion of this thesis. In particular, to Mr. Sam Brand for the timely acquisition of the data, to the staff of the W. R. Church Computer Center for their complete and generous cooperation, and to Capt. J. D. Shewchuk for his assistance with the finer points of computer programmi[illegible] And finally, my sincerest thanks to CDR. J. D. Jarrell for his [illegible]ss guidance, support, and patience; and the generous contribution of his experience to this project.

I. OBJECTIVES

As discussed in a request from the Director, Joint Typhoon Warning Center, Guam (JTWC) to the Naval Environmental Prediction Research Facility, Monterey (NEPRF), a need exists for a statistical analysis of the JTWC tropical cyclone forecasts to discover the existence of any significant trends. More specifically, the long range goals number four:

1. To identify situations where the forecasts are very good or very bad, to allow maximum concentration of resources for quick reduction of the largest errors.

2. To provide probability algorithms for an estimate of the forecast errors of warnings, to assist Western Pacific commanders in operational decisions regarding the protection and/or evacuation of military resources.

3. To stratify errors for 24- through 72-hour forecasts, based on various parameters such as location, time of year, speed of movement, intensity, and synoptic patterns.

4. To determine if the year to year variations in forecast accuracies for the 10 year period are real, or random deviations about a long term mean.

The more immediate short range goals of this research, as a first step toward the realization of the long range objectives, are:

1. To check the data for errors in recording, and to test it for reliability as a data base for statistical study.

2. To assemble and consolidate the data into a usable format.

3. To determine basic statistical relationships between parameters.

4. To manipulate the basic data, to create a set of parameters for further study of errors in the 24-, 48-, and 72-hour typhoon forecasts.

5. To perform discriminant and stepwise multiple linear regression analyses, to find parameters related to the forecast errors.

6. To summarize the results and test them by a preliminary application to 1976 data.

7. To make recommendations as to the direction for continued research toward the long range goals.

II. DESCRIPTION OF THE DATA

The JTWC Western Pacific tropical cyclone forecasts and best tracks were matched by date/time groups for 10 years as follows:

1966 - 34 storms

1967 - 35

1968 - 26

1969 - 23

1970 - 26

1971 - 35

1972 - 31

1973 - 21

1974 - 34

1975 - 25 storms

The term "storm" used herein refers collectively to tropical cyclones (tropical depressions, tropical storms, and typhoons) without regard to intensity. The 10 year total was 290 storms, or an average of 29 storms per year. The total number of best track positions at six-hourly intervals was approximately 6150.

In the process of matching forecasts with best tracks of the same time, the data was checked for errors and corrected, when necessary, using annual typhoon reports. Rarely was a report garbled beyond correction so as to require removal from the data. There were some storms, however, that were so short-lived as to provide no verifying forecasts. These were not represented in the verified case data that was statistically analyzed. For instance, to verify a 48-hour forecast and compute an error distance for the forecast, a best track must be available 48 hours later. If, in that 48-hour period, the storm dissipated and was no longer identified by a best track position, the forecast could not be verified. This also accounts for the fact that fewer cases were verified for 72-hour forecasts than for 48-hour, and fewer 48-hour than 24-hour forecasts. These cases numbered as follows:

4809 24-hour forecasts

3038 48-hour forecasts

1372 72-hour forecasts

As a minimum the following parameters were known for each case at forecast initiation time:

1.	Maximum Wind	MAX WIND
2.	Latitude	LAT
3.	Longitude	LONG
4.	West-East Component of Storm Movement	U MOVT
5.	South-North Component of Storm Movement	V MOVT
6.	Position Number on Storm Track	POS NO
7.	Number of Storms in Progress at Forecast Time	NO STM
8.	Month	MONTH
9.	Time-GMT of the Forecast	TIME
10.	Error Distance (Nautical Miles)	ERR DIS
11.	Direction from Verifying Position to Forecast Position	ERR DIR

The 1976 data was processed in a similar way, but retained separately for testing. There were 625 best track positions at six-hourly intervals from 25 storms and the verifying cases totaled:

524 24-hour forecasts
424 48-hour forecasts
332 72-hour forecasts

III. STATISTICAL ANALYSIS

A. BASIC STATISTICS

As a preliminary consideration, evidence of a climatic trend in the 10 year data was investigated. The number of occurrences of tropical cyclones in a fixed time (1 year) can be assumed to follow a Poisson distribution if two plausible conditions exist: (a) an occurrence is just as likely in one interval as another, and (b) the occurrence of an event has no effect on whether or not another occurs. A property of this distribution requires the population variance to equal the population mean. In this 10 year sample, the variance is 29.89, and the mean is 29.00. If a climatic change were occurring, then the sample variance should exceed the sample mean. As it does not, no climatic trend is evident within the 10 year sample.

The initial statistical analysis of the variables employed the UCLA Biomedical computer program BMD02R stepwise multiple linear regression (Dixon, 1970). Tables I and II summarize the means, standard deviations, variance explained, and the correlation matrix of the first 10 variables for the 24-, 48- and 72-hour forecasts. No correlation coefficient of any available predictor with the magnitude of the errors (at either 24, 48, or 72 hours) exceeded 0.185, and the total explained variance of the error distance did not exceed 11%. It was noted, however, that the variables contributing most to the explained variance were MAX WND, LAT, LONG, U MOVT, AND V MOVT. The concept of predicting the error directly was abandoned.

From these basic statistics, the average Western Pacific tropical cyclone moved from 19°N latitude, 135°E longitude to the WNW at seven knots with maximum winds of about 66 knots in late August.

B. MEAN ERROR STRATIFICATIONS

For more detail, the forecast errors were stratified and mean errors were computed for each stratification along the range of each variable. The distance to the nearest storm (STM DIS), for multiple storm cases, and initial position error (POS ERR) were added as variables. Significant trends were evident (Figures 1-6). In the figures, stratifications were selected to keep the number of cases in each group relatively high. Frequencies are typically a few hundred and are not indicated except where they drop below 100. As an aid in interpreting these figures, relative frequencies, all based on 4809 cases, are shown in Figures 1-6 for the 24-hour forecasts. Since there are in percentages of the total, they roughly apply also for 48- and 72-hour forecasts.

In each of the 24-, 48-, and 72-hour forecast situations, the mean forecast errors were minimal for lower latitudes (Figure 1a), gradually increasing with latitude. This indicates that storms are more accurately forecast before they recurve and move into higher latitudes. It is also consistent with a study by Sadler (1967), which distinguished between (a) storms originating in the vicinity of the mean August surface trough, between 5° and 20°N latitude, that moved mostly to the west; and (b) those beginning north of 20° N latitude, in the vicinity of the mean August Tropical Upper Tropospheric Trough (TUTT), which were more erratic with predominantly northerly components. Both of these concepts support smaller forecast errors for storms in lower latitudes.

Mean forecast errors decrease with decreasing east longitude (Figure 1b), or more westerly positions. Generally, a forecast for a storm in a more westerly position is one based on a longer than average history, perhaps in an area of better synoptic data coverage, given the proximity to the Philippines, Taiwan, and China, and other continental areas west of 130°E. Additionally, land radar enhances accuracy of location. Storms west of 130°E are less susceptible to large forecast errors when land is nearby to the north of the track, because storms unexpectedly recurving over China dissipate rapidly and hence are not generally reflected in forecast verification.

Maximum wind (Figure 2) is another important parameter. The mean errors decrease with increasing maximum wind speeds, indicating that better developed storms, again with longer histories and more accurate center locations, are more accurately forecast. This trend is visible for all three forecast times, with some increasing fluctuations and irregularities in the 48-, and 72-hour forecasts, which are based on progressively smaller sample sizes.

Relative to the U component of storm movement (Figure 3a), forecasts are generally better for storms moving west and becoming progressively more difficult as westward movement diminishes and becomes eastward as associated with recurvature. For the V component (Figure 3b), the best forecasts are centered at or near zero, again implying better forecasts when the storm is moving west with little or no deflection north or south. Errors increased markedly for storms moving south with any component, or to the north, as might be associated with the recurvature process.

Time of day (Figure 4b) showed no perceivable relationship with forecast errors at any of the forecast times. For the years 1969-1971 forecasts were issued at 0500 GMT plus every six hours, while in the remaining years forecasts were issued at 0000 GMT plus every six hours. For no obvious reason, forecasts in those three years appear to have been superior to forecasts issued at the more normal synoptic times. Time of year, or stratification by month (Figure 4a), did show that larger errors occurred with the largest freqeuncy of storms in late summer to fall. There is a consistent improvement in April for all three forecasts, but since this is based on only 5% of the cases, its significance is somewhat dubious. The factors of workload and personnel turnover seem to be reflected in Figure 4a. Most personnel changes occur in the spring to early summer months as the frequency of multiple storms increase. Mean errors subsequently increase by 20 to 30% in July and August, and then taper off through the rest of the season, as the workload stabilizes and as the newcomers gain experience. This trend is less pronounced in the 48- and 72-hour data, but in light of the case distribution, the argument is not negated. Support for this argument is shown by mean errors increasing with the number of storms occurring simultaneously (Figure 5a). This could be indicative of the aforementioned added workload on the forecasters, or perhaps due to complicated multistorm interaction not fully understood. With the progressively fewer number of cases considered for an increasing number of storms, the trend is not strongly supported, however. It is noted that four storm cases occurred in 1972 only and may reflect the year rather than the occurrence of four storms. The relationship is reinforced, however, in light of the larger errors occurring when less than 600 nautical miles separates two storms. This parameter, the distance to the nearest storm, is depicted in Figure 5b. It is in agreement with findings by Brand

(1968) that the Fujiwhara effect is not felt beyond 750 nautical miles. Beyond that distance, mean errors decreased and stabilized. The parameter of six-hourly point along the track (Figure 6a), a measure of the length of storm history, showed a trend congruous with that of maximum wind: as the storm's history and development increased, the forecast errors decreased. In this case, a minimum mean error occurs late in the third day of a storm. This stabilizes in the 24-hour situation, but decays for 48- and 72-hour forecasts, as might be expected with storm recurvature occurring late in the storm history.

For the last variable, initial position error (Figure 6b), a consistent and prominent trend shows the mean forecast errors increasing as the initial position errors increase. This supports the basic forecasting premise that accurate observations are necessary for accurate forecasts. This finding is in general agreement with that of Neumann (1975), who found that for Atlantic hurricanes, the initial position error was important in objective forecasts with its relative importance decreasing in longer-range forecasts.

C. ANNUAL VARIATION OF ERRORS

Figure 7 shows the mean forecast errors for each year of the 10 year sample, with the least squares linear trend lines. In each case, the trend line is too shallow to indicate conclusively an improvement during the 10 years. Correlations are negative but less in magnitude than 0.3 (about one standard deviation from 0.0). While the hypothesis that there has been no improvement over the 10 years is suspected, it cannot be rejected. Because these main error values were computed including all tropical cyclones, they differ from those published in the annual typhoon reports, which were based only on storms of typhoon intensity.

