NOAA TM ERL WMPO-20

NOAA Technical Memorandum ERL WMPO-20

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Environmental Research Laboratories

An Evaluation of the Accuracy of Tropical Cyclone Intensities and Locations Determined From Satellite Pictures

ROBERT C. SHEETS
PAUL GRIEMAN

Weather
Modification
Program Office
BOULDER,
COLORADO
February 1975

NHEML-108

RONMENTAL RESEARCH LABORATORIES

ATHER MODIFICATION PROGRAM OFFICE

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

NOAA Technical Memorandum ERL WMPO-20

AN EVALUATION OF THE ACCURACY OF TROPICAL CYCLONE INTENSITIES AND LOCATIONS DETERMINED FROM SATELLITE PICTURES

Robert C. Sheets
Paul Grieman

National Hurricane and Experimental Meteorology Laboratory

Weather Modification Program Office
Boulder, Colorado
February 1975

UNITED STATES
DEPARTMENT OF COMMERCE
Frederick B. Dent, Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
Robert M. White, Administrator

Environmental Research
Laboratories
Wilmot N. Hess, Director

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (CONT.)

AN EVALUATION OF THE ACCURACY OF TROPICAL CYCLONE INTENSITIES AND LOCATIONS DETERMINED FROM SATELLITE PICTURES

Robert C. Sheets and Paul Grieman[1]

An attempt is made to evaluate the accuracy of tropical cyclone intensity and location estimates determined from satellite information only. Data from three satellites (ESSA, ATS, DMSP) were used for this study. Operationally qualified groups from the National Environmental Satellite Services in Washington and Miami and from the U.S. Air Force Air Weather Service did the storm intensity and location evaluations. Some 50 storms including hurricanes, tropical storms, and Western Pacific typhoons comprised the data sample. The results indicate that higher resolution satellite data result in more accurate positioning but the accuracy of intensity estimates is not necessarily a function of the resolution. Logically, strong systems are more accurately located than weak systems. The intensity estimates, however, are most accurate for weak and average strength storms than for rapidly developing storms. The errors were also largest for the very strong storms contained in the Western Pacific DMSP sample.

I. INTRODUCTION

Techniques have been developed during the past few years to determine the intensity and location of tropical cyclones from satellite pictures. The Dvorak technique (Dvorak, 1973), appears to show the most skill among these schemes, and has been used extensively on an operational basis for the past 2 or 3 years. The success of this scheme has resulted in the implementation of the "Selective Reconnaissance Program" (SRP) in the Western Pacific. This program calls for substituting satellite "fixes" for reconnaissance aircraft "fixes" under certain conditions. The potential for reducing operational costs by use of this program leads to the desire to expand the procedure in the Pacific and the Atlantic. However, those individuals directly responsible for forecasts and warnings, especially in the Atlantic, do not feel that adequate information is now available for making intelligent decisions about if, when, and where these substitutions should be made. The purpose of this study is to provide such information.

---

[1] Presently at the University of Washington, Department of Atmospheric Science.

The Dvorak technique has been evaluated, but for various reasons, the studies do not provide the specific intensity and position accuracy information desired. For instance, operational storm classifications from satellite pictures are evaluated by the Department of Defense (DOD) Joint Typhoon Warning Center (JTWC) and the Miami Satellite Field Service Station (SFSS) of the National Environmental Satellite Service (NESS) groups. However, in most cases, the analyst had information other than satellite data available at the time of classification (e.g., past and present intensities measured by aircraft, and forecasts). How much this additional information influenced the subsequent classification is unknown. Also, sometimes only satellite data were available, therefore the classifications were essentially verified against themselves.

Erickson (1972) completed a study (unpublished report) using ESSA satellite data where the analysts were requested to use only the satellite pictures. The analysts were a mixture of trained and marginally trained personnel. The results of this study indicated a reasonable degree of skill in classification of storm intensities. However, information on the location of the storm centers for the various satellite products was not included in his study. Also, the Dvorak classification scheme has been modified, products of other satellites are being used, and considerable operational experience has been gained since that testing. For all of the above reasons, further evaluation seems necessary.

The testing reported in this paper includes three different satellite products. These are pictures obtained from the ESSA, ATS III, and DMSP satellite systems. The ESSA sample is from 1967, 1968, and 1970 with four storms chosen from the Atlantic Ocean (including Gulf of Mexico and Caribbean Sea) and six from the Western Pacific. The ATS III sample contains 19 Atlantic storms that occurred in 1971 through 1973. The DMSP sample of 21 cases was chosen from the 1972 and 1973 Western Pacific storm seasons.

The three groups of participants in the testing were operationally qualified in the application of the Dvorak technique. These included a group from Washington, D.C., chosen by Dvorak from the NESS operational and applications branches, an operational group from the Miami, Florida, NESS/SFSS, and a group from the U.S. Air Force Air Weather Service (AWS). The Washington, Miami, and AWS groups were primarily experienced with use of the ESSA, ATS, and DMSP products, respectively. The AWS group had not worked with ATS and the Miami group had not worked with DMSP data before our tests began. Some of the Washington group, however, had worked with both Atlantic and Pacific cases.

Two major tests determine the usefulness of any objective or semi-objective scheme such as the Dvorak technique: (1) does the system accurately perform its designed function, and (2) can individuals with a reasonable amount of training, obtain repeatable results with a reasonable effort? The second question can be answered rather easily for the Dvorak scheme, but a conclusive answer to the first question is much more difficult to obtain. That is, a measure of the consistency of the analysis scheme is rather easily determined, but a measure of the absolute accuracy of the intensity and location of the storm is more difficult. This condition results from "ground truth," i.e., the actual location and intensity of the storm at the time of the satellite coverage, being nearly impossible to ascertain. Therefore, the major portion of this study is aimed at determining the internal consistency of the system. However, certain bounds can be determined and reasonable speculations can be made based upon the information available.

Dvorak has also developed a forecast scheme, an extension of his analysis technique, which makes use of satellite photographs to determine the cloud structure of the storm and its environment. The authors did not consider the evaluation of the forecast scheme as an integral part of the basic evaluation study. However, at the urging of Dvorak, each analyst was asked to complete the forecast block of their work sheets. These forecast data were evaluated in the same manner as the intensity estimations and the results are discussed in appendix A.

2. PROCEDURE

All participants were asked to position and classify the storm relative to the superimposed grid for the three sets of satellite pictures mentioned above. The photographs were given to the participant one at a time, starting with the first picture available for the disturbance. The participant analyzed and classified the picture and was then given the picture for the same storm for the next day (one photograph per storm day). He was not permitted to change past classifications based upon later pictures, but retained past photographs; this procedure partially simulated operational conditions. No other information was supplied to the analyst.

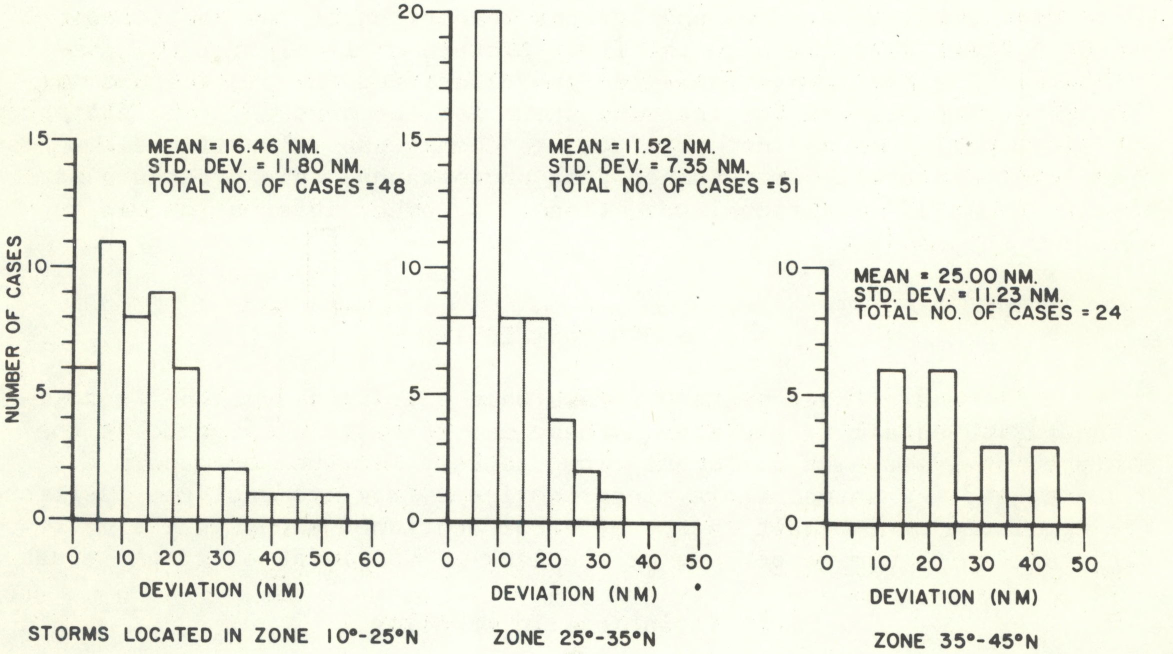
3. STORM POSITIONING

There are three primary errors when a tropical cyclone is positioned using satellite pictures. These are variations in gridding the picture; analysts pick different cloud centers for the same storm and picture; and the actual storm center deviates from its analyzed position. The magnitude of the last source of error mentioned above is the most difficult to determine because of the lack of knowledge of ground truth.

3.1 Variations in Gridding

The gridding error is believed to be dependent upon how close the storm is to land and the distance from the sub-satellite point (the point directly below the satellite) to the feature being positioned. Therefore, the gridding of ATS III pictures should be most accurate at low latitudes near 70 W and least accurate at high latitudes over the oceanic areas. The procedure used to determine the amount of the gridding error for the ATS product was to have four different photo-technicians grid identical sets of pictures (i.e., four possibly different grids for the same picture). Three analysts were then asked to pick the same cloud center on each of the four sets of pictures; the sets consisted of 1 year's worth of storm pictures. Each analyst then determined the location of the cloud center relative to the fixed grid on each of the four pictures. That is, the only variation in location should be from a different placement of the grid on the picture by the four photo-technicians. These data were analyzed using all combinations of photo-technicians as the reference. The statistics came out essentially the same regardless of which photo-technician's grids were used as the reference grid or which analyst did the positioning.

Figure 1 shows the variations in storm locations for one analyst; the data are stratified by latitude. As expected, the greatest variations occurred at the higher latitudes, but the smallest occurred at the mid- rather than low-latitudes. This condition probably resulted from the mid-latitude sample containing several cases in the Gulf of Mexico and along the east coast of the U. S. where land fixes were quite good. On the other hand, some of the lower latitude storms were located in the



STORMS LOCATED IN ZONE 10°-25°N

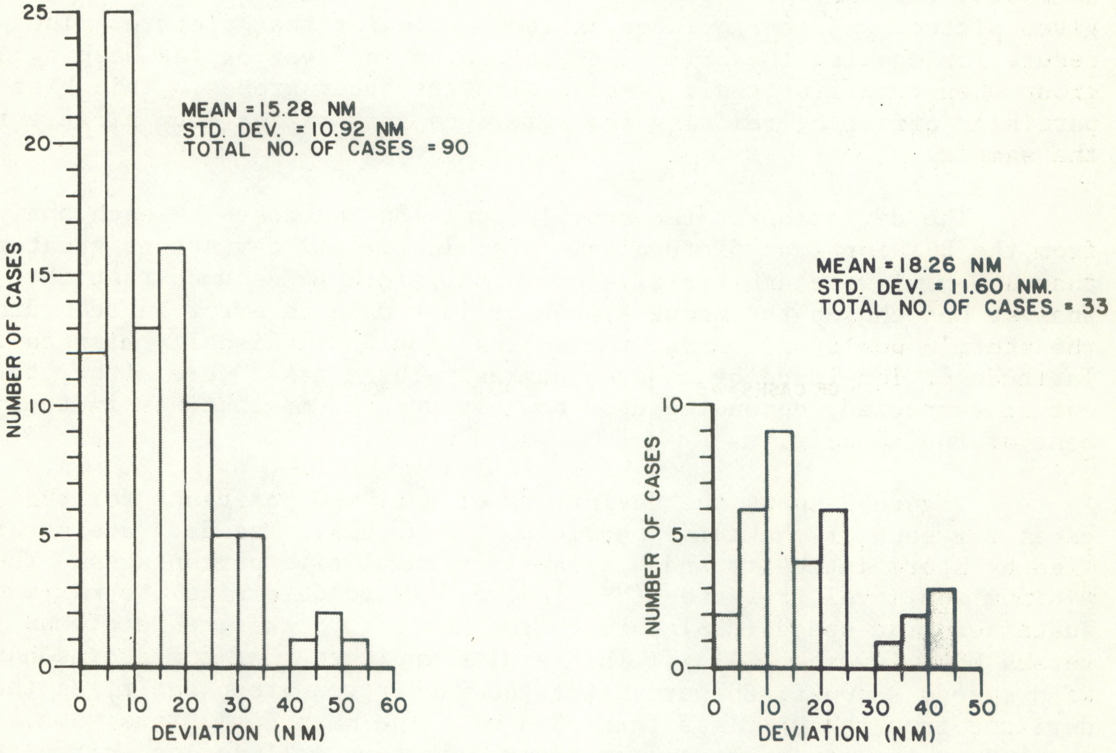
ZONE 25°-35°N

ZONE 35°-45°N

Figure 1. Deviations in storm locations due totally to variations in gridding of ATS satellite photographs stratified by latitudinal zones.

