THE COMPARISON OF ERROR STATISTICS AMONG ANALOG TYPHOON TRACK PREDICTION MODELS

Shi-Yang, Chen[1] *& Davis K. Cope*[2]

ABSTRACT

A study of position prediction methods for typhoon tracks is presented. The basic HURRAN method, the modified single weighting and double weighting methods, and least squares curve fitting methods are examined and compared. Error statistics for all methods are compiled using western North Pacific typhoons for 1976 and 1978 as sample years. The data base used as a source of analog typhoons consisted of all western North pacific typhoons over 1959-1978 omitting the sample years 1976 and 1978.

The HURRAN method was stable over the two sample years with respect to both bias and accuracy and was the most accurate of the predictors. The single weighting method over the two sample years was stable with respect to bias and shows good potential as a prediction method. Suggestions for further investigation of weighted analog methods are presented.

These results will be useful for typhoon track prediction in the operational meteorological center.

1. INTRODUCTION

Typhoons are a disastrous weather phenomenon of the tropical area and often cause severe damage in property and loss of life. At present, weather modification techniques cannot deal with typhoons. Early and accurate prediction is necessary to reduce damage. Among the various predictions we can make about a typhoon, track prediction is the most important one. A track consists of positions of the center of the typhoon at 6 hour intervals and is obtained from the post analysis of original observations of the center of the typhoon by various instruments. Meteorological satellite data, available since 1965, provides smoother tracks than those of earlier years. The track data used in this research came from the Annual Typhoon Report of Joint Typhoon Warning Center (JTWC) in Guam.

Track prediction has been divided into four catagories (Hope and Neumann, 1977): the analog method, the empirical

1 Central Weather Bureau, Taipei, Taiwan, Republic of China
2 Mathematical Division, North Dakota State University, Fargo, ND 58105, U.S.A.

method, the regression equation method, and numerical methods. The empirical method refers to position prediction from a track plot using the subjective estimates of experienced forecasters. A more objective approach is desirable. The regression equation method uses a regression equation to relate predicted positions to a large number of input variables. The regression equation is obtained by the statistical analysis of past typhoon tracks, and it is difficult to assign physical meaning to the regression coefficients. Numerical methods of position prediction refer to the numerical solution of the hydrodynamic equations of motion for an air mass. The accuracy is often influenced by sparse observations in the West Pacific Ocean area and errors of radiosonde reporting. Computational time – even for simplified models – is often extreme. However, more sophisticated models are under development (Hu Chung-Ying and Chen Shi-yang, 1976).

Tha analog method uses the partial track of the current typhoon to select similar tracks from an historical file. It has the advantage of simple data input, fast calculation, and reasonably good results (Neuman and Hope, 1972; Jarrel and Wagoner, 1973). The method has been widely adopted (Hope and Neuman, 1977).

This research examines the accuracy of position prediction by the analog method. The historical file used consists of all typhoons in the western North Pacific area over 1959-1978, excluding 1976 and 1978. These years provide a sample of real typhoon tracks for testing the method and were chosen simply as two recent years with some separation between them to reduce any year-to-year correlations. Predictions are made from 12 to 72 hours ahead; the 24 hour predictions is generally the one of most interest to an operational center. We concentrate on analyzing predictions for 12, 24, and 36 hours.

Section 2 does position prediction by a simple least-squares curve fit to the typhoon track using polynomials from first to fourth order. This work provides baseline error data for comparison with the analog method.

Section 3 gives a brief background review of the analog method.

Section 4 adapts the basic analog method of Neuman and Hope (1972) to the Western Pacific area data base and determines its prediction error.

Section 5 and 6 modify the basic analog method by weighting the analog tracks selected from the historical file according to their similarity to the current track. The more heavily weighted tracks make a heavier contribution to the position prediction. Prediction errors are determined for two weighting schemes.

Section 7 summarizes the results.

2. LEAST SQUARES PREDICTORS AND ERROR STATISTICS

a. Description of Data

The 19 tracks of 1976 and 19 tracks of 1978 which occured in the West Pacific Ocean over latitude 14 to 28 degree North and west of longitude 140 degree East were taken as a random sample of typhoons for testing track prediction methods.

b. Least Squares Fit

Position predictions were obtained as follows: given a current position on a current track, the lat/long (y-coordinate/x-coordinate) positions of the most recent several points, given at 6 hour intervals, were considered as functions of time. Least-squares fits for x, y as polynomials in the time were obtained. Thse polynomials were then used to predict positions 12, 24, and 36 hours ahead of the current point. For each prediction the error (Dx, Dy)= $(X_{true}, Y_{true}) - (X_{pred}, Y_{pred})$. Polynomial fits from first order through fourth order were examined, and the number of most recent points used varied from 3 to 13 as indicated in Tables 1-6.

c. Error Statistics of Curve Fitting

Tables 1, 2, 3, 4, 5, 6 summarize the results for 12, 24, and 36 hour predictions of 1976 and 1978, respectively. The sections of each table summarize the error statistics for the combinations of order of polynomial fit and number of points used in the fit. The statistics in each section are based on 163 to 242 total points for 1976 and 190 to 293 total points for 1978; the variation depends on the number of points used for the curve fit – requiring 13 points leaves fewer points for making predictions than requiring 3 points.