D. AUTOCORRELATION OF SUCCESSIVE FORECASTS

To this point, the forecasts (and hence forecast errors) have been tacitly assumed to be independent of each other. In reality, successive six-hourly forecasts for a particular storm are strongly correlated. Table III gives the estimated autocorrelation coefficients between errors from successive forecasts with lag times out to 36 hours. It is possible to adjust the number of related cases in a particular storm downward to an effective number of independent cases by a complicated relationship given by van der Bijl (1951). The ratio of these two values (the effective number of independent cases divided by the total cases) decreases with an increase in the autocorrelation coefficient as well as with increasing numbers of forecasts per storm, and increases with lag time between forecasts. For 24-hour forecasts, where typically 15 to 20 forecasts are made at six-hourly intervals, this ratio is approximately 1/3. In the 48-hour case where the autocorrelation is higher, but the typical number of forecasts per storm is lower (10-15), this ratio is 1/4 to 1/3. At 72 hours, the six-hourly autocorrelation coefficient is higher; however, the forecasts were usually issued at 12-hourly intervals and the typical number of forecasts was about five per storm, thus increasing the ratio to 1/3 to 1/2.

This ratio is important in significance testing where the square root of the number of cases (to be replaced by the effective number of independent cases) is found in the denominator of the test statistic. Whether 1/4, 1/3, or 1/2 of the number of cases is used as the effective number makes little difference in the test statistic, so 1/3 times the number of cases will be used as an arbitrary compromise estimate of the effective number of independent cases throughout for the purpose of significance testing.

E. FREQUENCY DISTRIBUTION OF FORECAST ERROR COMPONENTS

For the purpose of constructing probability ellipses, error components have been assumed to be distributed according to a Guassian, or normal, frequency distribution. Figures 8a and 8b show the observed cumulative frequency distributions of the West-East (U) and South-North (V) components of the errors plotted on probability scaled paper, where a normal distribution would be represented by a straight line. This presentation shows generally good agreement between the plotted observed and normal curves (straight lines) computed from estimates of the means and standard deviations. The maximum differences between the theoretical and observed cumulative frequencies are:

	U COMPONENT	V COMPONENT
24HR	4.5% errors < - 70 NMI	3.9% errors < - 70 NMI
48HR	2.3% errors < - 70 NMI	3.9% errors < - 30 NMI
72HR	2.2% errors < -110 NMI	5.1% errors < - 30 NMI

The Kolmogorov-Smirnov goodness of fit test (Massey, 1951) regards as significant at the 5% level, differences in observed and theoretical cumulative frequencies greater than $1.36/\sqrt{N}$ in absolute value. Using an effective number of cases of 1/3 (4809) = 1603 at 24 hours, 1/3 (3038) = 1013 at 48 hours, and 1/3 (1372) = 457 at 72 hours; the cutoff points for significance would be 3.4% at 24 hours, 4.3% at 48 hours, and 6.4% at 72 hours. Only the differences between the 24-hour theoretical curves and the observed plotted values are significant at the 5% level. Estimates of the third and fourth moments about the mean (Table IV) reveal that the 24-hour forecast errors are skewed west (forecasts are too far east) while all other skewness coefficients appear normal. Both components, however, appear to be leptokurtic. This is evident in Figure 8, where extreme occurrences fall counterclockwise with respect to the theoretical lines. This suggests that probability ellipses, based on the assumption of Gaussian distributions, may be slightly biased in that 24-hour verifying positions are more likely to fall out of an ellipse on the west side that the east side, and that inner and outer ellipses may not contain the proper proportion of the verifying positions. In general the 10% ellipses would be expected to contain more than 10% of the verifying positions and the area beyond the 95% ellipse to contain more than 5% of the verifying positions. It is not possible to make a statement about the intermediate ellipses between 10 and 95%, but Figure 8 suggests good agreement between the theoretical and observed cumulative frequencies there.

If probability ellipses, or other estimates of future probable error, integrated over an area are desired, the observed cumulative distribution could be used in place of the Gaussian cumulative distribution. The degree of complexity added by such a step, as well as the uncertainty in the representativeness of this particular 10 year sample (to the future) suggest such a step is not warranted.

IV. DISCRIMINANT ANALYSIS

A. CALCULATION OF DISCRIMINANT FUNCTIONS AND GROUP MEANS

Seeking to identify forecasts as either good or bad, discriminant analysis was used, namely, the UCLA Biomedical computer program BMDP7M (Dixon, 1974). With this approach, the cases were divided into groups and classification functions were found that best delineate the groups. These functions, linear combinations of the variables, would then be used to predict the classification of new cases. BMDP7M is the stepwise discriminant analysis which identifies the subset of variables that maximizes the difference between groups. Variables are entered into the classification function one at a time until there is no appreciable improvement in group separation.

For this study, U and V error components were either good, with the absolute values of the error less than or equal to the median; or bad, with the absolute value of the error greater than the median. The four possible combinations were resolved into three classifications:

GROUP 1: both U and V components good

GROUP 2: either U or V good

GROUP 3: both U and V components bad

The six variables contributing to the separation of groups included all those selected by the linear regression:

1.	Latitude	LAT
2.	Longitude	LONG
3.	Maximum Wind	MAX WND
4.	West-East Component of Movement	U MOVT

5. South-North Component of Movement V MOVT

6. Number of Storms in Progress at Forecast Time NO STM

Their means and standard deviations are given in Table V.

Previously established trends are consistently apparent in the data. Group 1 forecasts are associated with lower latitudes, more westerly longitudes, faster westerly movement, minimal N-S movement, the more intense storms (typhoons), and a fewer number of concurrent storms.

B. TESTING OF THE DISCRIMINANT FUNCTIONS AND GROUP MEANS

Each set of forecast cases was then tested by applying the group means and classification functions calculated from the BMDP7M program, using the 24 hour data. The classification functions derived from the 24-hour forecasts were also applied to 48- and 72-hour forecasts. This resulted in slight loss of discrimination at 48 hours, but has the advantage of classifying a forecast into the same category for forecasts at all time intervals.

Using the six resulting coefficients ($c_1, \ldots\ldots, c_6$) and a constant (c_7) for each group, operating on the six selected variables ($x_1, \ldots\ldots, x_6$), three functions (f) were evaluated for each group thusly:

$$f_i = c_{1i}x_1 + c_{2i}x_2 + \ldots\ldots + c_{6i}x_6 + c_{7i}\ , \quad i = 1,2,3$$

Each function value was subtracted from its corresponding component of the group means, and the differences were squared and summed to represent a vector distance from each group mean. Each case was then assigned to one of three groups according to which vector distance was minimal. The cases so sorted were counted, and the means and standard deviations of the error components and error magnitudes were calculated. These means and standard deviations are given in Table IV. The standard deviations of the components and the mean vector errors reflect the differences in the groups, with the component means mostly near zero. A Student's t test was applied to determine if the mean vector errors of each group were significantly different from those of the other groups. At the 5% level of significance for a one tail test, $t = 1.645$. Therefore, if the value of t between two groups exceeds that figure, the groups are deemed to be significantly different. Values of t computed on the group means are given in Table VII. It should be noted that the number of cases was reduced by a factor of 3 to account for autocorrelation as previously discussed.

The mean errors of Group 1 were found to be significantly less than those of Group 2, and those of Group 2 were significantly less than those of Group 3, except for the 72-hour data. There the mean V components of the errors were significantly non-zero and of different signs, giving unique spatial error distributions for the two groups with only slightly different mean absolute errors.

C. GROUP ANNUAL VARIATIONS

Mean forecasts errors per year per group, and all groups combined, are shown in Figure 9 with trend lines. Again the difference between Group 1 and 3 is substantial. The Group 2 average errors most closely approximate the pattern of all three groups combined. Year to year fluctuations are extremely large for Group 2, and somewhat less for Groups 1 and 3. The larger fluctuations in the Group 2 means and in the mean errors of longer-range forecasts in all groups is mostly attributable to relatively smaller sample sizes.

D. PROBABILITY ELLIPSE COMPUTATIONS

Having thus far separated forecasts into one of three groups, it was now desirable to present this information in a more graphical way; the goal being a useful operational application. Following previously established methods (Stevens and Palmer, 1963), the probability ellipse is such an application. Assume that errors in forecast position approximate a bivariate normal distribution (Section III.E). The expression for an ellipse is given by:

$$x^2 - 2rxy + y^2 = (1 - r^2)\, c^2 \ ,$$

with the normalized error components $x = (U-\bar{U})/s_u$, and $y = (V-\bar{V})/s_v$. U is the E-W error component; V is the N-S error component. $\bar{U}$, $\bar{V}$, s_u, and s_v are the estimates of the respective means and standard deviations; and r is the estimated correaltion coefficient between U and V. Probability = $1 - e^{-c^2/2}$, c = 1 approximates a 40% ellipse. Figure 10 shows the 40% probability ellipses for each group at each forecast interval. Distance dimensions are nautical miles, areas are in thousands of square nautical miles, and directions are in degrees north of east. A 40% probability ellipse means that a forecast position has a 40% probability of falling within its corresponding ellipse. Only the difference in size between Groups 1 and 3 is immediately obvious, but upon comparison of the areas, the distinctions are more pronounced. The 24-hour area of Group 3 is more than double that of Group 1, with the 48- and 72-hour areas being 97 and 57% larger, respectively.

The general NW orientation of the major axis of the ellipse indicates that for low latitude storms on a normal WNW track, errors are comprised of nearly equal components along the track (speed error), and across the track (track error). These storms are usually associated with Groups 1 and 2. Recurving and post-recurvature storms are almost always in Groups 2 and 3, with Group 3 predominating as storms are entrapped by the westerlies. During recurvature, when the track is nearly north, track error is dominant, whereas after recurvature, speed error dominates.

Using too large an ellipse (such as Group 2 for a Group 1 case) tends to dilute and spread the estimated probability density. This has the effect of overwarning those customers far removed from the forecast track, and underwarning those along the track. Conversely, using too small an ellipse (such as Group 2 for a Group 3 case) has the effect of overwarning those along the forecast track and underwarning those in the periphery. This case is the meteorologist's familiar dilemma when forecasts are taken too literally without adequate allowance for errors. Tailoring the ellipses to the expected forecasting difficulty has the effect of reducing both overwarning and underwarning.

V. TESTING OF INDEPENDENT DATA (1976)

The final step in this research was to apply the same procedures to an independent data set and compare the results.