Atlantic away from surrounding land features. Figure 2 shows these same positioning deviations now stratified by proximity to land. As expected, those storms closest to land have the smallest variations. It appears that, in general, the gridding error for the ATS photos is approximately 15 n mi except for storms at high latitudes or at considerable distances from land.

The pictures for the ESSA and DMSP samples were supplied pre-gridded. Considerable effort would have been required to obtain various gridded sets of the same picture. Therefore, no similar test was conducted for these satellites. However, knowledgeable people in the NESS and the AWS estimated the gridding error to be about 30 n mi for the ESSA pictures and about 10 n mi for the DMSP product. These values, of course, would vary depending upon the proximity of land and the sub-satellite point to the feature being positioned.



Deviations in storm location due to variations in gridding for storms located within ~600 nm. of land.

Deviations in storm location due to variations in gridding for storms located greater than ~600 nm. from land.

Figure 2. Deviations in storm locations due totally to variations in gridding of ATS satellite photographs stratified by proximity of land.

3.2 Variations of Storm Locations as Determined by Different Analysts

The second source of positioning error mentioned previously, is caused by the analysts picking a different storm center for the system in the picture. The point of reference used in this portion of the study is referred to as the "best" satellite track (BST), which was obtained from the analysis of the Washington group. We chose this sample primarily because of the greater experience of the analysts in this group and its larger number; thus, this created a larger sample. The BST position was not determined by a mean or mode value, but was subjectively chosen from the grouping of the data. That is, if the sample for a given picture contained 5 analysts and 3 chose a given point, the position determined by the analyst nearest the center of the group was chosen as the point of reference, regardless of the positions chosen by the other two analysts. Therefore, a different analyst's choice might be used for the reference point on successive pictures. In the later analyses, the analyst's position used as the point of reference for a given picture was not included in the sample for that picture. The procedure for chosing the BST builds in a bias in favor of the Washington group when comparing their results with the other groups. This bias is partially offset by removing the reference analyst for each picture from the sample.

The deviation of the storm's position as chosen by each analyst from the BST for each picture was determined. All deviations greater than 150 n mi were arbitrarily thrown out since we assumed that the analyst had chosen the wrong system or just made an error in recording the storm's position. Also, the analyst would occasionally misread the latitude or longitude by an even degree. These cases were either thrown out or corrected, depending upon how obvious the mistake was in the judgement of the authors.

Figure 3 shows the deviations of analyzed positions for the ATS cases for each of the three participating groups. The data are stratified by storm intensity and displayed as cumulative percentages. The minimum sea level pressures (MSLP) were used to determine the maximum sustained wind speed (MWS) through the use of Dvorak's table of MWS versus MSLP (Dvorak, 1973). The results indicate that for storms having wind speeds < 50 kt, 80 percent of the Washington group had fixes that deviated from the BST less than 30 n mi. The mean deviations then decreased as the storms became stronger with more well-defined circulations.

The results of similar analyses for the ESSA data (fig. 4) show much larger deviations for the weaker storms than were indicated for the ATS product. However, for storms having wind speeds > 50 kt the results were nearly the same as those for the ATS sample.

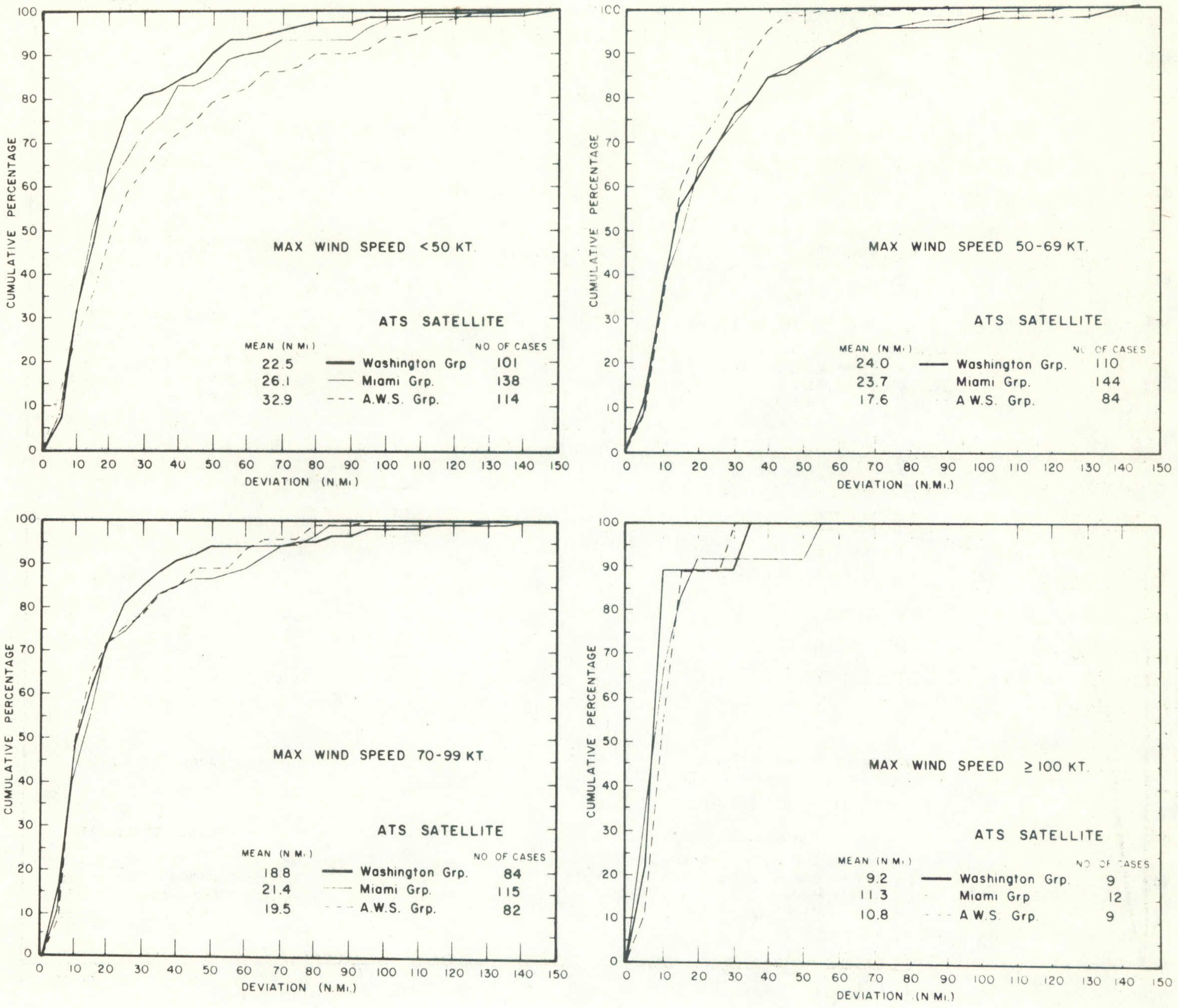


Figure 3. Deviations of analyzed storm location from the "Best Satellite Track" (BST) for the ATS data sample.

Figure 5 shows the results for the DMSP sample. The mean deviations are significantly smaller for medium to strong storms when compared with the ATS and ESSA results. This is probably a direct result of the higher resolution data produced by the DMSP system.

The deviations shown in the three previous figures either could result from a random scatter about the BST position or could be from a bias where a given analyst consistently picked a position that was always away from the reference position in the same direction. In such cases the vector movement would be the same for the reference track and the analyzed track. Some of the statistical prediction methods strongly depend upon the vector movement where as a shift of the fixed vector over

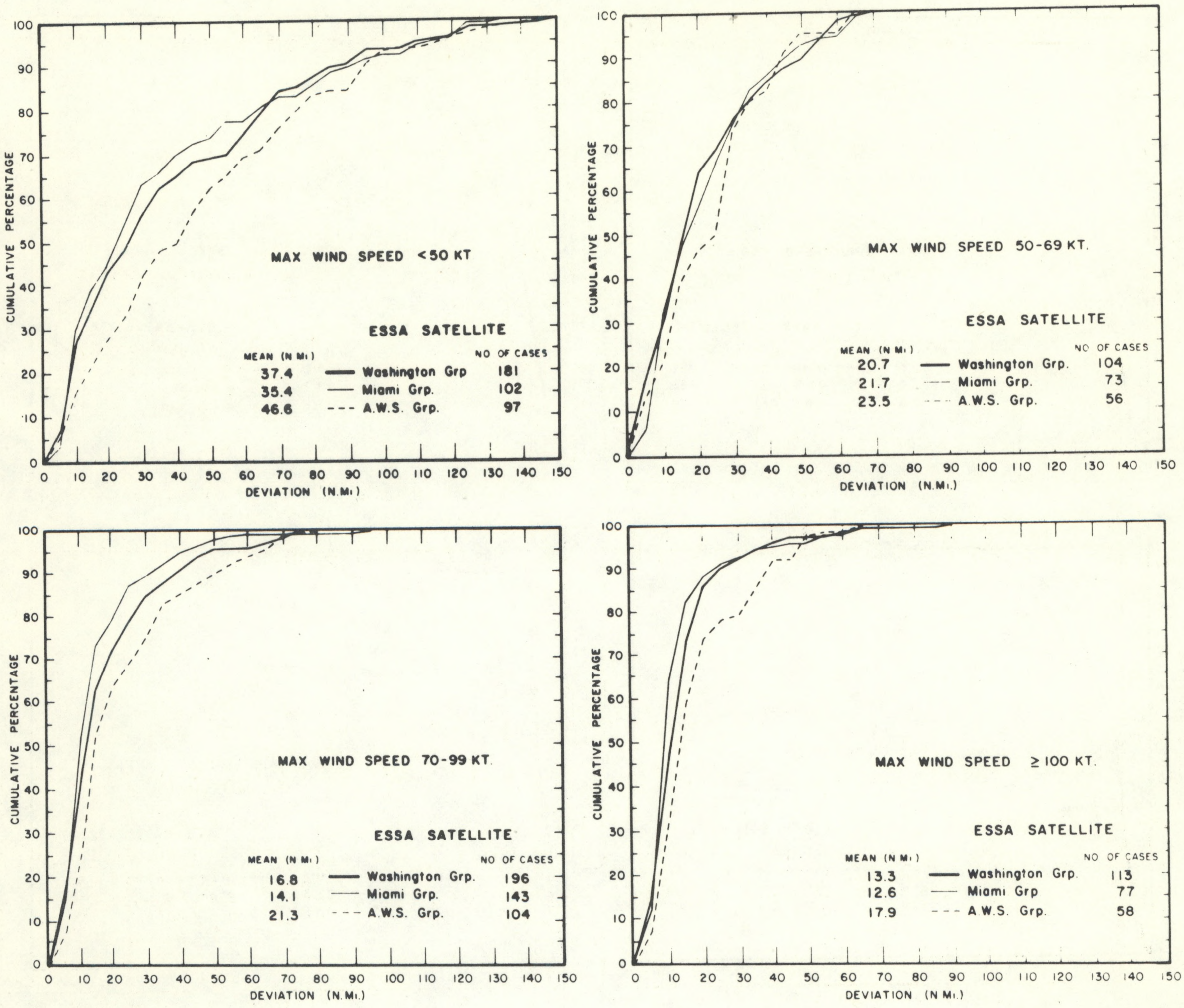


Figure 4. Deviations of analyzed storm location from the "Best Satellite Track" (BST) for the ESSA data sample.

a few miles relative to the earth's surface has very little effect on the predicted movement. Therefore, it is more important for these prediction methods to have relatively accurate vector movements than to have relatively accurate positions.