The statistics presented in each section of the table are the mean Dx, Dy – errors (used to check for biased predictions) $\overline{Dx}$, $\overline{Dy}$, the correlation coefficient (reported for completeness and not directly relevant to these error evaluations), and the semi-axes A, B for the 50% probability ellipse. $\overline{Dy}$, $\overline{Dx}$, A, B are given in degrees. The $\overline{Dy}$, $\overline{Dx}$ values give the bias of the predictions and the A, B values give the accuracy. The correlation coefficient is used to calculate the inclination of the ellipse in an actual prediction. Figure 1 shows a 50% probability ellipse calculated for a set of error values and the application of the ellipse to making a poistion prediction. The formulas used to calculate these quantities are given in Appendix. A.

Using these tables, first notice the difference in the size of the probability ellipse between corresponding entries for 1976 and 1978. The 1978 ellipses are considerably alrger than the corresponding 1976 ellipses. This difference is our first basic result. The absolute performance of thse predictors cannot be obtained from data for a single year because there may be large variations in performance for a given predictor from year to year, that is, the least squares predictors are unstable. This research does not attempt estimates of absolute performance. Instead the relative performance of estimators for each of the years 1976 and 1978 is examined, and this relative performance is consistent for the two years.

Second, increasing the order of the polynomial fit does not improve the accuracy. Although the results for third and fourth order polynomials show a tendency to improve as the number of data points used in the fit increases, they are still definitely worse than the first and second order predictors. Consequently, we concentrate on the first and second order cases.

Third, a polynomial predictor may require more data points as one tries longer prediction intervals. The first order predictor works best at 3 data points over all prediction intervals. The second order predictor uses 5-7 points for a best prediction over 12 hours, 9 points for a best prediction over 24 hours, and apparently something over 9 points for a best prediction at 36 hours.