A. DISTRIBUTION OF ERRORS

The 1976 forecasts and best tracks, having been processed like the 10 year sample, were analyzed using the same discriminant functions and group means to similarly arrive at three separated groups. the 1976 group statistics are listed in Table VIII. The contrast between group statistics is not as sharp as in the 10 year independent data sample (Table VI). The means vary more widely as compared to the dependent data, while the smaller standard deviations of Group 2 show the 1976 forecasts to appear significantly better for those cases.

So the question arises as to whether errors of 1976 are representative of the 10 year data sample. Statistically each group vector mean of the 1976 data was compared with its counterpart in the dependent data sample. Table IX lists the values of the normal test statistic, Z, found for each group to be compared at the 5% level of significance value of Z: 1.96. From this, Groups 1 and 3 of 1976 cannot be rejected as having come from the sample of the previous 10 years, but Group 2 forecasts appear to be significantly better than the preceding 10 year average.

On the other hand, Figure 9 shows Group 2 to have a wide annual variation, partially because of smaller frequency of occurrence. The relatively high number of Group 2 cases in 1976 may include only a few storms which can negate any significance in the above differences.

B. ELLIPSE TESTING

Forecast errors of 1976 were tested to determine the percentages of verifying positions that would fall within ellipses with probabilities specified at 25, 50, 75, 90, and 95%. These results appear in Figure 11. For Group 1, the observed closely follows the 45° expected line with the maximum deviation of observed from expected being 10% at 72 hours. Group 2 deviations were consistently conservative (above the 45° line) with deviations up to 18%. For Group 3 the 72-hour deviation was the greatest at 20%, also conservative. None of these differences are significant at the 5% level by the Kolmogorov-Smirnov goodness of fit test. Generally, the ellipses were too conservative. This is to be expected since a trend of decreasing errors was evident in the 10 year data.

For dramatic comparison, the same ellipse testing was repeated, but counting the number of Group 3 cases that fell into the smaller Group 1 ellipses and the number of Group 1 cases that fell into the larger Group 3 ellipses (Figure 12). The comparison shows the Group 3 ellipses to be ultra-conservative for Group 1 cases. Conversely, fewer Group 3 observed cases fell into Group 1 ellipses, by roughly the same percentages below the expected as the other situation was above. This contrast shows significant differences exist in the distribution of errors from forecasts specified in advance to be either Group 1 or Group 3.

VI. CONCLUSIONS

In light of the objectives outlined for this study, to some extent the long-range goals have been attained. It has been demonstrated that a small number of readily available parameters can, with reasonable effectiveness, classify a tropical cyclone forecast as likely resulting in either markedly above or below average errors. Group 1 forecasts have a high probability of below average errors with a low probability of above average errors. Group 2 forecasts have approximately equal probabilities of being above or below average. Group 3 forecasts have a low probability of below average errors with a high probability of above average errors.

The concept of using least squares regression to predict in advance the actual error (as opposed to a class of errors) appears to offer little chance of meaningful success. It is apparent that it is possible to isolate conditions contributing to forecast errors in the mean, but one must bear in mind that excellent forecasts are occasionally made under the worst conditions, and conversely, terrible forecasts can be made under the best of conditions.

VII. RECOMMENDATIONS

The examination of two additional parameters could improve the delineation between classifications. First, initial position error, shown to be directly related to the forecast errors, was not introduced as a discriminator because it is not generally known to the forecaster at the time of the forecast.

Second, the synoptic pattern associated with each storm has not been considered. Some parameter which accounts for the relative locations of semi-permanent features; such as the TUTT, subtropical ridges, and perhaps transient troughs in the westerlies; might prove to be a most important discriminator, especially as it relates to the track forecast errors and the basic problem of forecasting the recurvature of tropical cyclones.

While these results are to be considered preliminary, pending improvements and refinements, dissemination to typhoon forecast subscribers is recommended.

TABLE I. BASIC STATISTICS - RESULTS OF STEPWISE MULTIPLE LINEAR REGRESSIONS

VARIABLE	MEANS	STANDARD DEVIATION	VARIANCE EXPLAINED (%)		
			24-HR	48-HR	72-HR
MAX WND (kts)	65.9	28.8	3.4	1.4	0.4
LAT (°N)	18.7	6.3	4.2	1.1	0.6
LONG (°E)	135.0	14.6	1.8	2.3	2.3
U MOVT (kts)	-5.5	6.4	0.3	3.3	2.9
V MOVT (kts)	4.0	4.1	0.5	0.5	0.3
POS. NO.	13.5	10.6	NE*	NE	NE
NO. STM.	1.5	0.7	0.3	0.2	NE
MONTH	8.7	2.3	NE	NE	0.1
TIME GMT (hrs)	10.3	7.0	NE	NE	NE
ERR DIS: 24-HR (NMI)	125.7	80.8			
ERR DIS: 48-HR (NMI)	247.0	153.8			
ERR DIS: 72-HR (NMI)	369.4	226.0			
TOTAL VARIANCE EXPLAINED (%)			10.5	8.8	6.6

*NE: Was not entered in linear regression

TABLE II. CORRELATION MATRIX

	MAX WND	LAT	LONG	U MOVT	V MOVT	POS NO.	NO. STM.	MONTH	TIME GMT	24-HR ERR DIS	48-HR ERR DIS	72-HR ERR DIS
MAX WND	1.000	.177	.024	-.062	.153	.397	.022	.095	.036	-.185	-.102	-.052
LAT		1.000	.116	.354	.284	.367	.130	.069	-.006	.169	.173	.131
LONG			1.000	.119	.084	-.130	.061	-.001	.001	.152	.164	.167
U MOVT				1.000	.117	.161	.063	-.077	-.003	.150	.182	.171
V MOVT					1.000	.017	-.043	-.066	-.005	.101	.094	.061
POS. NO.						1.000	.072	.018	.031	-.053	-.003	.014
NO. STM.							1.000	.049	-.014	.073	.075	.045
MONTH								1.000	.028	-.019	.007	-.034
TIME GMT									1.000	-.008	-.021	.001

TABLE III. AUTOCORRELATIONS BETWEEN FORECAST ERRORS OF SUCCESSIVE FORECASTS

TIME LAG (HOURS)	24-HR	48-HR	72-HR
0	1.000	1.000	1.000
6	.665	.790	.838
12	.432	.587	.675
18	.291	.432	.476
24	.213	.305	.371
30	.173	.212	.127
36	.181	.171	.177

TABLE IV. SKEWNESS AND KURTOSIS COEFFICIENTS OF ERROR COMPONENTS

U	NORMAL	24-HR	48-HR	72-HR
3rd Moment Skewness Coef.	0.0	-1.58*	-.013	.046
4th Moment Kurtosis	3.0	4.434	3.753*	3.522*
V				
3rd Moment	0.0	0.18	-.061	.067
4th Moment	3.0	4.544*	3.881*	3.185

*Significantly different from the Gaussian values.

TABLE V. DISCRIMINANT ANALYSIS GROUP STATISTICS

	MAX WND(kts)	LAT(°N)	LONG(°E)	U MOVT(kts)	V MOVT(kts)	NR STMS	CASE COUNT
MEANS							
GROUP 1	78.1	17.7	132.4	-6.6	3.6	1.5	1423
GROUP 2	64.2	18.6	135.1	-5.2	3.8	1.5	2219
GROUP 3	62.2	20.0	138.0	-4.7	4.7	1.6	1167
TOTAL	65.9	18.7	135.0	-5.5	4.0	1.5	4809
STANDARD DEVIATIONS							
GROUP 1	29.8	5.8	14.4	5.6	3.7	0.7	
GROUP 2	28.5	6.4	14.7	6.5	3.9	0.7	
GROUP 3	26.9	6.6	14.1	6.7	4.8	0.8	
TOTAL	28.5	6.3	14.5	6.3	4.1	0.7	

TABLE VI. TEST GROUP STATISTICS

FORECAST 24-HR GROUP	CASES	MEANS U	MEANS V	STD. DEV. U	STD. DEV. V	MAGNITUDE OF VECTOR ERROR MEAN	MAGNITUDE OF VECTOR ERROR STD. DEV.
1	1834	-6.0	3.1	87.9	75.3	98.8	57.9
2	1310	-16.0	-5.1	119.4	95.9	130.4	84.0
3	1665	8.5	-3.1	129.6	118.4	151.1	89.8
48-HR							
1	1432	-15.0	4.1	195.9	151.4	211.1	130.3
2	733	-19.4	-31.4	237.0	179.8	242.2	181.3
3	873	19.2	15.6	258.7	223.0	295.7	172.7
72-HR							
1	623	-14.4	2.2	309.5	235.4	326.0	212.5
2	367	-8.1	-70.5	370.2	249.3	391.2	238.0
3	382	17.0	27.5	368.4	302.5	418.5	230.6

TABLE VII. STUDENT'S t VALUES COMPUTED ON GROUP MEANS (10 YEAR DATA)

	24-HR	48-HR	72-HR
Group 1 vs. 2	7.2	2.65	2.56
Group 2 vs. 3	3.7	3.48	0.92

TABLE VIII. 1976 TEST GROUP STATISTICS

FORECAST		MEANS		STD. DEV.		MAGNITUDE OF VECTOR ERROR	
24-HR GROUP	CASES	U	V	U	V	MEAN	STD. DEV.
1	144	-30.6	-15.9	84.8	80.6	104.6	62.6
2	240	-25.0	3.8	96.1	75.4	110.4	58.2
3	140	5.5	9.3	111.9	122.4	143.3	84.2
48-HR							
1	133	-75.1	-24.0	196.5	160.1	226.9	137.8
2	202	-57.6	13.2	192.3	145.9	216.9	121.4
3	89	42.2	33.4	214.7	195.6	265.2	130.0
72-HR							
1	109	-110.8	-35.2	328.0	240.1	357.3	225.9
2	164	-72.1	11.2	289.5	206.3	312.4	184.7
3	59	69.2	65.0	337.1	204.5	376.0	151.9

