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

OFFICE NOTE 261

A Threat Score Statistic for Hurricane Track Forecasts

Joseph P. Gerrity, Jr.
Development Division

JULY 1982

1. Introduction

I recently had occasion to read a report in which two sets of hurricane track forecasts were compared. The statistic used in the comparison was the distance between the observed and forecast position of the storm center. While it was easy to grasp the idea that a reduction in this statistic was indicative of an improved forecast, it was less obvious that the statistic is by itself a good measure of the utility of the forecast track of the storm.

Slow moving storms may be poorly forecast even though the displacement error is small compared to that obtained in a qualitatively better forecast of a fast-moving storm. This fact suggests that the displacement error should be normalized by the smaller distance traveled by either the forecast, or the observed storm center. Such a statistic would be analogous to the S_1 score used widely to evaluate numerical model circulation forecasts.

But since the hurricane threatens an area around the storm center, it occurred to me that another statistic, widely used in precipitation forecast verification, might find application in validating the utility of a numerical prediction of the track of a hurricane. The statistic I have in mind is the threat score which is simply defined as the ratio of the area correctly predicted to be threatened to the "union" of the area actually influenced and the area predicted to be influenced.

In order to be precise one must define the extent of the threatened area around the storm-track. At the present time, the numerical guidance produced by NMC does not indicate the extent of damaging winds and seas, thus for our purposes we have arbitrarily chosen to describe the threatened area to lie with a radius of one degree of latitude from the storm center.

To assess the virtues and faults of the threat score statistic, we have applied it to the 1978 Pacific storm FICO and to the NMC hurricane model forecasts made for that storm. In addition we have applied the statistic to the storms used in the comparison test mentioned in the first paragraph.

2. Method for Calculation

The storm tracks, observed and forecast, are plotted on an appropriate map. Since we are not concerned with achieving great precision, the variation of the map-scale in neglected, but we do lay out the radius of areal threat by using the 1° of latitude increment, as measured on the map at the latitude of the storm.

The envelope of threatened area is formed by drawing tangents to the circles which are centered on the forecast (or observed) storm center. A planimeter is used to measure the areas, enclosed within the appropriate curves, entering into the definition of the threat score statistic.

The statistic's measurement is subject to random error on the order of 10% which can be reduced by having the mechanical planimeter computation repeated independently a number of times.

By including the area around the initial position of the storm in the correctly predicted area, we have attributed some apparent skill to the forecast which actually results from the observation and initialization processes. Refinements to this procedure might be suggested.

3. Results for FICO

The twenty forecasts for storm FICO are summarized in the table below.

Initial Time	Forecast Period	Threat Score
12Z 17 Jul 78	48 hours	0.63
00Z 18 Jul 78	48 hours	0.28
12Z 18 Jul 78	30 hours	0.21
00Z 19 Jul 78	48 hours	0.63
12Z 19 Jul 78	48 hours	0.45
00Z 20 Jul 78	48 hours	0.41
12Z 20 Jul 78	48 hours	0.16
00Z 21 Jul 78	48 hours	0.23
12Z 21 Jul 78	48 hours	0.38
00Z 22 Jul 78	48 hours	0.25
12Z 22 Jul 78	48 hours	0.29
00Z 23 Jul 78	48 hours	0.62
12Z 23 Jul 78	48 hours	0.80
00Z 24 Jul 78	48 hours	0.43
12Z 24 Jul 78	48 hours	0.50
00Z 25 Jul 78	48 hours	0.23
12Z 25 Jul 78	48 hours	0.41
00Z 26 Jul 78	48 hours	0.26
12Z 26 Jul 78	48 hours	0.15
00Z 27 Jul 78	48 hours	0.19
Average		0.376

Table 1. Threat Scores for Typhoon FICO Track Forecasts.

The forecast and observed storm positions at six-hour intervals were used in defining the storm tracks and threatened areas. Two example maps used in computing Table 1 are given as Figures 1 and 2.