Fourth, the simple first order predictor using 3 data points gives the most ac-

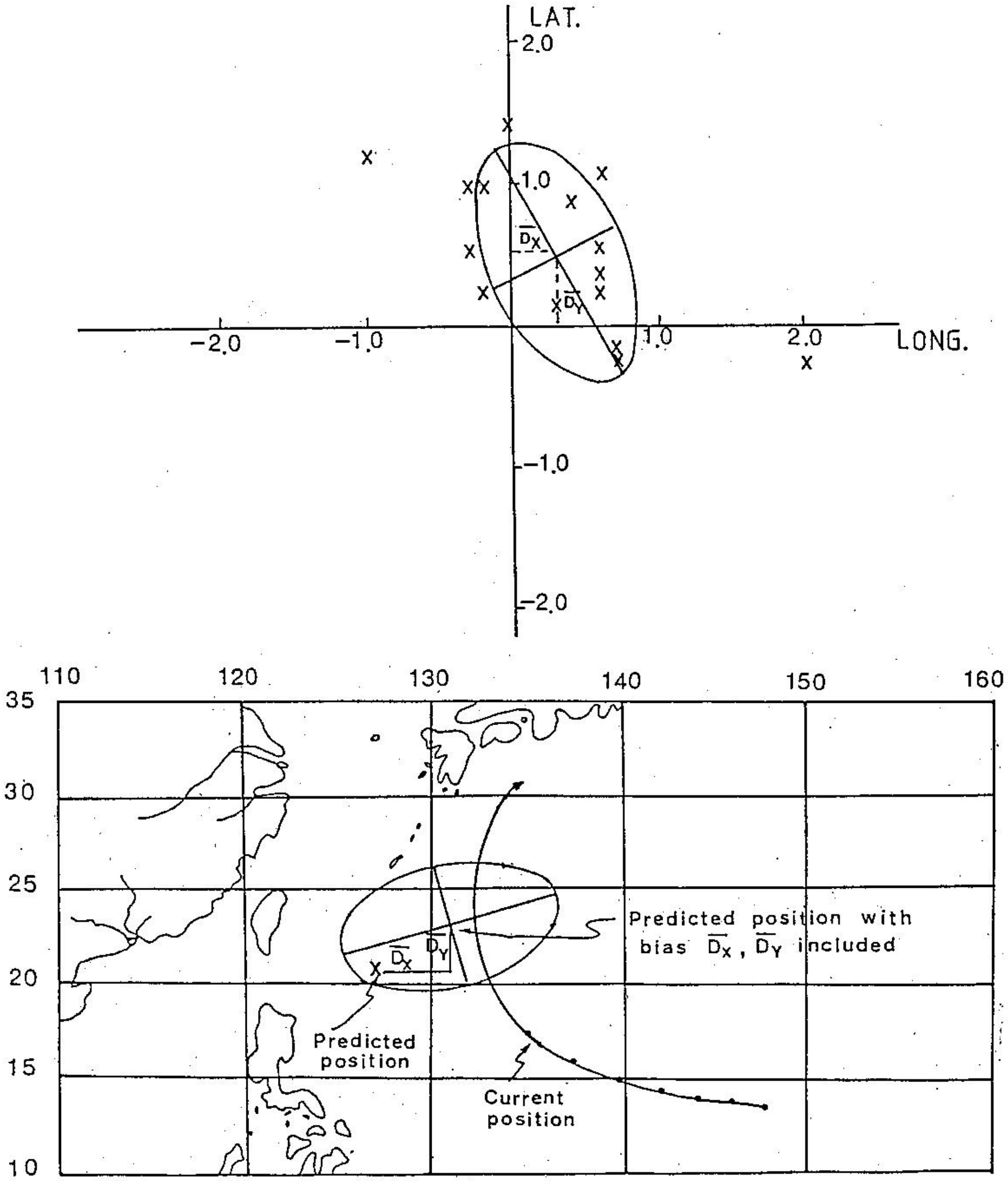

Figure 1: A 50% Probability Ellipse Showing Bias $\overline{Dx}$, $\overline{Dy}$ and Semiaxes A, B and the Application of the Ellipse To Obtaining A Position Prediction

curate results (smallest 50% probability ellipse) over all prediction intervals but it has a clearcut and sizeable longitudinal bias ($\overline{Dx}$-values). This bias must be subtracted from the predicted position for accurate results; values for the bias (average of 1976 and 1978 results) are 0.22 degree for 12 hour, 0.73 degree for 24 hour, 1.52 degree for 36 hour. This bias in the linear predictor is consistent with experience. Typhoons in the Western Pacific tend to move to the west and then curve to the north. This curving tends to place the longitude of the true position to the east of the linearly extrapolated position, in agreement with the bias observed here. Notice that the bias of the second order predictor is small compared with the first order. We could say that in choosing the first order over the second order we are choosing a gain in accuracy at the cost of a greater bias.

Finally, these results suggest that a further study of least-squares curve fits should involve first and second order predictors using 3-7 points. Such a study could compile error statistics over several years of data to confirm the instability in accuracy and bias. Such a study should also consider modified predictors which might combine the accuracy of the first order with the lack of bias of the second order, for instance, taking a linear fit for the latitude and a quadratic for the longitude.

3. PRINCIPLES OF THE ANALOG METHOD

The analog method first scans the historical file of past typhoons and chooses those most similar to the current typhoon. After making suitable track adjustments, the movements of the historical typhoons are used to predict the future movement of the current typhoon.

The method is very similar to the subjective forecasting procedure, because a forecaster makes predictions by an analog procedure (Jarrell and Wagoner, 1973). He finds similar conditions, then mentally determines the average results under these conditions, then makes his best estimate.

Two analog typhoon track prediction models, HURRAN and TYFOON, were developed in the 70s and are still used routinely. HURRAN (HURRicane ANalog), developed by Hope and Neuman, is based on a statistical study done by Haggard et al. in 1965 of the probability a hurricane will invade the Cape Kennedy space craft launching site and is a by-product of the American space project. After that, a similar model, TYFOON, was developed by the U.S. Navy and used by the Joint Typhoon Warning Center (JTWC), Guam in August, 1970 (Jarrell and Somervell, 1970). Now TYFOON has been improved to TYFN 78, and a similar method has been applied to track predictions of tropical cyclones in the Indian and South Pacific Ocean.

In general, the variations between analog models fall into the three areas: 1) the criteria used to select analog typhoons; 2) the adjustment of differences between the current typhoon and the analog typhoon; and 3) the composition of the adjusted analog tracks to a single position prediction. (Jarrell, Mauck and Renard, 1975).

4. THE HURRAN ANALOG METHOD AND ERROR STATISTICS

a. The HURRAN Technique

The HURRAN technique was developed

Table1. Error statistics for 12 hour prediction by curve fitting (1976 data).

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	0.029	0.046	0.073	0.107	0.148	0.206
	$\overline{Dx}$	0.271	0.506	0.792	1.050	1.308	1.590
	A	0.507	0.577	0.687	0.849	1.034	1.225
	B	0.847	1.045	1.248	1.567	1.938	2.345
	Corr. coef.	0.120	0.142	0.176	0.186	0.209	0.215
	Total point	242	225	209	194	179	163
order 2	$\overline{Dy}$	0.010	0.038	0.034	0.005	-0.030	-0.007
	$\overline{Dx}$	0.023	0.090	0.160	0.238	0.315	0.416
	A	0.823	0.663	0.637	0.684	0.765	0.864
	B	0.997	0.967	1.020	1.074	1.171	1.265
	Corr. coef.	0.035	0.162	0.102	0.010	-0.112	-0.185
	Total point	242	225	209	194	179	163
order 3	$\overline{Dy}$	—	0.023	0.030	0.019	0.014	0.011
	$\overline{Dx}$	—	0.011	-0.002	-0.057	-0.043	-0.002
	A	—	1.490	1.008	0.871	0.858	0.901
	B	—	1.592	1.420	1.348	1.324	1.388
	Corr. coef.	—	0.018	0.042	0.043	0.027	-0.075
	Total point	—	225	209	194	179	163
order 4	$\overline{Dy}$	—	-0.081	0.003	-0.003	0.041	0.122
	$\overline{Dx}$	—	-0.168	-0.065	-0.096	-0.065	-0.040
	A	—	4.279	2.115	1.526	1.255	1.144
	B	—	4.654	2.529	1.996	1.855	1.766
	Corr. coef.	—	0.001	0.126	0.024	-0.003	0.000
	Total point	—	225	209	194	179	163