TABLE IX. NORMALIZED DEVIATIONS OF 1976 GROUP MEANS FROM 10 YEAR GROUP MEANS

GROUP	24-HR	48-HR	72-HR
1	0.60	0.80	0.88
2	-2.12	-1.15	-2.45
3	-0.51	-0.96	-0.82

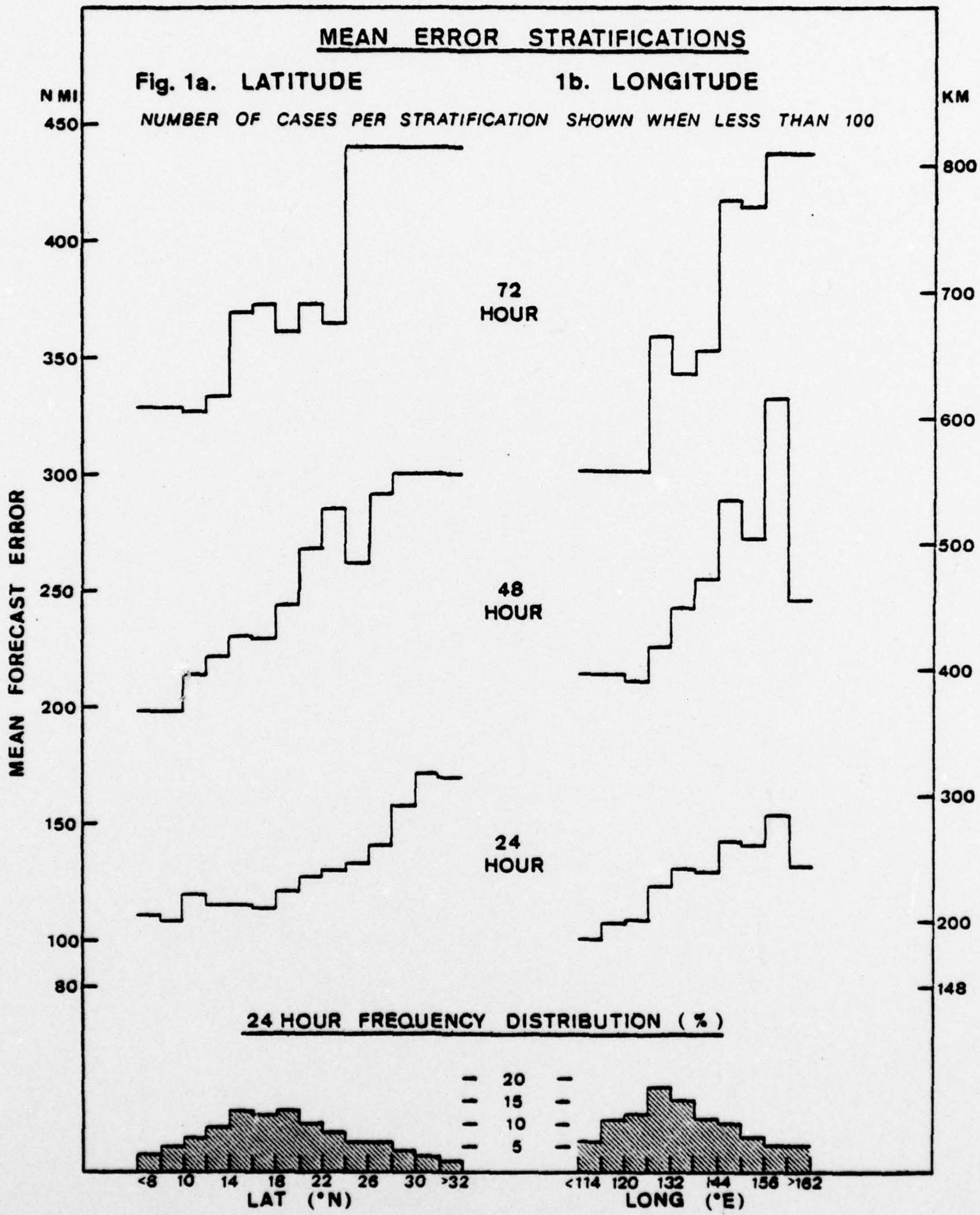

Figure 1. Mean Error Stratifications by Latitude, Longitude.

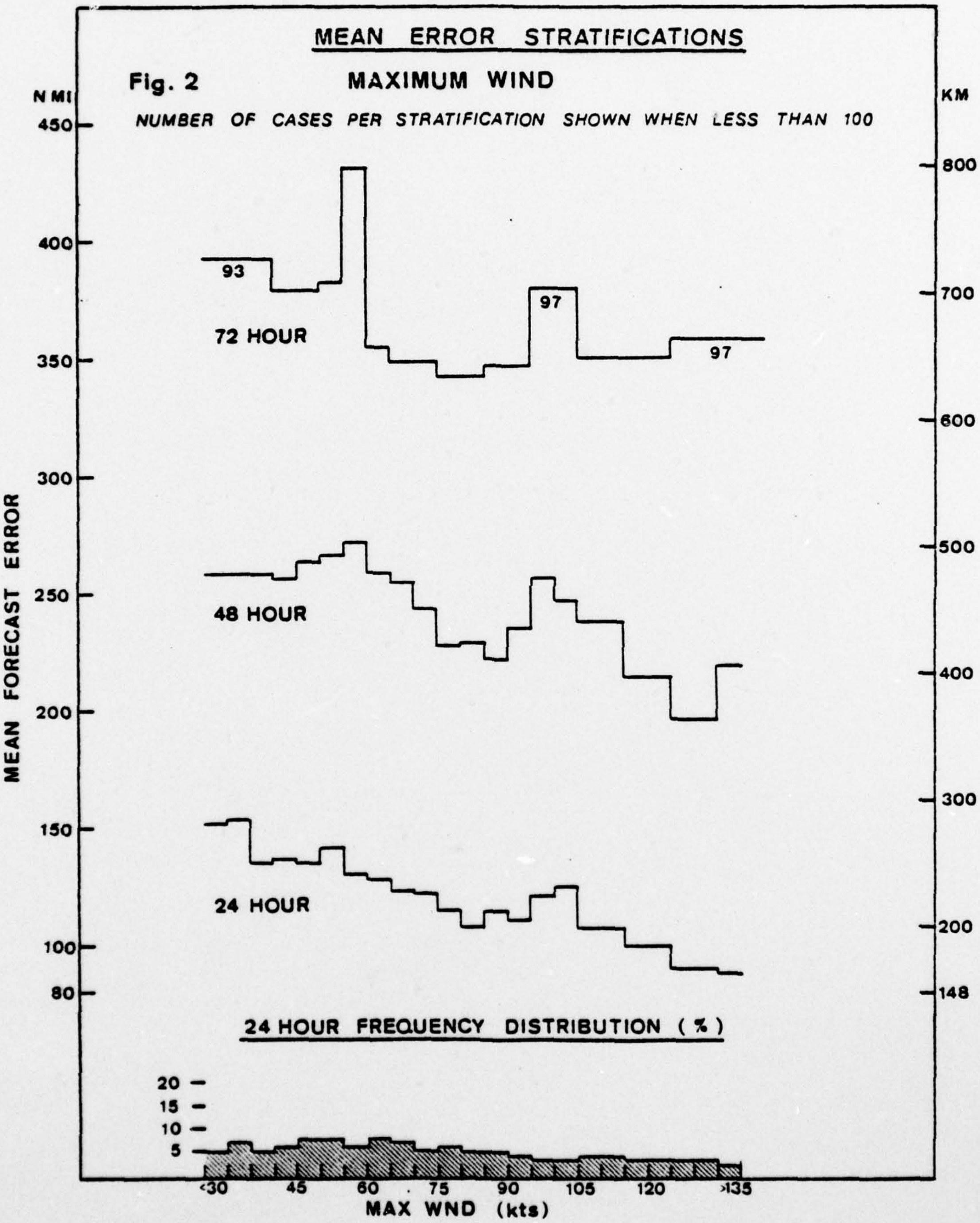

Figure 2. Mean Error Stratifications by Maximum Wind.

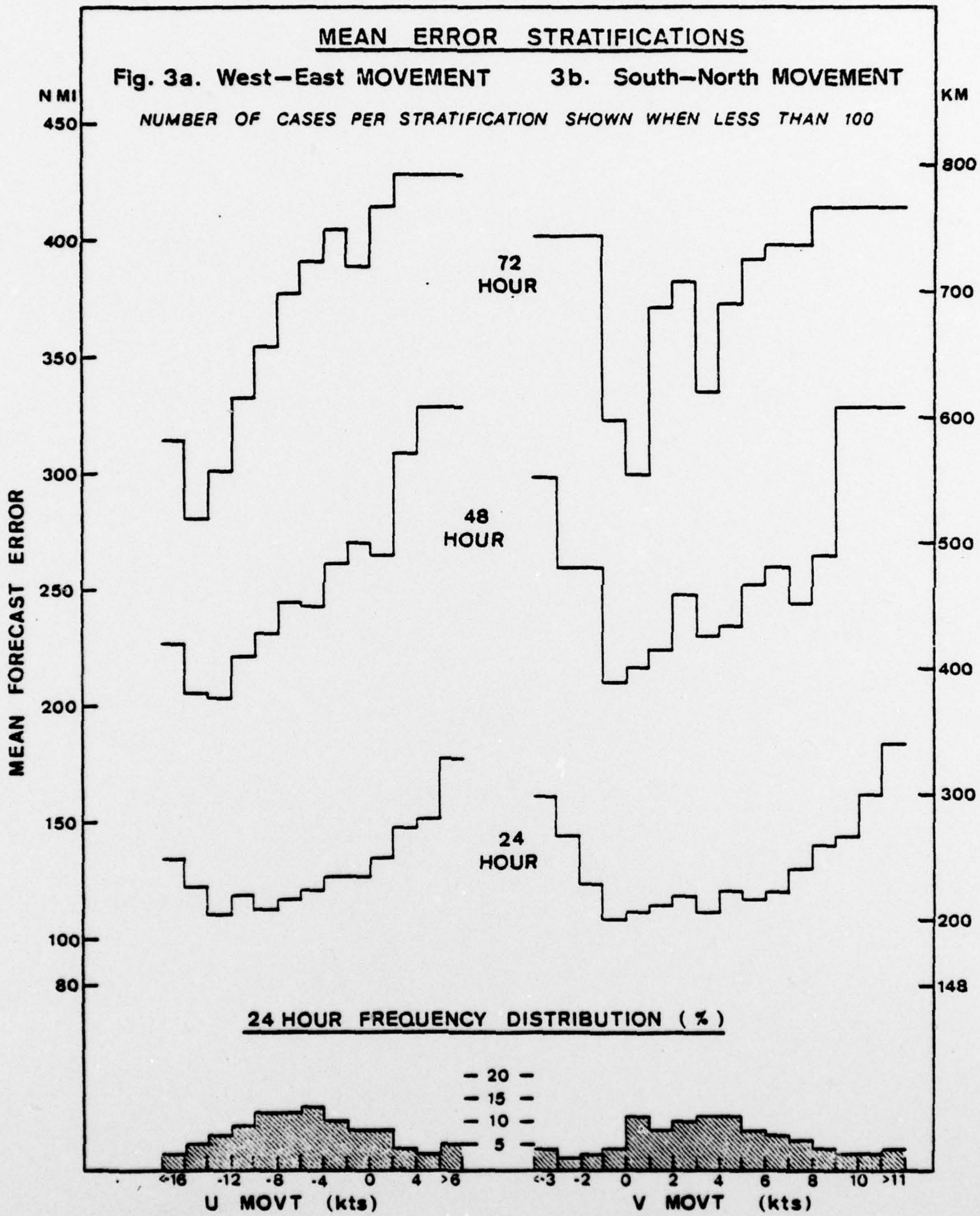

Figure 3. Mean Error Stratifications by West-East, South-North Components of Movement.