The average threat score 0.376 is broadly speaking consistent with a error of about 1° latitude in the displacement of the forecast track normal to the observed track.

In every forecast case the storm was predictd to the left of the observed track, usually too far southward of the westward moving storm. Notice that this is contrary to Rosenbloom's law for eastward moving extratropical cyclones.

4. Results for Spin-up Test Cases

For sixteen 24-hour forecasts of Atlantic hurricanes made during 1981 the utility of a new 3-dimensional spin-up initialization was evaluated by comparison with the original 2-dimensional spin-up forecasts. The results of this experimental evaluation were given in a memo for the record prepared by Donald Marks. Mr. Marks used the displacement error as the measure upon which to evaluate the improvement achieved by using the new initialization method.

With Mr. Marks assistance we were able to employ the basic data to calculate the two statistics--S_1, normalized displacement error and TS, threat score. The S_1 was calcualted using the observed storm displacement as the normalizing factor. The TS was computed using the method outlined earlier, but with only 12 hourly positions rather than the 6 hourly positions used in the FICO calculation.

The results are tabulated in Table 2, and in Figure 3 we have plotted the pair of values, TS,S_1 for each case.

The tabulated data and the graphed material both support Mr. Marks conclusion that the 3-dimensional spin-up produced a significant improvement in the average forecast in the studied sample.

It is of interest to note that the TS and S_1 statistics are well-correlated but the S_1 gives a more stretched-out measure than the TS. The threat score is compressed at its upper-end (good scores); the range of TS was 0.08 to 0.79, where as the S_1 ranged from 0.04 to 1.20.

5. Concluding Remarks

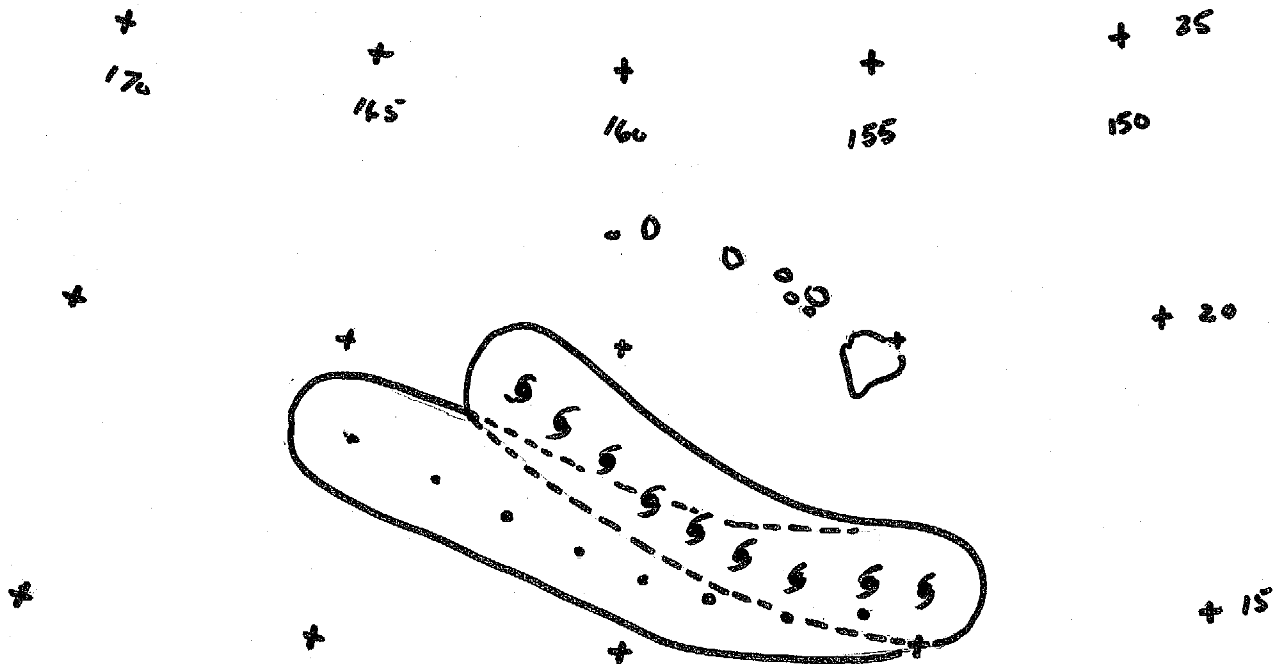
The verification of hurricane forecasts and the assessment of improvements made through research is a area of some concern; while it is possible to interpret most statistics in practical terms, it seems desirable to use statistics that rather directly measure those aspects of forecast performance which are of greatest practical concern. The threat score, presented here, seems to be a potentially useful tool for this purpose. It seems especially attractive to introduce a measure which will be adaptable to eventual use in assessing the skill with which advanced methods predict not only the track of the hurricane but its intensity as well. It is likely that a modification of the definition of the threat areas to fit the phenomena of concern (gale-force winds, etc.) would be especially useful.