The difference between the storm center picked by an individual analyst and that determined for the BST at the beginning of each 24 hour period was computed. This algebraic difference was then added to the storm center position picked by the same analyst 24 hours later. The deviation of these adjusted positions from the BST position for the same time (picture) then represents a corrected 24 hour vector deviation where any consistent bias has been removed. Figure 6 shows the results of this

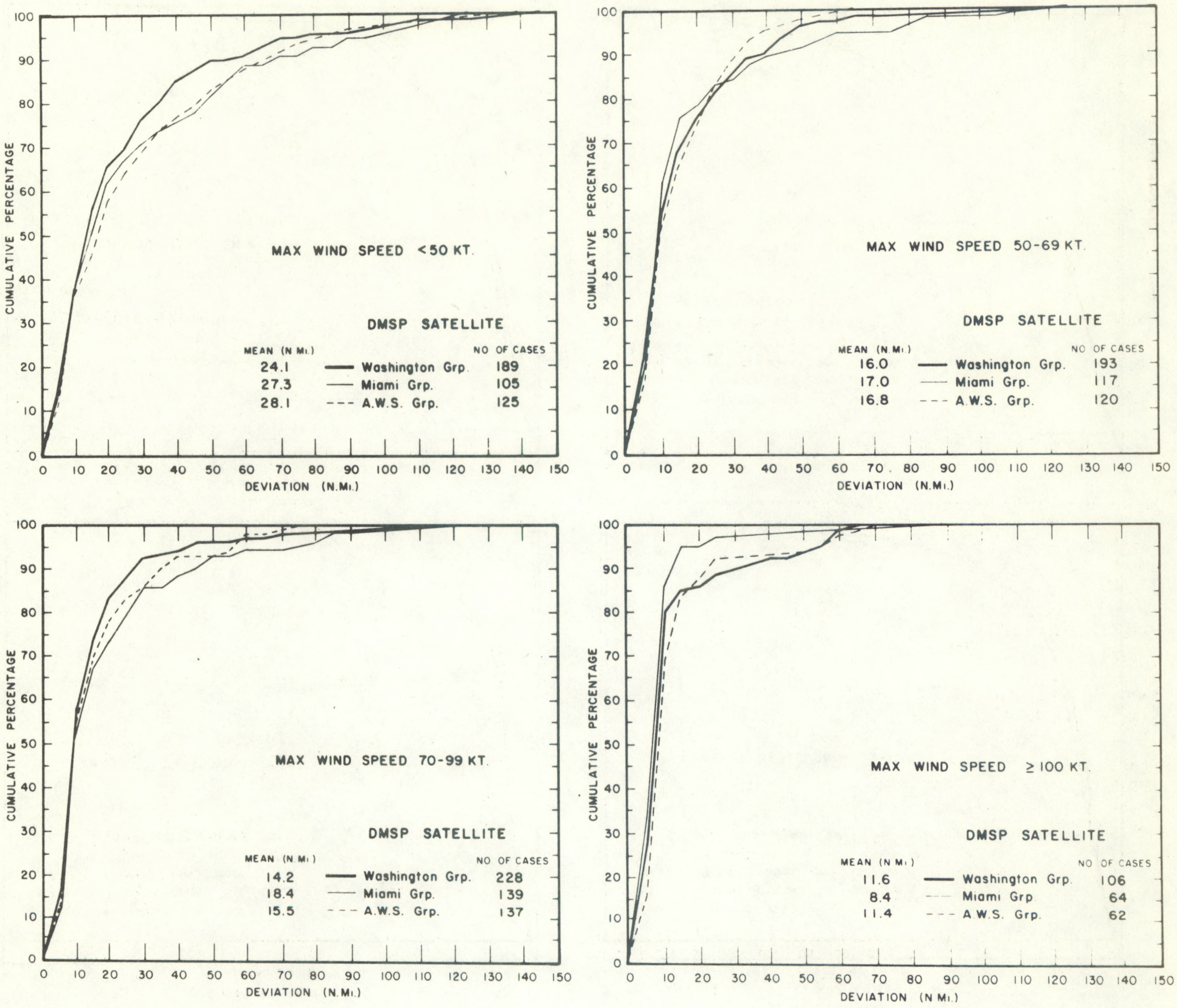


Figure 5. Deviations of analyzed storm locations from the "Best Satellite Track" (BST) for the DMSP data sample.

analysis. In nearly all cases, the mean 24 hour vector deviations are slightly greater than the comparable deviations of analyzed storm positions from the BST for the Washington group. That is, the deviations of analyzed positions from the BST are random.

3.3 Deviations from the Official Track

The third source of positioning errors is the deviation of the "best fit" satellite track from the actual track. The magnitudes of these deviations are only obtainable if ground truth is known. As mentioned earlier, ground truth is nearly impossible to obtain. However, an estimate of ground truth is available in the form of the official

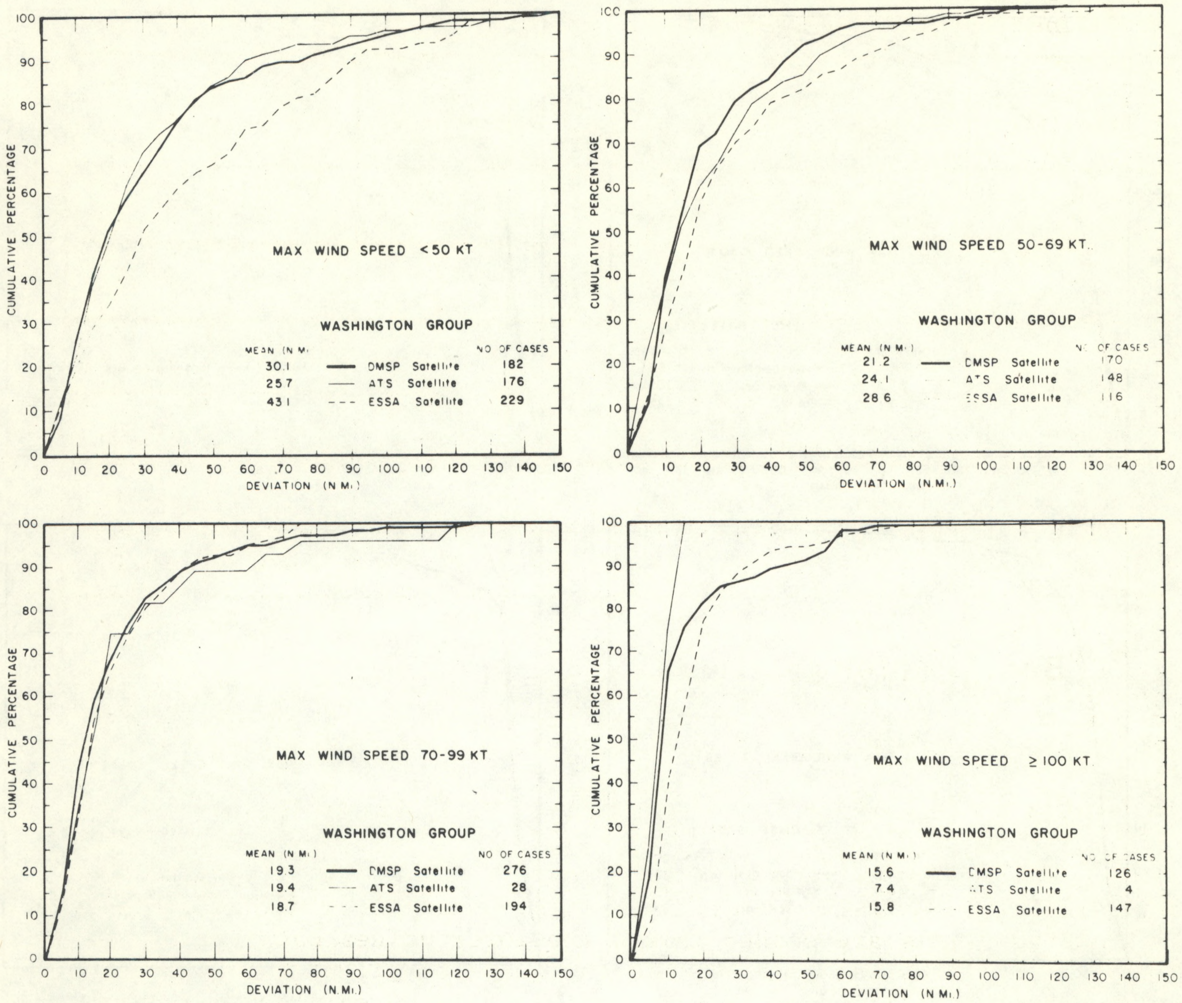


Figure 6. Deviations of the analyzed 24 hour movement vector from the "Best Satellite Track" (BST) for the Washington group and the DMSP, ATS, and ESSA data samples.

track. This is the smooth track constructed after the fact and includes all available information on the storm's location and intensity. Aircraft reconnaissance as well as satellite and land-based radar fixes are used to construct this track. An objective technique which applies various weights to the various types of fixes, is used to construct this track for the Pacific. The track in the Atlantic is made up subjectively by a hurricane forecast specialist. Obviously, a bias can exist in such a scheme, and table 1 indicates a possible bias in favor of positions determined from reconnaissance aircraft over those determined from satellite data.

The mean deviation of the BST positions from the official track positions is about 15 to 20 n mi greater than the mean position scatter among the satellite analysts. The summation of the errors due to scatter

Table 1. Deviations of the "Best Satellite Track" positions and the aircraft reconnaissance positions from the official track for all storms contained in the data sample.

Satellite	Group	No. of Cases	ABS MEAN (NM)	% LT 20 NM	% LT 40 NM	% LT 60 NM
DMSP	WASH.*	147	30.5	32.7	76.2	91.2
	RECON	388	12.6	85.6	98.5	99.5
ESSA	WASH.*	106	40.6	25.5	59.4	81.1
	RECON	370	13.8	82.7	95.9	97.8
ATS	WASH.*	105	37.1	33.3	61.9	84.8
	RECON	556	16.6	72.8	94.4	99.1

*** "Best" Satellite Track**

and due to gridding are comparable with the mean deviation of the BST from the official track.

3.4 Summary of Approximate Storm Positioning Errors

Table 2 summarizes the approximate storm positioning errors as determined from satellite products. This summary is somewhat subjective in that the values are shaded toward the lower values obtained by the groups. This bias was used because that some or all of the analysts rushed some of their measurements due to the large sample. How much this hurried approach affected the results is unknown. However, the mean deviations for all groups and each satellite sample were nearly the same (except for the AWS group with the ESSA satellite sample). The groups were generally more careful during the earlier portions of the testing than during the later portions and they were not always given the data in the same sequence, yet the results being similar would indicate no large degradation as a result of fatigue or rushing. Also, at least two of the Miami analysts worked at a much more leisurely pace and appeared to measure carefully throughout the tests. Their results were similar to those of the other well-trained analysts and again indicate no significant degradation due to fatigue or rushing.