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)

A, B = semi minor, semi major axes (degree) for 50% probability ellipse

Total point = number of predicted points

Order refers to the order of the polynomial fit (with respect to time)

Table 2. Error statistics for 12 hour prediction by curve fitting (1978 data).

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	-0.010	-0.013	0.019	0.068	0.130	0.163
	$\overline{Dx}$	0.179	0.283	0.410	0.545	0.714	0.951
	A	0.760	0.892	1.038	1.142	1.238	1.374
	B	1.019	1.203	1.433	1.770	2.140	2.523
	Corr. coef.	0.028	0.030	0.003	-0.027	-0.071	-0.132
	Total point	317	299	279	257	235	211
order 2	$\overline{Dy}$	-0.010	0.004	0.012	0.004	-0.001	-0.016
	$\overline{Dx}$	0.020	0.068	0.095	0.136	0.184	0.273
	A	1.446	1.091	1.046	1.077	1.178	1.218
	B	1.932	1.433	1.357	1.419	1.523	1.581
	Corr. coef.	-0.063	0.021	0.000	-0.099	-0.112	-0.075
	Total point	317	299	279	257	235	211
order 3	$\overline{Dy}$	—	-0.012	0.028	0.002	-0.039	-0.082
	$\overline{Dx}$	—	0.001	-0.017	0.004	0.060	0.080
	A	—	2.367	1.672	1.454	1.342	1.299
	B	—	3.018	2.259	1.890	1.759	1.791
	Corr. coef.	—	-0.082	0.013	0.006	-0.078	-0.076
	Total point	—	299	279	257	235	211
order 4	$\overline{Dy}$	—	-0.056	-0.055	-0.000	0.027	0.025
	$\overline{Dx}$	—	-0.113	-0.033	-0.000	0.015	0.042
	A	—	7.139	3.722	2.524	2.058	1.874
	B	—	9.398	4.792	3.592	2.846	2.320
	Corr. coef.	—	-0.134	-0.047	-0.006	-0.037	-0.058
	Total point	—	299	279	257	235	211

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)

A, B = semi minor, semi major axes (degree) for 50% probability ellipse

Total point = number of predicted points

Order refers to the order of the polynomial fit (with respect to time)

Table 3. Error statistics for 24 hour prediction by curve fitting (1976 data).

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	0.032	0.070	0.109	0.132	0.181	0.269
	$\overline{Dx}$	0.926	1.366	1.794	2.172	2.530	2.920
	A	1.032	1.075	1.233	1.447	1.679	1.903
	B	1.882	2.055	2.380	2.855	3.395	3.935
	Corr. coef.	0.084	0.170	0.183	0.215	0.225	0.219
	Total point	220	205	191	176	161	147
order 2	$\overline{Dy}$	-0.046	0.023	0.058	-0.008	-0.062	-0.001
	$\overline{Dx}$	0.285	0.405	0.519	0.668	0.800	1.013
	A	2.280	1.714	1.514	1.495	1.573	1.633
	B	2.967	2.621	2.388	2.312	2.357	2.484
	Corr. coef.	0.113	0.125	0.034	-0.098	-0.199	-0.299
	Total point	220	205	191	176	161	147
order 3	$\overline{Dy}$	—	-0.049	-0.066	-0.026	0.026	0.010
	$\overline{Dx}$	—	0.182	-0.050	-0.174	-0.105	-0.093
	A	—	5.875	3.387	2.616	2.328	2.404
	B	—	6.254	5.052	4.246	3.780	3.633
	Corr. coef.	—	0.023	0.043	0.022	-0.068	-0.119
	Total point	—	205	191	176	161	147
order 4	$\overline{Dy}$	—	-0.023	-0.147	-0.307	0.014	0.223
	$\overline{Dx}$	—	-0.693	-0.100	-0.466	-0.167	-0.069
	A	—	25.923	10.531	6.518	4.863	3.978
	B	—	29.272	12.577	8.822	7.490	6.462
	Corr. coef.	—	0.016	0.153	0.021	0.003	-0.039
	Total point	—	205	191	176	161	147

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)

A, B = semi minor, semi major axes (degree) for 50% probability ellipse

Total point = number of predicted points

Order refers to the order of the polynomial fit (with respect to time)

Table 4. Error statistics for 24 hour prediction by curve fitting (1978 data)