6. Acknowledgment

The assistance of Mr. Marks in obtaining the data used is appreciated. Mr. Keith Brightwell computed the threat scores using a planimeter provided by Mr. David Olson of Forecast Division. Dr. John Hovermale's interest motivated the preparation of this note.

Name	Initial Time	Initial Position	24 Hour Observed Movement	Error 2-D DISP/S_1'/TS	Error 3-D DISP/S_1'/TS
Dennis	12Z 8/17	25.8, 81.2	84.7	61.5/.72/.17	24.1/.28/.57
Dennis	00Z 8/19	28.7, 80.8	300.1	154.4/.51/.38	95.5/.32/.55
Dennis	00Z 8/20	33.4, 78.8	467.1	308.2/.66/.28	309.3/.66/.38
Emily	12Z 9/3	35.0, 65.8	203.4	245.0/1.20/.08	156.5/.77/.27
Emily	00Z 9/4	34.6, 63.6	252.3	197.8/.78/.13	203.8/.81/.27
Emily	00Z 9/5	38.2, 60.9	135.1	47.9/.35/.33	7.6/.06/.75
Emily	12Z 9/5	39.0, 60.8	168.1	128.1/.76/.17	151.8/.90/.18
Floyd	12Z 9/6	24.5, 69.1	236.2	49.6/.21/.67	10.6/.04/.55
Floyd	00Z 9/7	26.4, 69.1	232.8	139.1/.60/.33	105.7/.45/.50
Floud	12Z 9/7	28.4, 68.5	234.8	52.2/.22/.60	11.9/.05/.73
Gert	12Z 9/9	20.3, 70.0	322.9	255.8/.79/.22	324.0/1.00/.21
Gert	12Z 9/10	23.7, 74.5	345.2	58.0/.16/.69	27.7/.08/.79
Gert	00Z 9/11	26.3, 73.9	385.1	62.6/.16/.53	42.0/.11/.56
Harvey	00Z 9/13	20.2, 57.8	330.5	108.6/.33/.46	77.1/.23/.38
Harvey	12Z 9/13	22.2, 60.6	276.9	150.9/.54/.33	123.0/.44/.42
Harvey	12Z 9/14	26.4, 62.7	275.6	137.7/.50/.25	114.5/.42/.31
Averages			255.26	134.84/.531/.351	111.6/.414/.464

Table 2. Scores From Comparison of 2D and 3D Spin-up Experimental Forecasts in 1981. Distances are in Nautical Miles. DISP is Displacement Error Between Forecast and Observed Storm Position. S_1' is the Ratio of DISP to Actual Storm Movement. TS is the Threat Score.