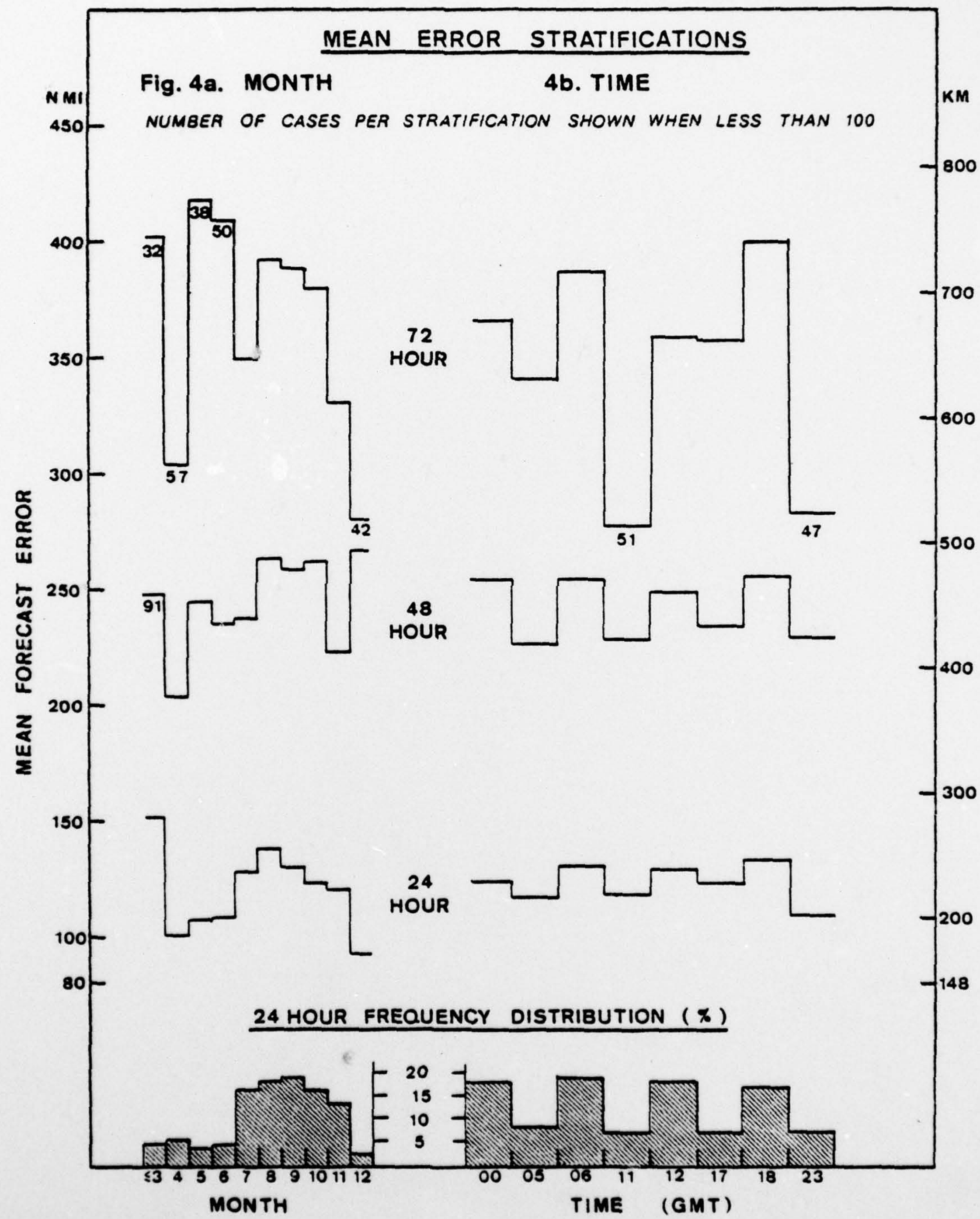

Figure 4. Mean Error Stratifications by Month, Time (GMT).

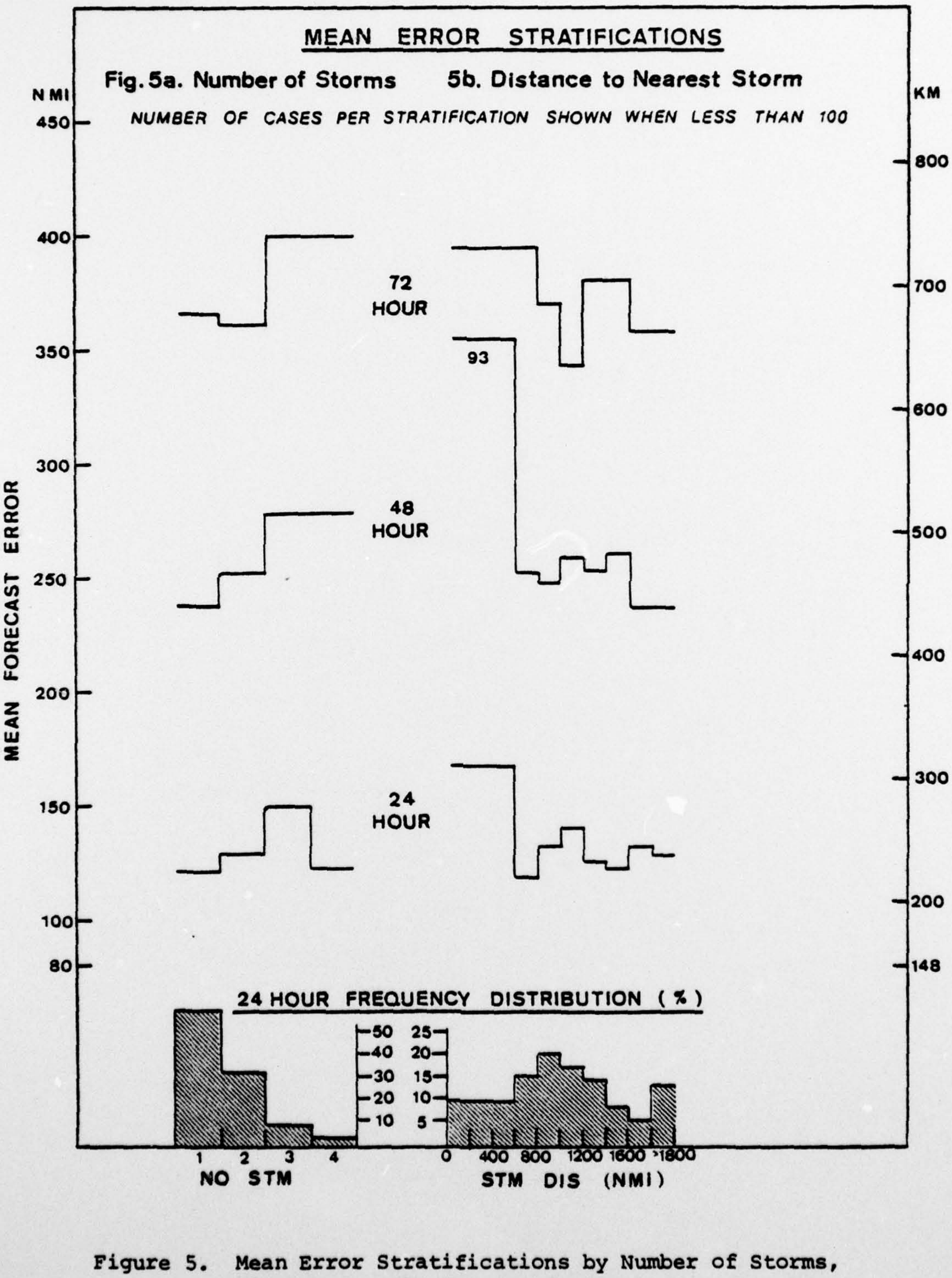

Figure 5. Mean Error Stratifications by Number of Storms, Distance to the Nearest Storm.

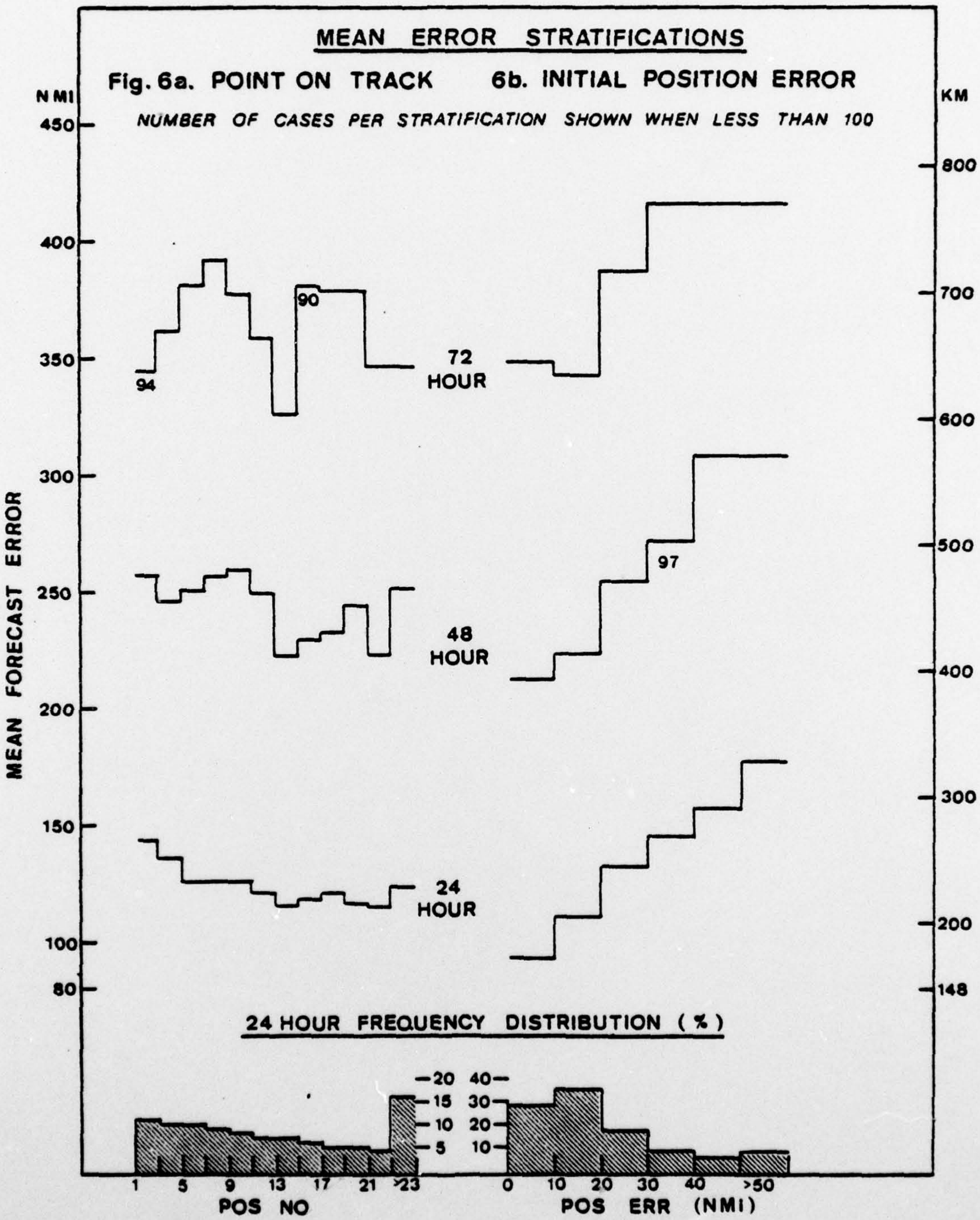

Figure 6. Mean Error Stratifications by Point on Track, Initial Position Error.