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	-0.022	-0.006	0.067	0.150	0.224	0.260
	$\overline{Dx}$	0.586	0.769	0.988	1.228	1.503	1.904
	A	1.608	1.761	1.852	1.888	1.984	2.106
	B	2.144	2.378	2.690	3.098	3.559	3.970
	Corr. coef.	0.000	0.011	-0.001	-0.023	-0.096	-0.157
	Total point	293	275	255	233	211	190
order 2	$\overline{Dy}$	0.070	0.033	0.041	0.025	0.002	-0.045
	$\overline{Dx}$	0.162	0.253	0.321	0.432	0.579	0.814
	A	4.210	2.793	2.484	2.468	2.503	2.430
	B	5.302	3.622	3.185	3.183	3.230	3.249
	Corr. coef.	-0.119	-0.008	-0.051	-0.110	-0.103	-0.035
	Total point	293	275	255	233	211	190
order 3	$\overline{Dy}$	—	0.217	0.111	-0.037	-0.173	-0.257
	$\overline{Dx}$	—	0.052	-0.073	0.030	0.148	0.218
	A	—	9.200	5.671	4.372	3.673	3.334
	B	—	11.321	7.338	5.522	4.926	4.646
	Corr. coef.	—	-0.113	-0.006	-0.024	-0.102	-0.077
	Total point	—	275	255	233	211	190
order 4	$\overline{Dy}$	—	0.065	0.123	0.160	0.071	0.172
	$\overline{Dx}$	—	-0.392	-0.018	0.072	0.121	0.101
	A	—	43.042	18.203	10.824	7.966	6.493
	B	—	55.912	23.011	14.837	10.563	8.110
	Corr. coef.	—	-0.185	-0.084	-0.030	-0.064	-0.101
	Total point	—	275	255	233	211	190

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse
Total point = number of predicted points
Order refers to the order of the polynomial fit (with respect to time)

Table 5. Error statistics for 36 hour precition by curve fitting (1976 data).

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	0.103	0.123	0.125	0.118	0.164	0.249
	$\overline{Dx}$	1.938	2.504	3.004	3.465	3.864	4.303
	A	1.519	1.647	1.866	2.137	2.413	2.655
	B	2.942	3.171	3.635	4.228	4.865	5.525
	Corr. coef.	0.231	0.237	0.251	0.255	0.235	0.206
	Total point	193	182	170	155	142	128
order 2	$\overline{Dy}$	0.065	0.255	0.246	0.112	0.033	0.105
	$\overline{Dx}$	0.535	0.825	1.065	1.264	1.502	1.795
	A	4.579	3.129	2.635	2.511	2.638	2.740
	B	5.855	4.789	4.241	3.938	3.955	4.030
	Corr. coef.	0.138	0.135	-0.008	-0.146	-0.238	-0.286
	Total point	193	182	170	155	142	128
order 3	$\overline{Dy}$	—	-0.012	0.050	0.195	0.325	0.342
	$\overline{Dx}$	—	-0.062	-0.165	-0.464	-0.462	-0.282
	A	—	15.241	7.790	5.518	4.792	4.820
	B	—	16.060	12.177	9.454	7.836	7.271
	Corr. coef.	—	0.030	0.023	-0.032	-0.099	-0.177
	Total point	—	182	170	155	142	128
order 4	$\overline{Dy}$	—	-0.700	-0.648	-0.613	0.155	0.674
	$\overline{Dx}$	—	-2.668	-1.289	-1.268	-0.491	-0.696
	A	—	90.087	33.241	18.494	12.411	9.710
	B	—	103.803	39.313	26.008	20.462	16.187
	Corr. coef.	—	-0.012	0.145	0.006	-0.018	-0.101
	Total point	—	182	170	155	142	128

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)

A, B = semi minor, semi major axes (degree) for 50% probability ellipse

Total point = number of predicted points

Order refers to the order of the polynomial fit (with respect to time)

Table 6. Error statistics for 36 hour prediction by curve fitting (1978 data).

		Number of points in regression					
		3	5	7	9	11	13
order 1	$\overline{Dy}$	-0.054	-0.010	0.099	0.181	0.259	0.322
	$\overline{Dx}$	1.216	1.489	1.832	2.218	2.671	3.151
	A	2.614	2.660	2.650	2.659	2.718	2.845
	B	3.486	3.806	4.196	4.662	5.118	5.483
	Corr. coef.	-0.038	-0.034	-0.039	-0.061	-0.114	-0.143
	Total point	267	249	229	207	188	171
order 2	$\overline{Dy}$	-0.000	0.098	0.116	0.077	-0.028	-0.045
	$\overline{Dx}$	0.383	0.584	0.631	0.901	1.228	1.735
	A	8.676	5.415	4.590	4.263	4.053	3.920
	B	10.899	6.761	5.601	5.354	5.296	5.262
	Corr. coef.	-0.124	-0.031	-0.080	-0.147	-0.141	-0.138
	Total point	267	249	229	207	188	171
order 3	$\overline{Dy}$	—	0.244	0.116	-0.156	-0.474	-0.541
	$\overline{Dx}$	—	0.064	-0.104	-0.120	0.102	-0.010
	A	—	23.455	13.298	9.464	7.558	6.629
	B	—	29.022	16.971	12.009	9.860	8.425
	Corr. coef.	—	-0.123	-0.027	-0.050	-0.093	-0.049
	Total point	—	249	229	207	188	171
order 4	$\overline{Dy}$	—	-2.490	0.135	-0.137	0.292	-0.013
	$\overline{Dx}$	—	-2.092	0.135	0.687	0.278	0.161
	A	—	143.794	56.240	30.037	20.085	15.782
	B	—	194.676	70.942	41.309	27.314	19.431
	Corr. coef.	—	-0.222	-0.089	-0.040	-0.102	-0.103
	Total point	—	249	229	207	188	171