The summary results indicate that the smallest positioning errors were obtained from the DMSP data. This result is probably caused by the higher resolution data obtained with this satellite. The minimum mean potential error, due solely to gridding and variations of analyses

Table 2. Summary of approximate mean absolute storm position errors where columns A and B represent the internal consistency of the satellite analysis

	SATELLITE					
	ATS		ESSA		DMSP	
STORM STRENGTH	Deviation (NM)		Deviation (NM)		Deviation (NM)	
	A	B	A	B	A	B
< 50 KT	25	26	36	43	25	30
50-69 KT	24	24	22	29	16	20
70-99 KT	20	19	16	19	15	19
≥ 100 KT	10	8	13	16	11	16
GRID	15		~ 30 ?		~ 10 ?	
"BEST" SAT. *	37		41		31	
RECON. *	17		14		13	

A - *Deviations from "best" satellite positions.* B - *Deviations of 24 hour movement from "best" satellite track.* * - *Deviations from official track.*

of the satellite products, ranges from approximately 21 to 35 n mi for the DMSP product, 25 to 40 n mi for the ATS pictures, and 43 to 66 n mi for the ESSA data. The smaller deviations are associated with the stronger storms, but the larger values were associated with the weaker systems.

The probable extreme upper bound of the mean absolute error can be obtained by summing the deviation of the BST position from the official track and the mean scatter of analyzed positions about the BST. The BST of course already contains the gridding error. The probable upper bound mean positioning errors would then be 50, 59, and 60 n mi for the DMSP, ATS, and ESSA products, respectively. Slightly larger values would be obtained for weak storms and smaller errors for strong systems.

4. STORM INTENSITY

There are two primary sources of error in determining the strength of a storm when the analyst uses techniques based upon the storm structure as revealed by satellite pictures. These are variations in analyses between analysts and the calibration of the system. The magnitude of the error caused by variations among analysts is rather easily obtained by having several analysts, trained in the technique being tested, analyze a significant number of duplicate satellite photographs of storms. That was the technique used and discussed in section 4.1. The calibration factor, however, is more difficult to obtain. That is, how well does the system intensity as determined from the satellite product, correlate to the actual storm intensity?

For the Dvorak technique, the Dvorak current intensity (CI) number and the actual maximum sustained wind speed (MWS) must be correlated. This requires a knowledge of the true strength of the storm. The MWS has traditionally been determined in two ways, either by the direct measuring of wind speed as reported from aircraft reconnaissance flights into these storms or by some pressure-wind relationship using the minimum sea level pressure (MSLP) as determined from the same reconnaissance flights.

The problems of using the direct reconnaissance reported wind speeds as a measure of storm intensity are many. For instance, the reports often come from different altitudes and different locations within the storm, the wind speeds are determined from various sources such as the state of the sea; Doppler, inertial systems, or combinations of each and; different crews sometimes define the MWS differently. Furthermore, even if all flight crews and organizations used the same equipment and techniques, large variations would still exist. The senior author has found MWS variations over a few minutes as large as 20 to 30 percent of the total wind speed. These variations in wind speed profiles obtained or repetitive tracks of research aircraft through the high energy portion of the storm where no apparent change in the actual intensity of the storm occurred. Figure 7 is a simple schematic illustration of how such observations could occur. The large bold vector represents the hurricane scale flow, while the small vectors represent a superimposed cumulonimbus scale flow at some arbitrary inflow level. The aircraft proceeding outward from the storm's center could pass on the north side of cell A where the two illustrated components act in opposite directions. The track of the inbound leg could be on the south side of cell A where the components act in the same direction. This condition, of course, could be observed with no actual change in the storm's strength and the measurements being made at the same geometric position relative to the storm center, since convective cells often circulate or form and dissipate around the vortex.

Minimum sea level pressures measured by aircraft using dropsondes or interpolated from height deviations of a given pressure surface change slowly and do not have the large fluctuations shown by wind speeds. Also, these measurements are generally done uniformly among organizations and cannot be easily influenced by the observer. Occassionally the calibration can be off by some constant factor that can be determined and corrected. The uniformity and conservativeness of the MSLP is quite evident in the data sample used in this study. The mean deviation of the official pressures from the MSLP's, as determined on reconnaissance aircraft flights, for these cases was generally less than 1 mb. Of course, the official pressure profiles were derived from the aircraft measurements, and therefore we are comparing dependent data. However, similar comparisons for winds show much larger fluctuations.

The authors believe, along with many others, that the MSLP is a much better measure of the strength of the hurricane or typhoon than spot measurements of wind speed. This belief is based upon the above discussion and the high correlation of MSLP to maximum sustained wind speeds (MWS) (Sheets, 1972; Holiday, 1969; Kraft, 1961). Even higher correlations are anticipated for wind speeds filtered to remove the small-scale or high-frequency components. Widespread variations of wind speed versus pressure are often noted during the weak stages of a tropical cyclone's development. These wind speed measurements can be misleading because they often appear to be dominated by contributions from individual thunderstorm cells (see fig. 7) rather than by the circulation of the meso- or larger scale features. The Dvorak technique is based upon the cloud structure of the hurricane or typhoon scale rather than that of individual cells. Therefore, it seems that the more conservative pressure measurements would be more highly correlated with the Dvorak intensity measurements. This was, in fact, generally found to be true for the sample used in this study.

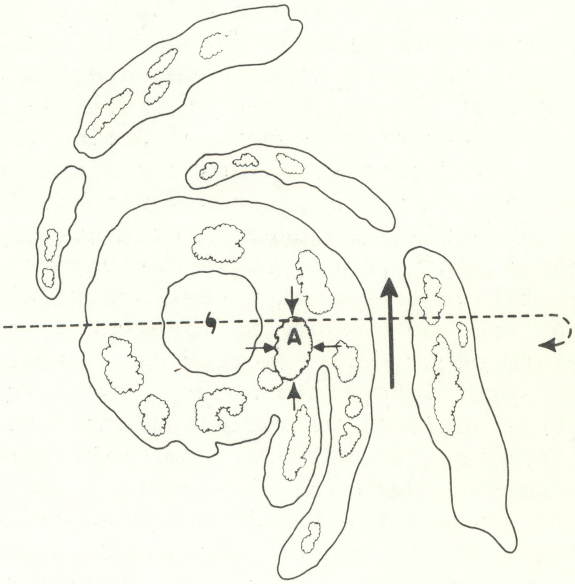


Figure 7. Schematic drawing of a hurricane illustrating probable sources of errors in measuring hurricane scale motions (winds).

Table 3 lists Dvorak's empirically determined relationships of his current intensity (CI) number and the MWS and MSLP. This table was used to convert from MSLP to MWS in the sections that follow. The analyzed CI's were computed versus both the official MWS and

Table 3. Dvorak calibration table of current intensity (CI) number versus the tropical cyclones maximum sustained wind speed (MWS) and minimum sea level pressure (MSLP) (from Dvorak, 1973).

C.I. Number	MWS (Knots)	MSLP (Atlantic)	MSLP (Pacific)
1.5	25K	1010 mb	1004 mb
2	30K	1007 mb	1001 mb
2.5	35K	1003 mb	997 mb
3	40K	998 mb	992 mb
3.5	50K	993 mb	987 mb
4	60K	988 mb	982 mb
4.5	72K	979 mb	973 mb
5	85K	970 mb	964 mb
5.5	97K	960 mb	954 mb
6	110K	948 mb	942 mb
6.5	122K	934 mb	928 mb
7	135K	920 mb	914 mb
7.5	150K	906 mb	900 mb
8	170K	891 mb	885 mb

the MWS derived from the official MSLP. In general, however, only those computations using pressure are presented for all the reasons listed above. Also, since the Dvorak technique uses a picture-comparison method, the analysis for the first picture in each sequence is not included in the results that follow.

4.1 Variations of Storm Intensity as Determined by Different Analysis

Table 4 gives a measure of the internal consistency of the Dvorak intensity classification scheme. Here, the data are stratified by MWS (determined from pressure), satellite, and analysis group. The "Best Satellite Analyses" (BSA) were determined from the Washington sample as described in section 3.2, except that the CI was used rather than the position. This procedure, of course, provides a bias in favor of the Washington group when compared with the other groups. The affects of this bias appears primarily in the first of two categories. Again, the groups analyzed the same sets of photographs so that any variations in results are solely a function of the analysis.

Table 4. *Deviations of the analyzed current intensity (CI) numbers from the "Best Satellite Analyzed" values (measure of internal consistency of the Dvorak scheme). See table 3 for the wind speed or pressure deviations corresponding to CI deviations.*

MAXIMUM WIND SPEED < 50 KNOTS

Satellite	Analysis Group	No. of Cases	Deviations						
			0	± 0.5		± 1.0		± 1.5	
			% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %
DMSP	WASH.	166	53.0	32.5	85.5	12.1	97.6	1.8	99.4
	MIAMI	115	32.2	46.9	79.1	13.9	93.0	6.1	99.1
	AWS	84	31.0	36.9	67.9	20.2	88.1	9.5	97.6
ESSA	WASH.	232	61.6	33.7	95.3	4.3	99.6	0.4	100
	MIAMI	128	42.2	40.6	82.8	14.1	96.9	1.5	98.4
	AWS	99	28.3	32.3	60.6	28.3	88.9	7.1	96.0
ATS	WASH.	170	65.3	30.0	95.3	4.7	100	-	-
	MIAMI	173	36.4	45.7	82.1	14.4	96.5	3.5	100
	AWS	129	33.3	40.2	73.5	18.9	92.4	3.8	96.2

MAXIMUM WIND SPEED 50-69 KNOTS

Satellite	Analysis Group	No. of Cases	Deviations						
			0	± 0.5		± 1.0		± 1.5	
			% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %
DMSP	WASH.	143	53.2	37.0	90.2	9.1	99.3	0.7	100
	MIAMI	101	31.7	46.5	78.2	15.9	94.1	4.9	99.0
	AWS	77	28.6	48.0	76.6	18.2	94.8	5.2	100
ESSA	WASH.	96	52.1	36.4	88.5	11.5	100	-	-
	MIAMI	63	31.8	38.0	69.8	20.7	90.5	4.7	95.2
	AWS	48	31.3	44.6	75.9	7.4	83.3	12.5	95.8
ATS	WASH.	131	60.3	32.8	93.1	5.4	98.5	1.5	100
	MIAMI	139	29.5	58.3	87.8	11.5	99.3	0.7	99.3
	AWS	103	24.8	39.9	64.7	22.0	86.7	10.4	97.1

MAXIMUM WIND SPEED 70-99 KNOTS

Satellite	Analysis Group	No. of Cases	Deviations						
			0	± 0.5		± 1.0		± 1.5	
			% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %
DMSP	WASH.	260	49.6	37.3	86.9	12.0	98.9	1.1	100
	MIAMI	179	28.5	43.0	71.5	17.3	88.8	8.4	97.2
	AWS	135	28.2	40.7	68.9	21.5	90.4	6.6	97.0
ESSA	WASH.	196	49.0	38.2	87.2	12.3	99.5	0.5	100
	MIAMI	112	25.0	38.4	63.4	27.7	91.1	5.3	96.4
	AWS	83	37.4	38.5	75.9	18.1	94.0	4.8	98.8
ATS	WASH.	28	57.1	39.3	96.4	3.6	100	-	-
	MIAMI	28	35.7	42.9	78.6	7.1	85.7	10.7	96.4
	AWS	21	19.1	52.3	71.4	19.1	90.5	9.5	100

MAXIMUM WIND SPEED > 99 KNOTS

Satellite	Analysis Group	No. of Cases	Deviations						
			0	± 0.5		± 1.0		± 1.5	
			% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %
DMSP	WASH.	129	51.9	37.3	89.2	10.8	100	-	-
	MIAMI	91	18.7	41.7	60.4	25.3	85.7	8.8	94.5
	AWS	69	21.7	43.5	65.2	20.3	85.5	11.6	97.1
ESSA	WASH.	145	52.4	32.4	84.8	13.1	97.9	2.1	100
	MIAMI	82	31.7	51.2	82.9	13.9	96.8	2.0	98.8
	AWS	62	27.4	35.5	62.9	30.7	93.6	3.2	96.8
ATS	WASH.	4	50.0	50.0	100	-	-	-	-
	MIAMI	4	0.0	75.0	75.0	0.0	75.0	25.0	100
	AWS	3	33.3	0	33.3	33.4	66.7	33.3	100

The results indicate that 95 to 100 percent of all cases (ATS, ESSA, DMSP), are analyzed to be within ± 1.5 CI numbers of each other. For storms with maximum wind speeds of < 40 kt, this would mean variations of ± 15 kt, and for approximate 100 kt storms variabilities of ± 35 kt would occur less than 5 percent of the time. The internal consistency of the Washington group analyses seems to be best using the ESSA and ATS satellites for the weaker cases (< 50 kt), and using the DMSP and ATS satellites for the stronger cases. The Miami group, in general, shows greater consistencies with the ATS product with which they are most familiar. However, their DMSP results run a close second. The AWS group results are a little more erratic, especially for the ATS and ESSA products.