FICO 12Z 20 JUL 78

TS = 0.16

FIG 1. SAMPLE THREAT SCORE COMPUTATION

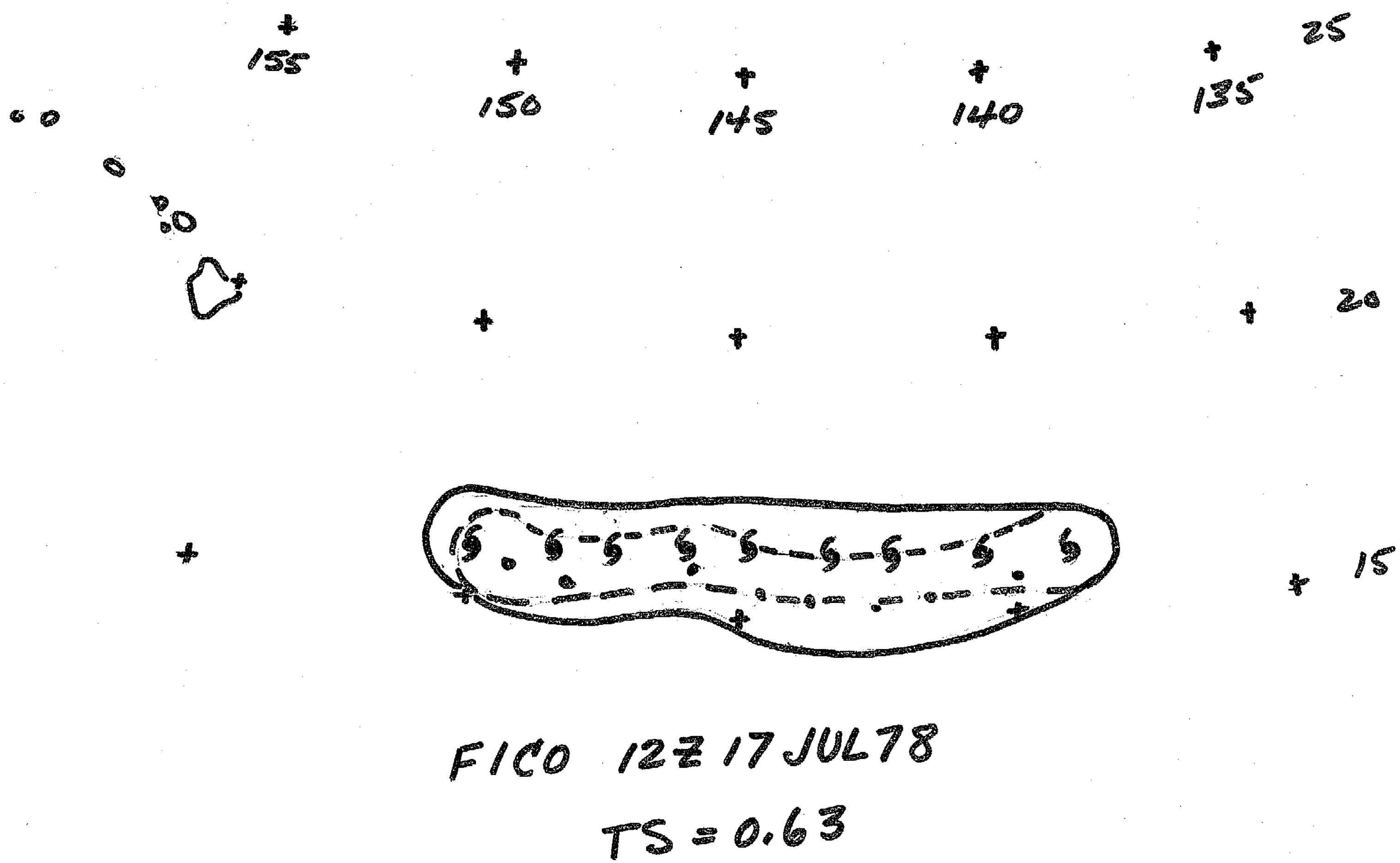


FIG 2. SAMPLE THREAT SCORE COMPUTATION

FIG. 3: THREAT SCORE TS VS NORMALIZED DISPLACEMENT ERROR S_1 FOR COMPARISON OF 2D AND 3D SPIN-UP HURRICANE 24 HR FORECASTS.