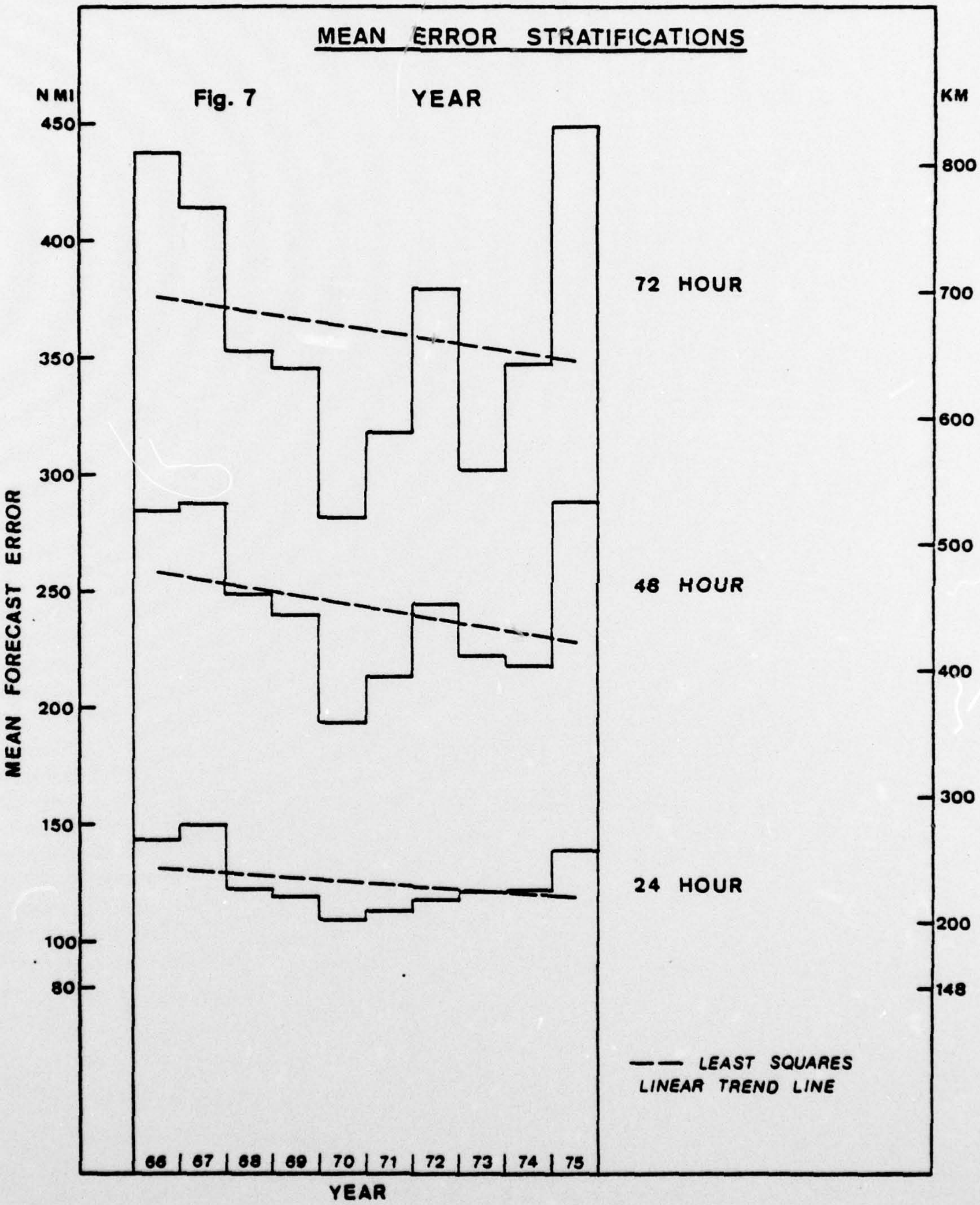

Figure 7. Mean Error Stratifications by Year.

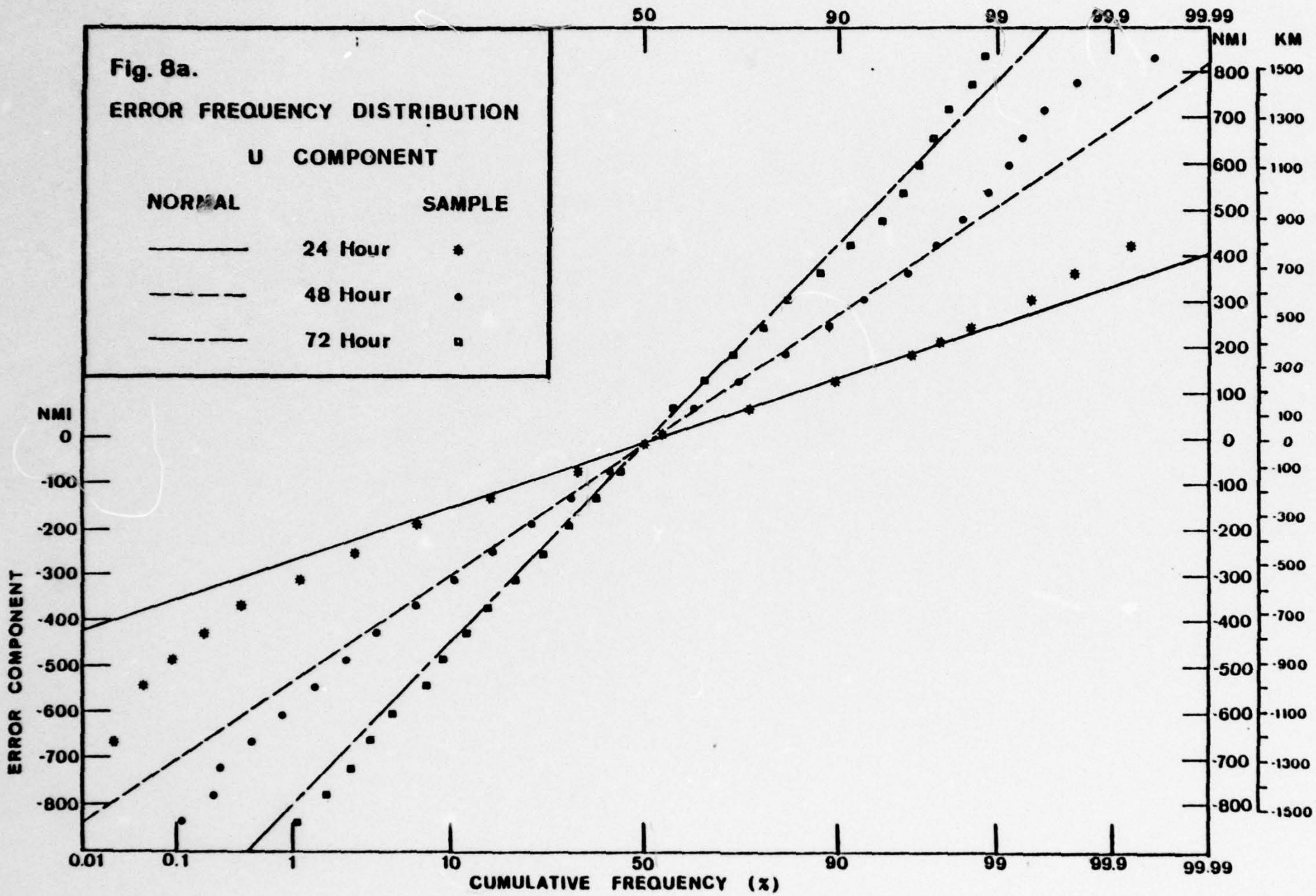

Figure 8a. Error Frequency Distribution - U Component.