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse
Total point = number of predicted points
Order refers to the order of the polynomial fit (with respect to time)

by Hope and Neumann (1970) for the prediction of the movement of North Atlantic tropical cyclones.

1) The Criteria for Selection of Analogs

The selected analog must satisfy four criteria:

i) Distance:

We find the position on the historical track which is closest to the current typhoon position. Let this point be (X_k, Y_k). The distance between these two positions must be less than 300 nautical miles.

ii) Speed:

The vector velocity is based on the vector difference between the present and preceding position (6 hours). One degree per hour equals 60 knots. Speed and heading are calculated from the vector velocity.

We calculate the speed of the historical and current typhoon at (X_k, Y_k). Let D_v be the speed difference between current and historical typhoons and v be the speed of the current typhoon. Then the analog must satisfy:

$D_v < 5$ knots if $V < 10$
$D_v < 10$ knots if $10 <= V < = 20$
$D_v < 15$ knots if $V > 20$

iii) Heading:

The historical track must have a heading within 22.5 degree of either side of the heading of the existing storm.

iv) Date:

The historical storm must be within 30 days before or after the current date.

2) The Adjustment of Analog Tracks

The HURRAN technique uses a simple method of adjustment. It takes the point which is closest to the current position to be the "best position" and translates (without rotation) the historical track until the best position coincides with the current typhoon position, as shown in Figure 2.

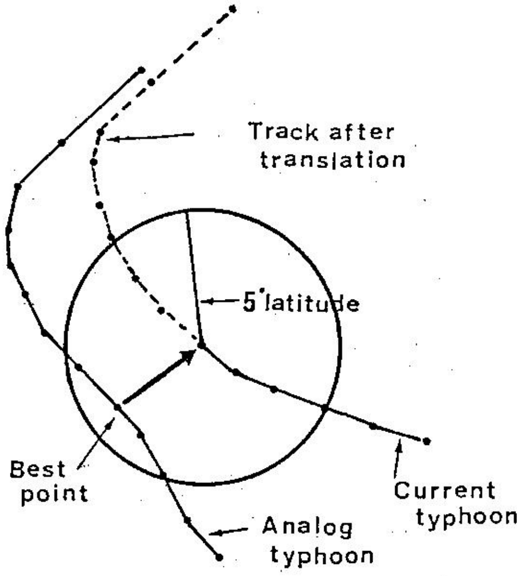

Figure 2: Track Translation of the Analog Typhoon

3) Composition of Tracks

Each historical track is composed with the current track to produce a position prediction. The arithmetic mean of these position predictions is the final predicted position for the current typhoon. The composition proceeds as follows: The initial position of the composite track is the current position, and new positions (at 6 hour intervals) on the composite track are formed by adding position increments. Each poistion increment is a weighted average of the last 6 hour position change of the current typhoon with the position change at the appropriate future time of the analog track. The contribution of the current typhoon is referred to as "persistence"; the analog track contribution is called the "climatology"

Hope and Neumann (1970) prove that if persistence is included over the first 36 hours, average forecasting results will be improved. Including persistence after 36 hours will have no effect on the forecast. Therefore, the initial weighting is almost entirely "persistence" but the weighting shifts towards the climatological component until it is entirely climatological at 36 hours and beyond.

Here (X_j, Y_j), (X'_j, Y'_j), (X''_j, Y''_j) are the (longitude, latitude) coordinates of the current track, historical track, and composite track respectively at 6 hour intervals, and j=k corresponds to the current position.

For $0 < i < 6$,

$$X''_{k+i} = X''_{k+i-1} + (X_k - X_{k-1})(1-i/6) + (X'_{k+i} - X'_{k+i-1})(i/6)$$

$$Y''_{k+i} = Y''_{k+i-1} + (Y_k - Y_{k-1})(1-i/6) + (Y'_{k+i} - Y'_{k+i-1})(i/6)$$

For $6 <= i$,

$$X''_{k+i} = X''_{k+i-1} + X'_{k+i} - X'_{k+i-1}$$

$$Y''_{k+i} = Y''_{k+i-1} + Y'_{k+i} - Y'_{k+i-1}$$

The predicted typhoon position $(XMEAN_{k+i}, YMEAN_{k+i})$ at 6i hours will be:

$$XMEAN_{k+i} = \frac{1}{N} \sum_{\text{all composites}} X''_{k+i}$$

$$YMEAN_{k+i} = \frac{1}{N} \sum_{\text{all composites}} Y''_{k+i}$$

Here, N = number of analog typhoons.

b. Error Statistics for the HURRAN Analog Method

The HURRAN method was applied to the same random sample of typhoons in 1976 and 1978. Besides the error statistics, we also calculate the vector error, right angle error, and direction error of forecasting.