4.2 Deviations Analyzed CI Number From the Official Intensity

The previous section showed the internal consistency of the Dvorak technique for the three types of satellite products. This section indicates the accuracy of the system versus the official intensity and, for all the reasons mentioned in section 3.0, probably gives a good representation of the absolute accuracy of the technique as applied.

Contingency tables were constructed for each analyst and each satellite; the analyzed values are plotted versus the official intensity and table 3 is used to convert pressure to a CI number. Table 5 shows such a contingency table constructed for a climatological analysis of the DMSP sample. These climatological values were computed for comparison with the results obtained by the various analysts. These analyses were done by using climatological tables constructed by Michaels (1971; 1973) for the Pacific and Atlantic typhoons and hurricanes. These data are stratified by storm intensity, location, and direction of movement. That is, given yesterday's storm intensity, direction of movement, and location, we apply the 24 hour change determined from the climatological sample to obtain today's intensity. In one case, the official pressure is used to obtain yesterday's intensity, the 24 hour climatological change is applied, and table 3 is used to convert to a CI number. In the other case, the official maximum wind speed is used. You should be aware that this is not cumulative blind climatology but dependent climatology. That is the climatology is updated to the "known" storm strength before each 24 hour period. It is in fact a 24 hour forecast based on climatology where the initial condition is known. Therefore, do not conclude that this is how well an analyst could do with climatology alone, but realize that some other source of information is needed to determine the strength of the storm at the initial time. This information is presented here only because most forecasters are familiar with the relative accuracy of a 24 hour forecast based on climatology and can therefore compare a known quantity with the skill of the analyst. Cumulative blind climatology would be equivalent to 24 hour, 48 hour, 72 hour, etc. forecasts based upon climatology alone and such forecasts are known to deteriorate rapidly with time.

Table 5. A contingency table constructed for climatology and each analyst for each satellite data sample to obtain the skill scores etc. shown in tables 6 through 11. The number in each block is the number of cases that fell in the particular square.

24 HOUR UPDATED PACIFIC CLIMATOLOGY C.I.

OFFICIAL OBSERVED C.I.

	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	TOTALS
1.5		1	4												5
2.0		6	3	4											13
2.5		1	1	3	3				1						9
3.0			1	2	6										9
3.5		1	1	6	9	1	4								22
4.0		1		5	6	9	7	4							32
4.5					4	6	11	10			2				33
5.0						6	8	7	8	1	1				31
5.5						1	1	3	6	5	2	2			20
6.0								2	6	1	6	1			16
6.5									1	1	4	4			10
7.0						1				3	1		1		6
7.5										1		1			2
8.0															0
TOTALS	0	10	10	20	28	24	31	26	22	12	16	8	1	0	56/208
BIAS	0	.77	1.11	2.22	1.27	.75	.94	.84	1.1	.75	1.6	1.3	.5	%	COR=26.9

Table 6 shows the results of the intensity analyses for the ATS satellite sample. The analysts are listed by letter and group. Analysts K through Q comprised the Washington group and analysts G, H, J, and R were Miami and S, T, and U made up the AWS groups. The category deviations correspond to the contingency tables where 0 indicates that the analyzed value is the same as the observed or official value, catagories 1, 2, 3, 4, and 5 correspond to analyzed values deviating from the official value by ±0.5, ±1.0, ±1.5, ±2.0, and ±2.5 CI numbers. The deviation skill score was computed using the formulation derived by Vernon (1953). This skill score is similar to the Heidke skill score except that the greater the deviation of the analyzed value from the observed value, the greater the penalty. The skill is computed versus that value expected by chance and versus that value expected by the climatology of the sample. The final two columns measure the bias of the analyst. The column headed LT means that the analyst gave a CI value less than what was observed; i.e., underestimated the strength of the storm, and vica versa for GT (greater than).

Table 6. *Deviations of Analyzed Current Intensity From Official Intensity (Pressure)-ATS-Atlantic***

Analyst	No. of Cases	Number of Category Deviations (C.I.)												Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥6				
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs.	GT Obs.
Clima. (Press)	91	38.5	47.2	85.7	9.9	95.6	2.2	97.8	0	97.8	1.1	98.9	1.1	.570*	.578*	28.6	33.0
Clima. (Wind)	91	18.7	42.8	61.5	29.7	91.2	6.6	97.8	0.9	98.9	0	98.9	1.1	.384*	.345*	8.8	72.5
K	82	32.9	40.3	73.2	18.3	91.5	4.8	96.3	2.5	98.8	1.2	100	-	.507	.488	24.4	42.7
L	88	28.4	42.1	70.5	17.0	87.5	10.2	97.7	2.3	100	-	-	-	.452	.430	28.4	43.2
M	77	37.7	41.5	79.2	16.9	96.1	1.3	97.4	1.3	98.7	1.3	100	-	.560	.564	24.7	37.7
N	86	19.8	48.8	68.6	16.3	84.9	11.6	96.5	3.5	100	-	-	-	.374	.362	38.4	41.9
G	87	31.0	39.1	70.1	26.5	96.6	1.1	97.7	2.3	100	-	-	-	.490	.487	42.5	26.4
H	91	30.8	40.6	71.4	15.4	86.8	11.0	97.8	2.2	100	-	-	-	.465	.443	15.4	53.8
J	91	28.6	43.9	72.5	22.0	94.5	4.4	98.9	1.1	100	-	-	-	.470	.481	36.3	35.2
R	75	26.7	37.3	64.0	21.3	85.3	10.7	96.0	4.0	100	-	-	-	.340	.400	52.0	21.3
S	91	15.4	37.3	52.7	20.9	73.6	9.9	83.5	5.5	89.0	3.3	92.3	7.7	.305	.032	9.9	74.7
T	88	15.9	36.4	52.3	27.2	79.5	13.7	93.2	6.8	100	-	-	-	.290	.227	34.1	50.0
U	82	24.4	46.3	70.7	13.4	84.1	13.5	97.6	2.4	100	-	-	-	.430	.396	41.5	34.1

* *These skill scores do not represent values for cumulative blind climatology. See section 4.2 for a detailed explanation before drawing conclusions about the relative accuracy of climatology versus the satellite analyzed values.*

** *The MSLP and table 3 are used to determine the cyclone intensity and the corresponding CI number. A category deviation of 1 represents a CI number deviation of ±0.5, 2 is ±1.0 etc.*

The direct comparison of the dependent climatological skill scores with the analyst's skill scores (tables 6 through 8) is unfair as indicated above. That is, the climatological computations are based upon a knowledge of the storm's intensity, 24 hours before each analysis period. The analyst has only past and present satellite photographs of the disturbance upon which to base his intensity estimate. However, the comparison does give a relative indication of the effectiveness of the analysis scheme, which has meaning to most operational forecasters.

The results for the ATS data indicate comparable skill scores for the Washington and Miami groups except for two analysts. It is also interesting that the Washington group tended to overestimate the strength of the storm, but the Miami group's estimates showed more variability. However, we find that 85 to 95 percent of the sample for this satellite was analyzed to be within ±1 CI number of the official value. Also, only two cases out of more than 670 deviated by more than 2 CI numbers from the official value.

Table 7 shows the intensity analysis results for the ESSA satellite sample. Here, more analysts were available for the Washington group. The skill scores are considerably higher than for the ATS data, but a greater percentage of misses in the higher categories also occurred. That is, the number of misses greater than ±1 CI number ranges from 10 to 20 percent of the total sample as compared with 5 to 15 percent for the ATS sample. The skill scores are higher because the data cover a wider range of values than those of the ATS sample and are therefore less likely to be picked by chance. The scores are also quite uniform and comparable with the dependent climatological[2] scores showing essentially equal skill for most of the analysts. Also, except for two or three analysts, the bias was less than for the ATS sample.

Table 8 shows the results for the DMSP sample. Here, the skill scores again are generally uniform and comparable with the dependent climatology score, except for a few of the analysts. The scores are slightly smaller than for the ESSA sample, but larger than for the ATS sample. In general, the number of misses in the higher categories for the Washington group is less for these cases than for the ESSA sample. Also, the biases for most analysts are smaller than for the ATS or ESSA samples.

Tables 9, 10, and 11 show the skill in determining the 24 hour change in intensity by use of the Dvorak technique. That is, the analyzed 24 hour change (today's CI number minus yesterday's CI number) is compared with the official change for the same time period. Here,

---

[2] See previous explanation of dependent climatology in section 4.2.

Table 7. *Deviations of Analyzed Current Intensity From Official Intensity (Pressure) -- ESSA-Atlantic-Pacific***

Analyst	No. of Cases	Number of Category Deviations (C.I.)											Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6			
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Clima. (Press)	93	25.8	40.9	66.7	21.5	88.2	7.5	95.7	4.3	100	-	-	-	.640*	.651*	40.9	33.3
Clima. (Wind)	96	22.9	50.0	72.9	15.6	88.5	8.4	96.9	3.1	100	-	-	-	.648*	.669*	34.4	42.7
K	92	30.4	32.6	63.0	21.8	84.8	8.7	93.5	5.4	98.9	0.0	98.9	1.1	.600	.636	43.5	26.1
L	93	25.8	37.6	63.4	22.6	86.0	9.7	95.7	4.3	100	-	-	-	.628	.641	33.3	40.9
M	98	28.6	46.9	75.5	11.2	86.7	8.2	94.9	5.1	100	-	-	-	.674	.690	35.7	35.7
N	101	24.8	39.6	64.4	19.8	84.2	9.9	94.1	4.9	99.0	1.0	100	-	.644	.636	33.7	41.6
O	97	24.7	39.2	63.9	16.5	80.4	11.4	91.8	5.1	96.9	3.1	100	-	.601	.611	26.8	48.5
P	97	32.0	41.2	73.2	15.5	88.7	3.1	91.8	6.1	97.9	1.1	99.0	1.0	.674	.680	24.7	43.3
Q	95	27.4	37.9	65.3	24.2	89.5	4.2	93.7	2.1	95.8	2.1	97.9	2.1	.635	.641	35.8	36.8
G	94	31.9	34.1	66.0	15.9	81.9	9.6	91.5	5.3	96.8	2.3	98.9	1.1	.622	.634	51.1	17.0
H	101	29.7	37.6	67.3	15.9	83.2	9.9	93.1	2.9	96.0	3.0	99.0	1.0	.639	.644	30.7	39.6
J	101	24.8	41.5	66.3	16.9	83.2	10.9	94.1	4.9	99.0	1.0	100	-	.614	.638	39.6	35.6
R	97	15.5	31.9	47.4	29.9	77.3	13.4	90.7	7.2	97.9	1.1	99.0	1.0	.508	.524	60.8	23.7
S	101	18.8	36.6	55.4	16.9	72.3	10.9	83.2	10.9	94.1	3.9	98.0	2.0	.505	.514	24.8	56.4
T	101	17.8	42.6	60.4	17.8	78.2	6.9	85.1	9.9	95.0	3.0	98.0	2.0	.526	.549	28.7	53.5
U	97	19.6	38.1	57.7	20.7	78.4	15.4	93.8	5.2	99.0	1.0	100	-	.547	.582	48.5	32.0

* *These skill scores do not represent values for cumulative blind climatology. See section 4.2 for a detailed explanation before drawing conclusions about the relative accuracy of climatology versus the satellite analyzed values.*

** *Same as for table 6 except that the analyses are for the ESSA satellite data sample.*

Table 8. *Deviations of Analyzed Current Intensity From Official Intensity (Pressure) -- DMSP-Pacific***