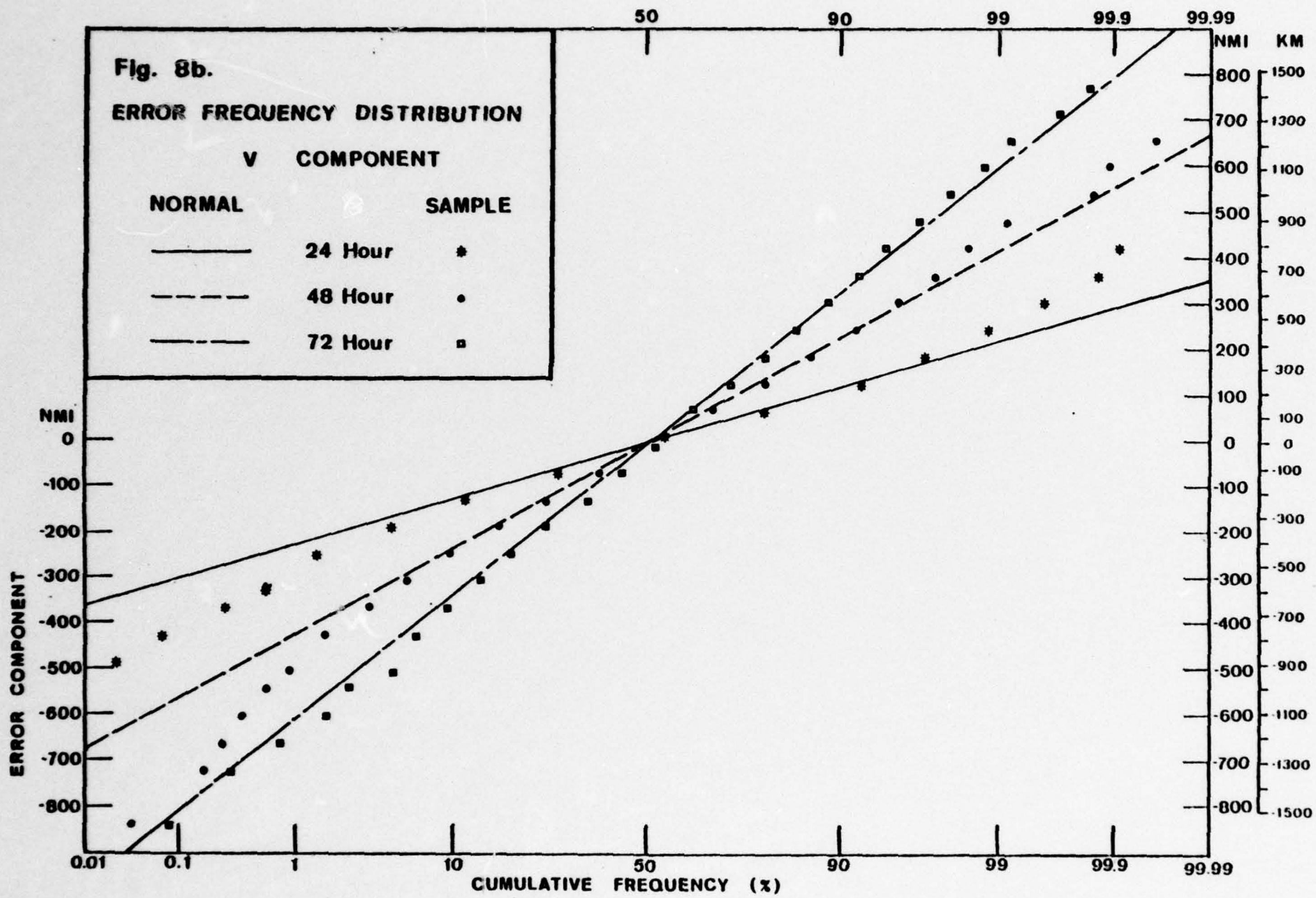

Figure 8b. Error Frequency Distribution - V Component.

Fig. 9

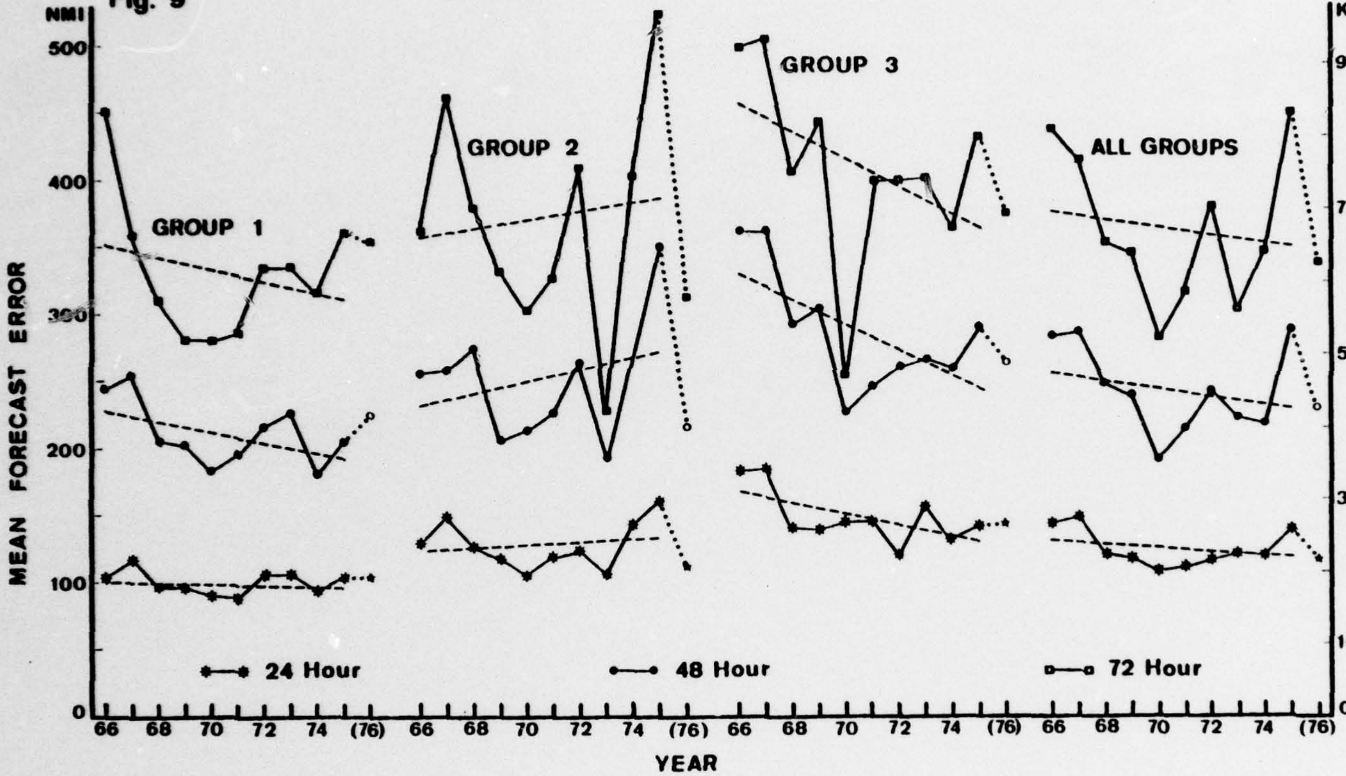

Figure 9. Mean Annual Forecast Error Stratifications Per Group [with Least Squares Linear Trend Lines (---)].

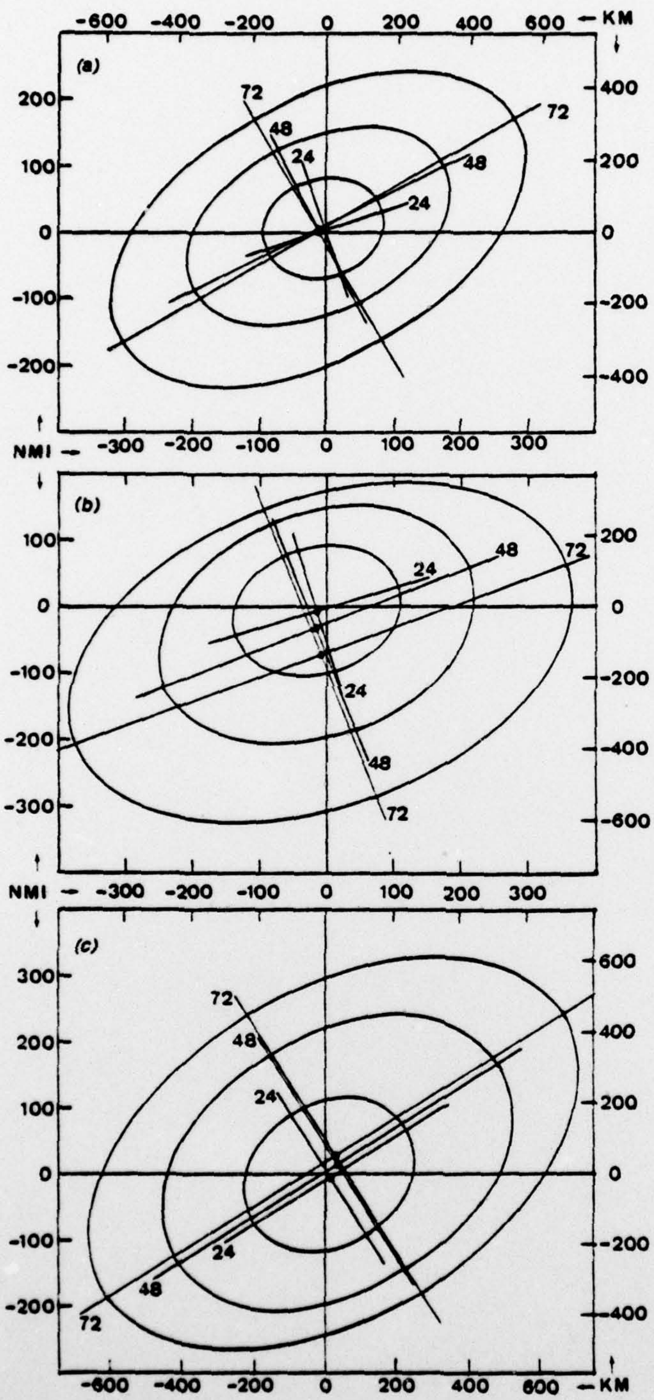

Fig. 10

GROUP 1

HR	24	48	72
$\bar{U}$	-6	-15	-15
$\bar{V}$	3	4	2
S_u	90	196	310
S_v	74	151	235
r	.12	.35	.46
DIR	19°	27°	30°
SEMI AXES	90	209	339
	74	133	191
AREA	20.6	87.3	203.4

GROUP 2

HR	24	48	72
$\bar{U}$	-16	-19	-8
$\bar{V}$	-5	-31	-70
S_u	119	237	370
S_v	97	182	252
r	.14	.26	.34
DIR	17°	22°	20°
SEMI AXES	121	247	386
	94	170	228
AREA	35.9	131.9	275.9

GROUP 3

HR	24	48	72
$\bar{U}$	8	19	17
$\bar{V}$	-3	16	28
S_u	130	259	368
S_v	118	223	302
r	.20	.31	.41
DIR	32°	32°	32°
SEMI AXES	137	280	405
	110	196	252
AREA	47.3	172.1	319.9

ALL DISTANCES: NMI
AREAS: THOUSANDS OF SQ. NMI
DIR: LOCATES MAJOR AXIS(°N of E)
$\bar{U}$, $\bar{V}$: LOCATE ELLIPSE CENTER

Figure 10. 40% Probability Ellipses Per Group Per Forecast.