Vector error is the difference in distance between the true future location and the predicted location. The right angle error is the smallest distance from the predicted location to the correct track of the typhoon. Vector and right angle errors are given in kilometers. The direction error is the difference (degrees) in direction between the vectors from the curent position to the predicted and true position. The direction error is negative when the predicted position is to the left of the true position and positive when the predicted position is on the right. The average direction error in Tables 9, 10 is based on absolute values.

From Tables 7 and 8, we first note that the differences in bias and accuracy between 1976 and 1978 is definitely less than the difference for the least squares predictors in section 2. The HURRAN predictor is more stable from year to year than the least squares predictors.

Second, the HURRAN accuracy is consistently better than the least squares predictors of section. 2. HURRAN has a clearcut bias in longitude at all prediction intervals; a latitudinal bias begins to show clearly around 36 hours. The bias values (based on averaging the 1976 and 1978 values) are $(\overline{Dx}, \overline{Dy})$ = (−0.04, 0.20), (−0.16, 0.59), (−0.49, 0.78) for 12, 24, and 36 hours respectively.

Further work on the HURRAN predictor would include verification of its stability and determination of the bias over a large sample of years. It would also be desirable to investigate the source of the bias — there does not seem to be any direct natural explanation as there was for

Table 7. Error statistics for 12 to 72 hour prediction by HURRAN (1976 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	-0.033	-0.155	-0.506	-0.353	-0.596	-0.124
$\overline{Dx}$	0.248	0.730	0.859	1.354	2.356	3.424
A	0.503	1.064	1.840	1.994	2.715	2.819
B	0.613	1.548	2.906	4.227	5.973	7.588
Corr. coef.	0.0000	0.223	0.371	0.359	0.405	0.366
Total point	122	104	88	72	58	39

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% proabability ellipse
Total point = number of predicted points

Table 8. Error statistics for 12 to 72 hour prediction by HURRAN method (1978 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	-0.043	-0.170	-0.486	-0.886	-1.281	-1.996
$\overline{Dx}$	0.164	0.499	0.729	1.099	1.486	1.018
A	0.593	1.167	1.797	2.413	2.845	3.565
B	1.004	1.784	2.800	4.113	5.582	7.047
Corr. coef.	0.148	0.124	0.155	0.209	0.393	0.327
Total point	188	166	141	120	103	84

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% proabiliity ellipse
Total point = number of predicted points

Table 9. Mean forecast erros for 12 to 72 hour by HURRAN method (1976 data).

	Hours of prediction					
	12	24	36	48	60	72
Vector	62.587	155.470	272.853	378.242	539.392	645.563
Right angle	42.946	108.727	188.369	268.444	394.583	433.802
Direction	14.970	18.931	22.308	25.902	30.216	30.215
Negative angles	78	64	52	45	39	31
Zero angles	1	1	0	0	0	0
Positive Angles	44	40	38	27	19	11
Total point	123	105	90	72	58	42

Vector error = mean radial error.
Right angle error = mean distance from precited position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

Table 10. Mean forecast errors for 12 to 72 hour by HURRAN method (1978 data)

	Hours of prediction					
	12	24	36	48	60	72
Vector	81.852	170.774	271.601	391.001	494.004	625.434
Right angle	51.166	111.255	174.803	245.347	284.582	316.460
Direction	16.505	18.007	19.058	22.657	25.710	30.908
Negative angles	91	80	58	48	38	25
Zero angles	3	1	0	2	0	0
Positive Angles	95	86	85	72	67	61
Total point	189	167	143	122	105	86

Vector error = mean radial error.
Right angle error = mean distance from precited position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

the linear predictor bias in section 2.

From Tables 9 and 10, the prediction errors of 1978 are similar to those of 1976. For the 1976 data, predictions tend to deviate to the left of the track; in 1978 the deviation is to the right of the track at longer times. It seems that the method has no systematic direction error in prediction.