Analyst	No. of Cases	Number of Category Deviations (C.I.)												Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6				
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Clima. (Press)	122	32.8	43.4	76.2	12.3	88.5	7.4	95.9	3.3	99.2	.8	100	--	.677*	.670*	39.3	27.9
Clima. (Wind)	125	20.0	40.8	60.8	25.6	86.4	8.0	94.4	4.8	99.2	0	99.2	.8	.540*	.566*	32.0	48.0
K	116	21.6	27.5	49.1	26.8	75.9	14.6	90.5	8.6	99.1	.9	100	--	.451	.502	52.6	25.9
L	117	29.9	38.5	68.4	21.3	89.7	7.7	97.4	2.6	100	--	--	--	.611	.646	30.8	39.3
M	107	29.0	42.0	71.0	20.6	91.6	5.6	97.2	1.9	99.1	.9	100	--	.601	.636	32.7	38.3
N	121	26.4	42.2	68.6	18.2	86.8	9.1	95.9	3.3	99.2	.8	100	--	.600	.617	35.5	38.0
O	122	19.7	37.7	57.4	30.3	87.7	9.0	96.7	1.7	98.4	.8	99.2	.8	.518	.561	41.8	38.5
P	121	30.6	36.3	66.9	24.0	90.9	6.6	97.5	1.7	99.2	.8	100	--	.610	.643	35.5	33.9
G	122	21.3	36.1	57.4	19.6	77.0	12.3	89.3	9.9	99.2	.8	100	--	.494	.517	48.4	30.3
H	122	23.0	43.4	66.4	18.0	84.4	12.3	96.7	1.7	98.4	.8	99.2	.8	.547	.585	41.8	35.2
J	123	24.4	22.8	47.2	31.7	78.9	13.8	92.7	5.7	98.4	1.6	100	--	.418	.507	40.7	35.0
R	123	11.4	21.1	32.5	30.1	62.6	19.5	82.1	13.8	95.9	2.5	98.4	1.6	.298	.324	74.0	14.6
S	124	21.8	35.5	57.3	22.5	79.8	15.4	95.2	2.4	97.6	2.4	100	--	.495	.537	39.5	38.7
T	124	19.4	28.2	47.6	18.5	66.1	21.8	87.9	7.3	95.2	3.2	98.4	--	.372	.421	35.5	45.2
U	120	18.3	31.7	50.0	24.2	74.2	18.3	92.5	5.0	97.5	1.7	99.2	.8	.452	.468	62.5	19.2

* *These skill scores do not represent values for cumulative blind climatology. See section 4.2 for a detailed explanation before drawing conclusions about the relative accuracy of climatology versus the satellite analyzed values.*

** *Same as for table 6 except that the analyses are for the DMSP satellite data sample.*

*Table 9. Deviations of Analyzed 24 Hour Change of Intensity From Official Change (Pressure) -- ATS-Atlantic **

		Number of Category Deviations (C.I.)												Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6				
Analyst	No. of Cases	% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs.	GT Obs.
Clima. (Press)	73	28.8	47.9	76.7	12.3	89.0	5.5	94.5	4.1	98.6	0	98.6	1.4	.221	.197	41.1	30.1
Clima. (Wind)	73	34.2	35.7	69.9	17.8	87.7	6.8	94.5	2.8	97.3	1.3	98.6	1.4	.203	.168	34.2	31.5
K	60	35.0	53.3	88.3	8.4	96.7	3.3	100	-	-	-	-	-	.379	.373	31.7	33.3
L	69	47.8	34.8	82.6	14.5	97.1	2.9	100	-	-	-	-	-	.483	.477	31.9	20.3
M	58	48.3	41.4	89.7	10.3	100	-	-	-	-	-	-	-	.520	.547	31.0	20.7
N	66	34.8	53.1	87.9	10.6	98.5	1.5	100	-	-	-	-	-	.393	.409	39.4	25.8
G	68	36.8	48.5	85.3	13.2	98.5	0	98.5	1.5	100	-	-	-	.386	.416	30.9	32.4
H	73	42.5	39.7	82.2	15.1	97.3	2.7	100	-	-	-	-	-	.451	.455	32.9	24.7
J	73	34.2	46.6	80.8	17.8	98.6	0	98.6	1.4	100	-	-	-	.342	.388	30.1	35.6
R	58	24.1	50.0	74.1	20.7	94.8	5.2	100	-	-	-	-	-	.181	.227	36.2	39.7
S	73	35.6	32.9	68.5	24.7	93.2	5.4	98.6	1.4	100	-	-	-	.336	.273	30.1	34.2
T	70	24.3	45.7	70.0	17.1	87.1	11.5	98.6	1.4	100	-	-	-	.244	.160	31.4	44.3
U	60	28.3	41.7	70.0	26.7	96.7	1.6	98.3	1.7	100	-	-	-	.177	.271	36.7	35.0

** Same as for table 6 except that the deviations are for 24 hour changes only.*

Table 10. *Deviations of Analyzed 24 Hour Change of Intensity From Official (Pressure) -- ESSA-Atlantic-Pacific**

Analyst	No. of Cases	Number of Category Deviations (C.I.)											Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6			
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Clima. (Press)	82	29.3	28.0	57.3	23.2	80.5	9.7	90.2	6.1	96.3	0.0	96.3	3.7	.249	.308	41.5	29.3
Clima. (Wind)	85	21.2	45.9	67.1	21.1	88.2	7.1	95.3	4.7	100	-	-	-	.330	.385	38.8	40.0
K	77	33.8	32.4	66.2	23.4	89.6	7.8	97.4	0.0	97.4	0.0	97.4	2.6	.343	.418	27.3	39.0
L	81	24.7	43.2	67.9	22.2	90.1	5.0	95.1	3.7	98.8	0.0	98.8	1.2	.329	.367	35.8	39.5
M	86	31.4	36.0	67.4	22.1	89.5	7.0	96.5	2.3	98.8	1.2	100	-	.398	.425	31.4	37.2
N	90	22.2	44.5	66.7	18.9	85.6	7.7	93.3	5.6	98.9	0.0	98.9	1.1	.380	.391	42.2	35.6
O	86	24.4	41.9	66.3	22.1	88.4	5.8	94.2	3.5	97.7	1.1	98.8	1.2	.318	.366	39.5	36.0
P	85	36.5	37.6	74.1	15.3	89.4	2.4	91.8	7.0	98.8	0.0	98.8	1.2	.431	.451	31.8	31.8
Q	82	26.8	40.3	67.1	20.7	87.8	6.1	93.9	4.9	98.8	1.2	100	-	.357	.382	39.0	34.1
G	83	25.3	49.4	74.7	15.7	90.4	6.0	96.4	1.2	97.6	1.2	98.8	1.2	.386	.439	36.1	38.6
H	90	30.0	36.7	66.7	18.9	85.6	8.8	94.4	4.5	98.9	1.1	100	-	.383	.411	34.4	35.6
J	90	28.9	44.4	73.3	16.7	90.0	5.6	95.6	3.3	98.9	0.0	98.9	1.1	.432	.486	31.1	40.0
R	84	33.3	41.7	75.0	17.9	92.9	2.3	95.2	2.4	97.6	1.2	98.8	1.2	.429	.497	28.6	38.1
S	90	17.8	44.4	62.2	23.4	85.6	4.4	90.0	6.7	96.7	1.1	97.8	2.2	.301	.326	41.1	41.1
T	90	21.1	33.3	54.4	25.6	80.0	13.3	93.3	2.3	95.6	3.3	98.9	1.1	.293	.296	35.6	43.3
U	85	24.7	36.5	61.2	22.3	83.5	9.4	92.9	3.6	96.5	2.3	98.8	1.2	.229	.324	37.6	37.6

* *Same as for table 9 except that the analyses are for the ESSA satellite data sample.*

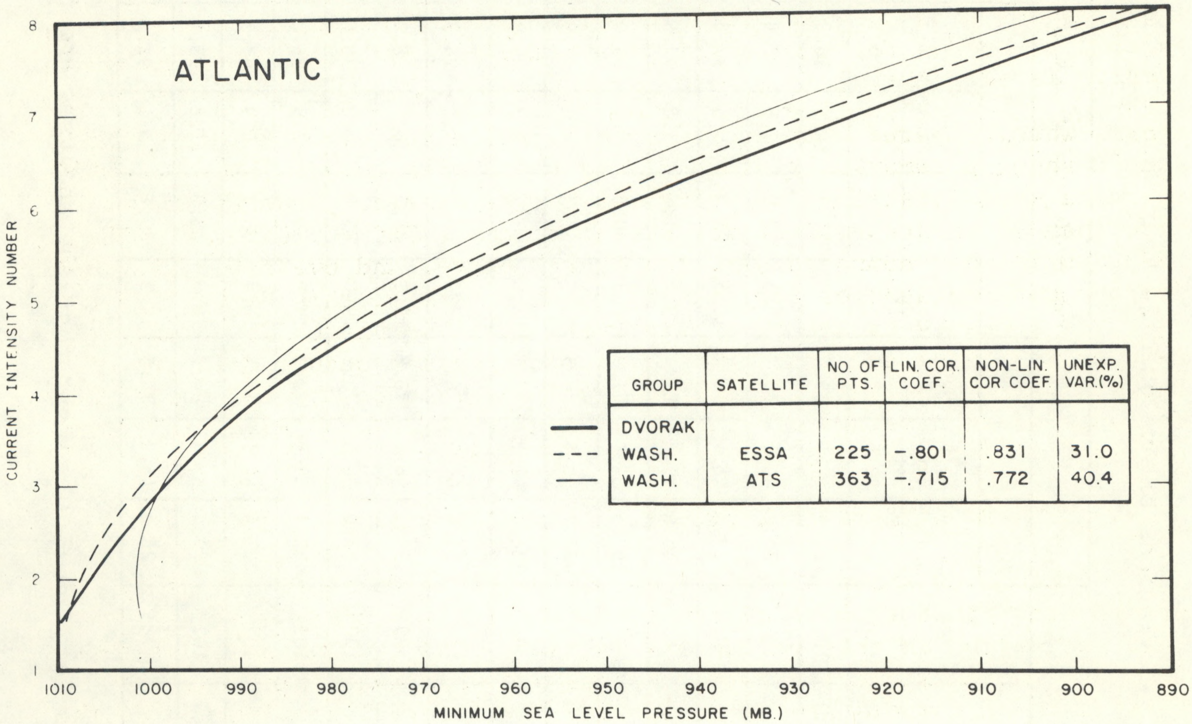
*Table 11. Deviations of Analyzed 24 Hour Change of Intensity From Official Change (Pressure) -- DMSP-Pacific**

Analyst	No. of Cases	Number of Category Deviations (C.I.)											Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6			
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Clima. (Press)	101	22.8	42.5	65.3	19.8	85.1	5.0	90.1	5.9	96.0	1.0	97.0	3.0	.241	.244	34.7	42.6
Clima. (Wind)	104	27.9	37.5	65.4	16.3	81.7	14.5	96.2	0	96.2	.9	97.1	2.9	.249	.282	31.7	40.4
K	90	32.2	44.5	76.7	13.3	90.0	7.8	97.8	1.1	98.9	1.1	100	-	.389	.485	26.7	41.1
L	91	29.7	46.1	75.8	16.5	92.3	5.5	97.8	1.1	98.9	1.0	100	-	.414	.478	38.5	31.9
M	85	30.6	51.8	82.4	9.4	91.8	5.8	97.6	1.2	98.8	1.2	100	-	.459	.511	25.9	43.5
N	96	37.5	34.4	71.9	18.7	90.6	8.4	99.0	1.0	100	-	-	-	.415	.476	32.3	30.2
O	98	32.7	43.8	76.5	15.3	91.8	4.1	95.9	3.1	99.0	1.0	100	-	.382	.476	28.6	38.8
P	96	34.4	42.7	77.1	14.6	91.7	7.3	99.0	0	99.0	1.0	100	-	.437	.501	27.1	38.5
G	98	29.6	46.9	76.5	13.3	89.8	5.1	94.9	2.0	96.9	3.1	100	-	.371	.438	31.6	38.8
H	99	30.3	46.5	76.8	13.1	89.9	8.1	98.0	1.0	99.0	1.0	100	-	.416	.459	30.3	39.4
J	100	34.0	36.0	70.0	19.0	89.0	4.0	93.0	6.0	99.0	1.0	100	-	.319	.417	27.0	39.0
R	101	32.7	39.6	72.3	19.8	92.1	2.9	95.0	2.0	97.0	3.0	100	-	.328	.435	26.7	40.6
S	102	31.4	44.1	75.5	14.7	90.2	6.9	97.1	1.9	99.0	1.0	100	-	.435	.460	32.4	36.3
T	102	29.4	35.3	64.7	18.6	83.3	11.8	95.1	2.9	98.0	2.0	100	-	.334	.346	30.4	40.2
U	98	28.6	45.9	74.5	15.3	89.8	3.1	92.9	7.1	100	-	-	-	.372	.433	29.6	41.8