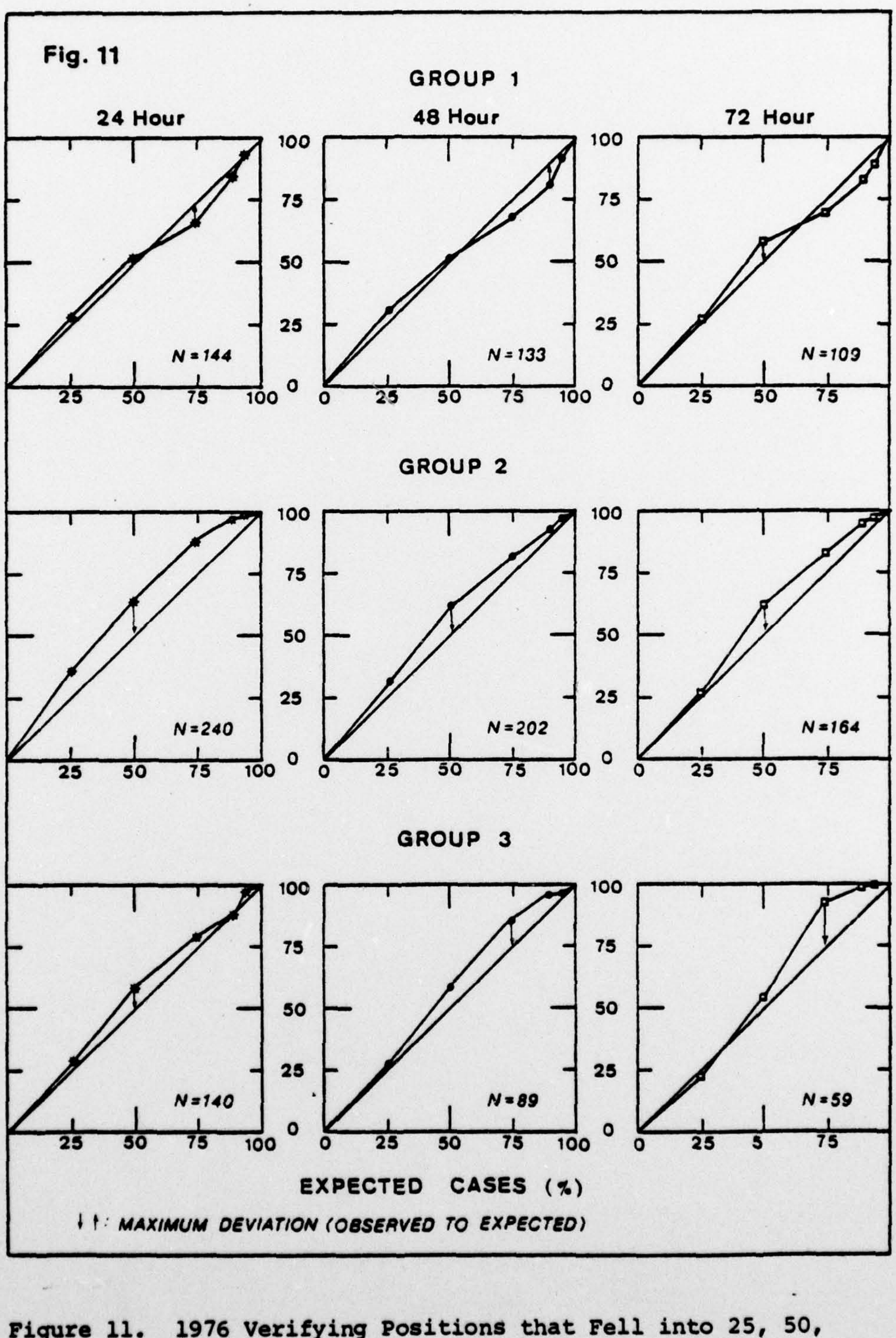

Figure 11. 1976 Verifying Positions that Fell into 25, 50, 75,90, and 95% Probability Ellipses by Group and Forecast (24, 48, and 72 hours).

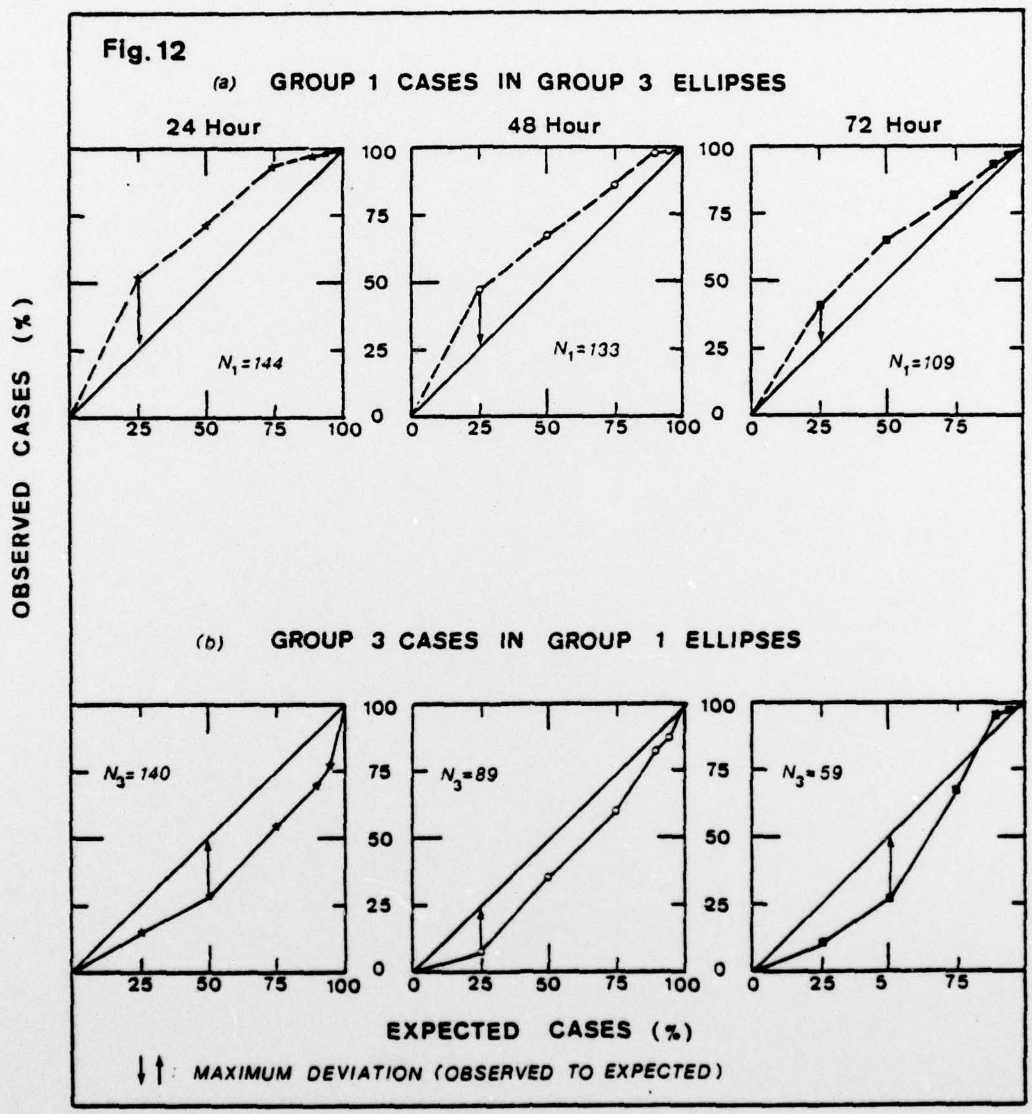

Figure 12. Contrast of Group 1 and Group 3 1976 Verifying Positions that Fell into 25,50,75,90, and 95% Probability Ellipses for Group 1 and Group 3 by Forecast (24, 48, and 72 hours).

BIBLIOGRAPHY

Annual Typhoon Reports, 1966-1976: Fleet Weather Central/Joint Typhoon Joint Typhoon Warning Center, Guam.

Atkinson, G. D., Forecasters Guide to Tropical Meteorology, Technical Report 240, Air Weather Service (MAC) USAF, April 1971.

Brand, S., Interaction of Binary Tropical Cyclones of the Western North Pacific Ocean, NAVWEARSCHFAC Technical Paper No. 26-68, September 1968.

Brooks, C. E. P., and Carruthers, N., Handbook of Statistical Methods in Meteorology, Her Majesty's Stationery Office, London, 1953.

Burington, R. S., and May, D. C., Handbook of Probability and Statistics with Tables, Handbook Publishers, Inc., Sandusky, Ohio, 1953.

Dixon, W. J., BMD Biomedical Computer Programs, University of California Press, Berkeley, 1970.

_____, BMDP Biomedical Computer Programs, University of California Press, Berkeley, 1975.

Jarrell, J. D., Personnel Communication, 1977.

Massey, F. J., "The Kolmogorov-Smirnov Test for Goodness of Fit," Journal of the American Statistical Association, Vol. 46, p. 68-78, 1951.

Neumann, C. J., The Effect of Initial Data Uncertainties on the Performance of Statistical Tropical Cyclone Prediction Models, NOAA Technical Memorandum NWS SR-81, March 1975.

_____, A Statistical Study of Tropical Cyclone Positioning Errors with Economic Applications, NOAA Technical Memorandum NWS SR-82, March 1975.

Panofsky, H. A., and Brier G. W., Some Applications of Statistics to Meteorology, University Park, PA., 1968.

Sadler, J. C., The Tropical Upper Tropospheric Trough as a Secondary Source of Typhoons and a Primary Source of Tradewind Disturbances, Contract No. AF19(628)-3860, Final Report, Hawaii Institute of Geophysics, University of Hawaii, March 1967.

_____, A Role of the Tropical Upper Tropospheric Trough in Early Season Typhoon Development, NAVENVPREDRSCHFAC Technical Paper No. 2-74, June, 1974.

Sadler, J. C., Tropical Cyclone Initiation by the Tropical Upper Tropospheric Trough, NAVENVPREDRSCHFAC Technical Paper No. 2-76, February 1976.

Stevens, W., and Palmer, C. A., An Examination of the Distribution of Hurricane Forecast Errors Using Probability Ellipses, NWRF 12-1063-81, October 1963.

Tropical Cyclone Conference Proceedings Reports, 1974 and 1976, Environmental Group Pacific Command.

Van der Bijl, W., "Fünf Fehlerquellen in Wissenschaftlicher Statistischer Forschung," Annalen der Meteorologie, Hamburg, Vol 4, p. 183-212, 1951.

Yamane, T., Statistics, An Introductory Analysis, Harper and Row, New York, 1964.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. G. J. Haltiner, Code 63Ha Chairman, Department of Meteorology Naval Postgraduate School Monterey, California 93940	1
4.	CDR. J. D. Jarrell, Code 35 Naval Postgraduate School Monterey, California 93940	4
5.	Mr. Samson Brand Naval Environmental Prediction Research Facility Naval Postgraduate School Monterey, California 93940	1
6.	Air Weather Service AWVAS/TF Scott AFB, Illinois 62225	2
7.	Capt. Harry Hughes AFIT/CIPF Wright-Patterson AFB, Ohio 45433	1
8.	Director JTWC Box 17 FLEWEACEN, COMNAVMAR FPO San Francisco 96630	2
9.	Dr. W. van der Bijl, Code 63Vb Naval Postgraduate School Monterey, California 93940	4
10.	Capt. J. D. Shewchuk Det 1, 1WW COMNAVMAR Box 17 FPO San Francisco 96630	1
11.	1Lt. Donald S. Nicklin 3304 Coffey Avenue Omaha, Nebraska 68123	5