** Same as for table 10 except that the analyses are for the DMSP satellite data sample.*

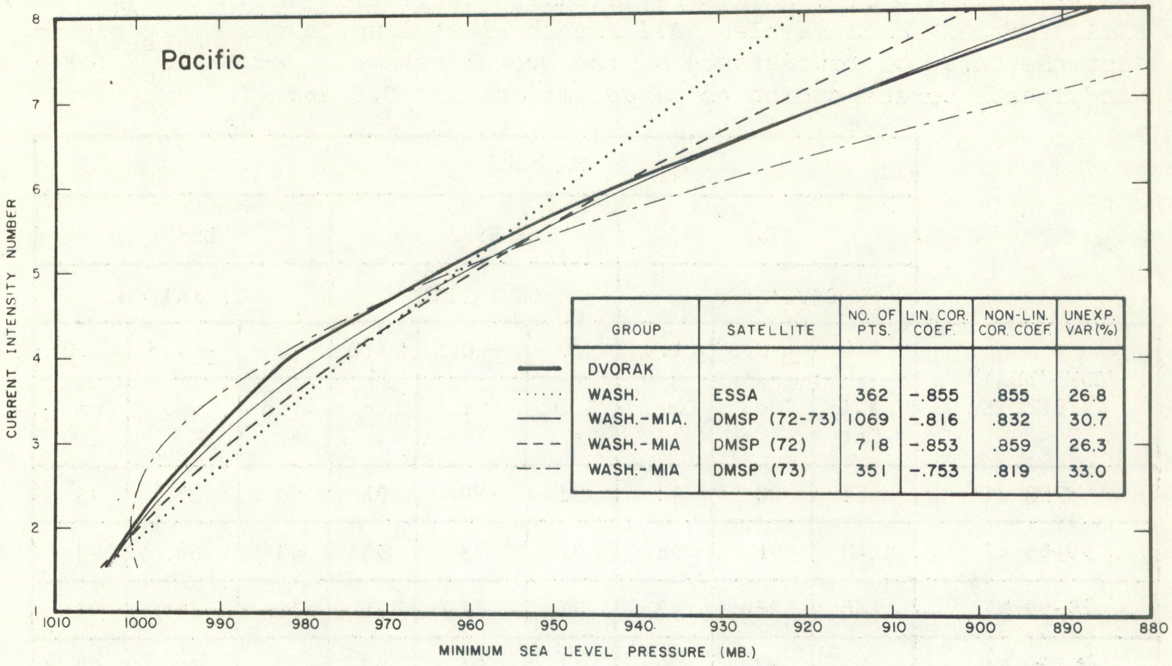
a direct comparison can be made with the climatological values. Also, constant biases no longer affect the results. The skill scores seem to be more variable for the 24 hour change data than for the data shown in the three previous tables. However, note that the average analyst's skill scores are consistently much higher than those obtained for climatology. The deviations are smallest for the ATS sample and largest for the ESSA data. Remember that the ESSA and DMSP samples contained more strong storms than the ATS sample.

A least-square parabola of CI number versus MSLP was computed for the ESSA and ATS Atlantic cases from the Washington group data (fig. 8). These parabolas are compared with Dvorak's curve plotted from table 3. The small overestimation bias previously mentioned is shown for both cases. The linear and non-linear correlation coefficients are relatively large, but some 30 to 40 percent of the variance remains unexplained by this fit. Figure 9 shows similar computations for the ESSA and DMSP Pacific cases. The ESSA data were those of the Washington group, while the DMSP fit was obtained using the Miami and Washington group, but excluding analysts K and R who showed strong biases. The ESSA Pacific data show almost a straight line fit considerably different than for the ESSA Atlantic cases. This curve indicates a large overestimate of the intensity for strong storms and an underestimate for weak



	GROUP	SATELLITE	NO. OF PTS.	LIN. COR. COEF.	NON-LIN. COR COEF	UNEXP. VAR.(%)
—	DVORAK					
- - -	WASH.	ESSA	225	-.801	.831	31.0
—	WASH.	ATS	363	-.715	.772	40.4

Figure 8. Parabolic least-squares fit of the Washington group analyzed current intensity (CI) numbers versus minimum sea level pressures for the Atlantic data sample. The Dvorak calibration curve is illustrated as a point of reference.



GROUP	SATELLITE	NO. OF PTS.	LIN. COR. COEF.	NON-LIN. COR. COEF.	UNEXP. VAR (%)
DVORAK					
WASH.	ESSA	362	-.855	.855	26.8
WASH.-MIA.	DMSP (72-73)	1069	-.816	.832	30.7
WASH.-MIA	DMSP (72)	718	-.853	.859	26.3
WASH.-MIA	DMSP (73)	351	-.753	.819	33.0

Figure 9. Parabolic least squares fit for the Washington and Miami groups analyzed current intensity (CI) numbers versus minimum sea level pressures for the Pacific data sample. The Dvorak calibration curve is illustrated as a point of reference.

storms when compared with Dvorak's calibration curve. The total DMSP sample shows a remarkably close fit to the Dvorak curve. However, when the data were stratified by year, the curves varied. Underestimates of more than 20 mb are evident for strong storms (MSLP > 920 mb) for the 1973 sample when compared with the Dvorak curve, and overestimates result for the same catagories for the 1972 data. Part of this disparity, however, is probably caused by a relatively small sample of low pressures in the 1973 cases and that extreme intensification occurred in two of those storms. However, these trends are observed in the 920 mb to 950 mb range where a relatively large sample was available indicating some problems with the strong and very strong cases.

4.3 Summary of Approximate Storm Intensity Errors

Table 12 shows a summary of the approximate storm intensity errors determined from satellite products. This summary again is somewhat subjective in that values are shaded toward the consensus of the analysts who obtained the better scores and were least biased. The first sets of numbers are those showing the internal consistency of the scheme, while the last line represents an estimate of the absolute accuracy of the system as applied by skilled analysts. The internal consistency values are 10 to 15 percent higher and 5 to 10 percent higher than the analyzed values versus the official intensity for CI number

Table 12. Summary of approximate intensity estimation errors for the ATS, ESSA, and DMSP data sample. All but the last line in the table represent the internal consistency of the Dvorak scheme. See table 3 for wind speeds corresponding to CI deviations of ±0.5 and ±1.0.

	SATELLITE								
	ATS			ESSA			DMSP		
	DEVIATION			DEVIATION			DEVIATION		
FROM "BEST" SATELLITE	0	± 0.5	± 1.0	0	± 0.5	± 1.0	0	± 0.5	± 1.0
	% in Cat	Cum. %	Cum. %	% in Cat	Cum. %	Cum. %	% in Cat	Cum. %	Cum. %
< 50 KT	51	88	98	52	88	98	40	82	95
50-69 KT	45	91	98	42	79	95	43	84	97
70-99 KT	46	88	93	37	77	95	38	79	93
≥ 100 KT	50	87	100	42	84	97	37	72	92
FROM OFFICIAL	32	73	92	28	67	85	25	63	86

deviations of ±0.5 and ±1.0, respectively. The deviations being slightly larger for the ESSA and DMSP as compared with the ATS sample data is probably caused by the greater number of more intense storms in the ESSA and DMSP samples. This conclusion results from the quadratic fit of the data implying greater problems with the upper end of the scale (stronger storms) than the lower end (weaker storms).

5. CONCLUSIONS

The position errors summarized in table 2 show no major surprises. That is, the higher resolution data result in smaller analysis errors. Also, as storms become stronger and more well defined or closer to land, the position errors decrease. The Air Weather Service (1974) indicated mean position errors of from 15 to about 50 n mi for operational fixes obtained with the DMSP satellite photographs. The low values correspond to eye fix positions with geographical gridding, while the upper values correspond to poorly defined circulation centers. These values are comparable with those obtained by a summation of the grid and internal consistency errors (table 2) that range from 20 to 35 n mi.

The major questions that remain to be answered are how can these positioning errors be decreased and how well can we expect to do in the future? One of the dominant sources of error shown in the previous sections was the variations in center positions caused by different analyses of the same picture. This error is particularly large for the weaker and less well-defined systems and should be considerably decreased by the use of the GOES photographs. The movie loop capability of the GOES product which has the resolution of the DMSP, and the enhancement techniques now becoming available to operational units should make locating the center of the circulation less ambiguous. The "scatter" error for the less well-defined storms should then approach that of the well-defined ones. Also, as was mentioned, the vector movement of the storm may be more important for some applications than the approximate location of the instantaneous circulation center. Movie loops should help in obtaining this vector movement. In addition, defining some cloud envelope as representing the hurricane scale circulation, possibly could make these measurements even more consistent. This has been previously suggested by Dvorak and others, but more work must be done in this area.

The intensity estimates from the satellite pictures are relatively consistent among the more experienced analysts regardless of the satellite data used. This does not mean that all analysts came up with the same intensity estimate for the same picture; but their percentages of correct, scatter, and skill scores are comparable. These results are not significantly different to those obtained previously by Erickson (1972) for the better analysts in his study (unpublished) and only slightly larger than the results obtained by Gaby (1974) in his evaluation of the Miami/NESS SFSS analyses for 1974. The skill scores for the ESSA and DMSP products in general were the highest and probably represent a close approximation of what will be attained using the Dvorak technique and the GOES products with similar resolution. However, recall that the analyses for the weaker storms were better with the lower resolution data. Perhaps some minor modification of the system may be necessary for the weaker systems when using the higher resolution data.

The analyst's results were comparable in accuracy to a 24 hour dependent climatological forecast (see section 4.2 for an explanation of dependent climatology). That is, given yesterday's intensity, location, and direction of movement, we can predict what today's intensity should be based on the 24 hour climatological change. The skill scores for these climatological values are quite high. It is the authors' opinion that the intensity classification technique could be significantly improved by taking advantage of this information as a first-guess intensity for the analysis scheme. For instance, in a known area of potentially rapid development, a greater change from the Dvorak model CI would be permitted than for a statistically average development area etc.

The Dvorak intensity estimation scheme seems to be best for weak and average strength storms and worst for the very strong or extremely fast developing storms. Perhaps new research and additional experience with strong storms will result in better estimates for these cases in the future. The consistency of the results and the skill scores attained when compared with climatology indicate the usefulness of this product. To what degree these observations can or should replace or supplement aircraft reconnaissance information of course remains a decision for others. The data are presented so that the people making these decisions can determine the probability of a given error.

6. ACKNOWLEDGEMENTS

The authors appreciate the efforts of the participating analysts, as well as the National Environmental Satellite Services and the U.S. Air Force Air Weather Service for supplying the photographs and other material. Also, special thanks to Vern Dvorak and Don Gaby of NESS and Major Ernie Dash of AWS for coordinating activities with their organizations.

7. REFERENCES

Air Weather Service (1974): Air Weather Service Technique Development Summary, Dec. 73 - Feb. 74; Directorate of Aerospace Services DCS/Aerospace Sciences Headquarters Air Weather Service (MAC) U. S. Air Force. Scott AFB, Ill. 15 pp.

Dvorak, V. F. (1973): A Technique for the Analysis and Forecasting of Tropical Cyclone Intensities from Satellite Pictures, NOAA Tech. Memo./NESS 45, (U.S. Department of Commerce, National Hurricane Research Laboratory, Miami, Florida), 19 pp.

Erickson, C. O. (1972): A further Evaluation of the Dvorak Technique for Tropical Cyclone Classification, (NESS, Meteorological Satellite Laboratory, Washington, D. C. Unpublished Report).

Gaby, D. C. (1974): Atlantic Tropical Cyclone Classifications for 1974, (NESS Satellite Field Services Station, NOAA, Miami, Florida Unpublished Report).

Holiday, C. (1969): On the Maximum Sustained Winds Occurring in Atlantic Hurricanes, ESSA Tech. Memo./WMBTM-SR-45, (U.S. Department of Commerce, Weather Bureau Southern Region, Ft. Worth, Texas), 6 pp.

Kraft, R. (1961): The Hurricane's Central Pressure and Highest Winds, Mariners Weather Log, 5, No. 5, 157 pp.

Michaels, J. (1971): Intensity Changes of Maximum Winds in Tropical Cyclones in the Western North Pacific, (U. S. Department of Commerce, National Hurricane Research Laboratory Unpublished Report).

Michaels, J. (1973): Changes of the Maximum Winds in Atlantic Tropical Cyclones as Deduced from Central Pressure Changes, NOAA Tech. Memo./ERL WMPO-6, (U. S. Department of Commerce, National Hurricane Research Laboratory, Miami, Florida), 43 pp.

Sheets, R. C. (1972): Some Statistical Characteristics of the Hurricane Eye and Minimum Sea-Level Pressure, Project STORMFURY Annual Report, 1971, Appendix I (U. S. Department of Commerce, National Hurricane Research Laboratory, Miami, Florida), pp. 143-156.

Vernon, E. M. (1953): A New Concept of Skill Score for Rating Quantitative Forecasts, Monthly Weather Review, 81, No. 10, 326-329.

APPENDIX A

EVALUATION OF THE DVORAK FORECAST SCHEME

We mentioned in the Introduction that Dvorak has developed a forecast scheme based upon the storm's cloud structure as observed in the satellite photographs. This scheme is essentially an extension of his intensity analysis scheme discussed in the main text. All analysts participating in the evaluation study were also asked to complete the forecast scheme as devised by Dvorak. These forecasts were evaluated in the same manner as the 24 hour changes of current intensity (C.I.) described in section 4.2.

Our procedure was to determine the change in the current intensity estimated for the next 24 hour period. That is, if the day's analyzed CI number was 3.0 and the 24 hour forecast CI number was 4.0, then the change of CI was +1.0. This procedure removed constant biases of the analyst, as discussed in section 4.2, and enabled a direct comparison between a simple 24 hour forecast change based on climatology. The results of these analyses are shown in tables A-1, A-2, and A-3. The interpretation is the same as for tables 9 through 11 in the main text, except that the 24 hour forecast intensity change, rather than the analyzed 24 hour change, is being evaluated.

Table A-1 shows the results for the ATS sample. A comparison of the analysts' forecast values with values obtained by a direct application of the 24 hour climatological change shows an average improvement over the climatological-based forecast of 6 percent for the 0 deviation category (change forecast was change observed) for the top eight analysts. However, only three analysts showed improvements of more than 5 percent. The cumulative percentages for the deviation category of 1 ($\leq$ 0.5 CI number deviations from the observed change), showed an average improvement over climatology of only 1 percent for these same eight analysts. The major improvement appears in category 2, ($\leq$ 1 CI number deviation from the observed change).

The forecasts for the ESSA data (table A-2) show a higher percentage of correct calls (0 deviation) for the climatological-based forecasts than for any of the analysts. However, the analysts in general slightly exceeded the cumulative percentages associated with the climatological-based forecasts in the first deviation category ($\leq$.5 CI number deviations).

The analysts' forecasts for the DMSP data (table A-3) show slightly improved results over those obtained through use of straight climatology. These improvements are similar to those obtained for the ATS data sample with only small average percentage increases in each category.

The results of this analysis indicate that only small improvements are obtained by the analyst over simple climatologically-based forecasts. These average differences are generally less than 5 to 10 percent for any category or satellite. With two exceptions, the more skilled analysts' results were relatively consistent. This exception was the ATS sample where two analysts obtained significantly higher correct forecasts (0 deviation) than other analysts for the same data set or for their own analyses of the ESSA and DMSP data. This condition may have resulted from better "behaved" storms in the ATS sample and/or recall of specific cases which probably influenced some forecasts. More subjectivity enters the forecasts than the intensity analyses and, therefore, permits some conscious or subconscious bias from prior knowledge of a given storm. The fact that, in general, the percentages correct for the ATS sample were larger than for the ESSA and DMSP sample indicates that the ATS sample contained better "behaved" storms.

The skill scores generally indicated a higher skill for the analysts' forecasts than for climatological-based forecasts. This resulted primarily from a greater number of climatological-based forecasts in the higher deviation categories. This number is small, but the deviation skill scores penalize greatly for large misses. We do not mean to imply that the analysts were able to accurately forecast large changes. In fact, an examination of individual forecast changes for the Washington group revealed that seldom were large changes (2 CI numbers) forecast correctly within two categories. In summary, this forecast scheme appears best for well "behaved" systems and shows only minimal improvement over climatologically-based forecasts in its present form.

Table A-1. *Deviations of Forecast 24 Hour Change of Intensity From Official Change (Pressure) -- ATS-Atlantic**

Analyst	No. of Cases	Number of Category Deviations (C.I.)												Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6				
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Climatology	73	28.8	47.9	76.7	12.3	89.0	5.5	94.5	4.1	98.6	0	98.6	1.4	.221	.197	41.1	30.1
K	67	31.3	49.3	80.6	14.9	95.5	1.5	97.0	1.5	98.5	1.5	100	-	.295	.278	40.3	28.4
L	71	40.8	40.9	81.7	14.1	95.8	1.4	97.2	2.8	100	-	-	-	.363	.405	38.0	21.1
M	56	44.6	41.1	85.7	12.5	98.2	1.8	100	-	-	-	-	-	.462	.407	42.9	12.5
N	71	35.2	40.9	76.1	19.7	95.8	2.8	98.6	0	98.6	1.4	100	-	.330	.303	43.7	21.1
G	72	33.3	47.3	80.6	15.2	95.8	2.8	98.6	0	98.6	1.4	100	-	.340	.343	37.5	29.2
H	73	32.9	46.6	79.5	13.7	93.2	4.1	97.3	2.7	100	-	-	-	.345	.321	42.5	24.7
J	73	31.5	39.7	71.2	22.0	93.2	1.3	94.5	4.1	98.6	1.4	100	-	.303	.225	41.1	27.4
R	60	28.3	51.7	80.0	16.7	96.7	3.3	100	-	-	-	-	-	.239	.324	41.7	30.0
S	71	21.1	42.3	63.4	31.0	94.4	2.8	97.2	1.4	98.6	1.4	100	-	.243	.124	46.5	34.2
T	73	13.7	37.0	50.7	35.6	86.3	6.9	93.2	5.4	98.6	1.4	100	-	.176	-	42.5	43.8
U	63	30.2	36.5	66.7	25.4	92.1	6.3	98.4	0	98.4	0	98.4	1.6	.305	.221	36.5	33.3

* *The MSLP and table 3 are used to determine the cyclone intensity and the corresponding CI number. A category deviation of 1 represents a 24 hour forecast CI number change deviation of ± 0.5, 2 is 1.0 etc.*

Table A-2. *Deviations of Forecast 24 Hour Changes of Intensity From Official Change (Pressure) -- ESSA-Atlantic-Pacific**

Analyst	No. of Cases	Number of Category Deviations (C.I.)											Deviation Skill Score		Percent Analyzed	
		0	1		2		3		4		5		≥ 6			
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Climatology	82	29.3	28.0	57.3	23.2	80.5	9.7	90.2	6.1	96.3	0.0	96.3	3.7	.249	.308	41.5	29.3
K	82	22.0	41.4	63.4	24.4	87.8	6.1	93.9	3.7	97.6	2.4	100	-	.269	.345	40.2	37.8
L	83	25.3	37.4	62.7	24.0	86.7	9.7	96.4	1.2	97.6	1.2	98.8	1.2	.324	.358	31.3	43.4
M	87	25.3	37.9	63.2	26.5	89.7	4.6	94.3	4.6	98.9	1.1	100	-	.331	.359	34.5	40.2
N	90	18.9	38.9	57.8	23.3	81.1	6.7	87.8	6.6	94.4	3.4	97.8	2.2	.234	.272	38.9	42.2
O	89	18.0	40.4	58.4	29.2	87.6	5.7	93.3	4.5	97.8	1.1	98.9	1.1	.248	.313	41.6	40.4
P	88	21.6	39.8	61.4	23.8	85.2	5.7	90.9	6.8	97.7	2.3	100	-	.328	.351	38.6	39.8
Q	87	21.8	40.3	62.1	20.7	82.8	8.0	90.8	6.9	97.7	1.2	98.9	1.1	.279	.291	32.2	46.0
G	85	16.5	38.8	55.3	34.1	89.4	5.9	95.3	1.2	96.5	3.5	100	-	.280	.328	37.6	45.9
H	91	15.4	42.8	58.2	25.2	83.5	8.8	92.3	6.6	98.9	1.1	100	-	.286	.315	38.5	46.2
J	90	22.2	38.9	61.1	24.5	85.6	7.7	93.3	3.4	96.7	3.3	100	-	.301	.366	34.4	43.3
R	87	27.6	42.5	70.1	18.4	88.5	5.8	94.3	2.3	96.6	2.3	98.9	1.1	.322	.430	32.2	40.2
S	90	17.8	37.8	55.6	28.8	84.4	5.6	90.0	7.8	97.8	1.1	98.9	1.1	.255	.296	46.7	35.6
T	89	10.1	38.2	48.3	36.0	84.3	7.8	92.1	4.5	96.6	2.3	98.9	1.1	.220	.233	37.1	52.8
U	85	12.9	42.4	55.3	27.1	82.4	4.7	87.1	7.0	94.1	3.4	97.6	2.4	.174	.221	40.0	47.1

* *Same as for table A-1 except that the 24 hour forecast changes are for the ESSA satellite data sample.*

Table A-3. *Deviations of Forecast 24 Hour Changes of Intensity From Official Change (Pressure) -- DMSP-Pacific**

		Number of Cagegory Deviations (C.I.)												Deviation Skill Score		Percent Analyzed	
Analyst	No. of Cases	0	1		2		3		4		5		> 6				
		% in Cat	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	Cum. %	% in Cat	vs. Chance	vs. Clima	LT Obs	GT Obs
Climatology	101	22.8	42.5	65.3	19.8	85.1	5.0	90.1	5.9	96.0	1.0	97.0	3.0	.241	.244	34.7	42.6
K	86	29.1	40.7	69.8	20.9	90.7	4.6	95.3	2.4	97.7	0	97.7	2.3	.331	.400	34.9	36.0
L	94	23.4	42.6	66.0	24.4	90.4	4.3	94.7	3.2	97.9	0	97.9	2.1	.288	.341	41.5	35.1
M	87	25.3	44.8	70.1	19.6	89.7	6.9	96.6	2.3	98.9	0	98.9	1.1	.360	.402	36.8	37.9
N	100	21.0	39.0	60.0	22.0	82.0	13.0	95.0	2.0	97.0	1.0	98.0	2.0	.265	.255	43.0	36.0
O	101	28.7	43.6	72.3	21.8	94.1	2.9	97.0	0	97.0	1.0	98.0	2.0	.358	.420	39.6	31.7
P	100	27.0	41.0	68.0	21.0	89.0	8.0	97.0	1.0	98.0	0	98.0	2.0	.339	.377	42.0	31.0
G	101	34.7	35.6	70.3	19.8	90.1	4.0	94.1	4.9	99.0	1.0	100	--	.390	.430	28.7	36.6
H	101	25.7	51.5	77.2	14.9	92.1	2.9	95.0	2.0	97.0	1.0	98.0	2.0	.392	.407	29.7	44.6
J	102	31.4	37.2	68.6	18.7	87.3	8.8	96.1	2.9	99.0	0	99.0	1.0	.337	.392	35.3	33.3
R	100	30.9	43.3	74.2	16.5	90.7	6.2	96.9	2.1	99.0	0	99.0	1.0	.348	.431	24.7	44.3
S	102	21.6	37.2	58.8	27.5	86.3	10.8	97.1	0.9	98.0	0	98.0	2.0	.275	.270	45.1	33.3
T	101	19.8	35.6	55.4	29.7	85.1	6.0	91.1	3.9	95.0	2.0	97.0	3.0	.212	.188	37.6	42.6
U	102	19.6	38.2	57.8	28.5	86.3	6.8	93.1	4.0	97.1	1.9	99.0	1.0	.226	.251	42.2	38.2

* *Same as for table A-1 except that the 24 hour forecast changes are for the DMSP satellite data sample.*