5. THE WEIGHTED ANALOG METHOD

a. Introduction

This section and the next studies an analog method for the prediction of typhoon tracks. (Shi-Yang, Chen, 1980, 1983) This method scans the historical file for typhoons analogous to the current one using a specified space-time envelope criterion (section 5.2), then makes use of 28 weighting factors to compute a similarity index for each analog typhoon (sections 5.3 and 5.4). The similarity index is first computed using different positions on the historical track as "current positions" and the best similarity index then determines the portion of the analog track most similar to the current typhoon. These best similarity indices are then used to rank the analogs in order of their similarity to the current typhoon. After adjustments for both initial position and velocity difference have been made for each analog typhoon (section 5.5), their adjusted tracks are weighted according to their rankings and then combined to produce the predicted positions of the current typhoon for every 12 hours up to 72 hours (section 5.6). Two different ways of combining position predictions are examined and these are referred to as the single weighting and double weighting methods (not to be confused with the weights used to calculate the similarity index).

b. Analog Selection Criteria

A pure analog scheme is one which searches the historical file for a single track similar to the current partial track. The subsequent behavior of this analog track is then used directly as a forecast. A pure analog scheme is undesirable for two reasons. First, good analogs are not common enough to insure that one can be found for most forecast situations. Second, the results obtained by using a pure analog scheme are inferior to those obtained by subjective estimation or a weighted average of several similar typhoons (Jarrell and Somervell, 1970; Jarrell, Mauck and Renard, 1975). This poor performance apparently results from the fact that two tracks may be very similar over a portion of the path yet diverge widely in the future.

Therefore, we look for criteria to find a group of acceptable analogs. Generally, analog schemes screen track parameters to separate analogs into two groups: "good enough" and "not good enough" The "not good enough" group is then ignored and the "good enough" group forms the basis of the forecast. As the criteria for "good enough" are relaxed, the number of analogs increases, but by including less similar tracks. Obviously, it is ideal to define the "good enough" cutoff point as that point where the improvement brought about by increasing the number of analogs exactly balances the detrimental effect of including worse analogs and beyond which the net result is to decrease accuracy.

Typhoon tracks in the Western Pacific area show both regional and seasonal variation. Typhoons tend to originate in the southeastern low latitudes, travel approximately northwest for some period, then dissipate or continue curving towards the northeast. The tracks show a seasonal migration also. Tracks tend to lie in the eastern area in the winter and spring and tend to lie to west in the summer and fall. This suggests selecting historical typhoons which occurred in the same area and at about the same time of year as the current typhoon. In fact, such selection criteria were used in TYFOON 72, 73 (Jarrell and Wagoner, 1973). After trying various limits, a time-space envelope of ±10 days, ± 2.5 degrees latitude, ± 5.0 degree longitude was chosen. That is, an historical track is chosen as an analog if 1) the initial (terminal) date of its track lies within + 10 days (– 10 days) of the date of the current position on the current typhoon track; and 2) some point on the historical track is within ±2.5 degree latitude, ±5.0 degree longitude of the current position on the current track.

If it should happen that three or fewer analogs occur, the time-space envelop in the program is enlarged, first increasing the time limits to ±30 days, then the longitude limits to ±7 degrees, then the latitude limits to ±3.5 degrees, until more than three analogs are found. The reason fo the adjustment in this order is that empirical observations while choosing the time-space envelope indicated least sensitivity to date and most sensitivity to latitude. An insufficient number of analogs may occur when the current typhoon has a strange route or if it happens during a month when typhoons are unusual. Storage in the program limits the selection to 80 analogs.

c. Weighting of Analogs

The degree of similarity of the historic to the current typhoons is determined by the analog parameters, a set of physically meaningful values calculated from the historical track for comparison with the corresponding values of the current track, and their weights, used to combine the analog parameters into a single similarity index.

In 1979 the Japan Meteorological Agency (JMA) developed a PC (persistency and climatology) typhoon track prediction model. This model includes 48 physically meaningful track parameters which are related to the persistency and climatology of typhoon movement. These track parameters are the input variables for several regression equations for predicting the track and central pressure; different equations correspond to different spatial areas and time periods during the typhoon season. The analog parameters for the weighted analog method studied here were selected from the 48 PC track parameters. The JTWC data file provides only time and space data on hitorical tracks, which is not sufficient data to calculate all 48 PC parameters. Instead, 28 parameters which can be calculated from time and space data alone are used. These parameters are basically velocities, accelerations, angular velocities, etc. at various positions along the track and are not independent parameters. These 28 analog parameters are listed in detail in Appendix B.

The weights assigned to the 28 parameters are based on a subjective estimate of the importance of the parameter in the PC model regression equations. Since there are different equations for different spatial areas and time periods during the typhoon season, there is a similar division of weightings. Appendix B contains a detailed description of the weights. This reasearch examines only this one selection of weights.

d. The Similarity Index and Ranking of Analogs

We shall describe the process of ranking analogs in more detail. Let c(1) to c(28) be the 28 track parameters calculated for the current position on the current track.

Our first step is to determine the position in the analog track most similar to the current position of the current track. For each psoition on the hostorical track we calculate a set of track parameters p(1) to p(28). For each index i a value RKVE (i) is computed: The absolute values of c(i)-p(i) are arranged in order from smallest to largest and RKVE(i) is the location of the difference within this ordering. The minimum value of difference is 1, then the second minimum is 2, etc. If a position does not have certain parameter, the total number of positions within the space-time envelope is assigned to it. There are 28 weights w(i) for the 28 parameters and these are used to calculate the similarity index SI:

$$SI = \sum_{i=1}^{28} W(i)*RKVE(i)$$

Each position on the analog track now has a similarity index associated with it, and the position with mininum similarity index is the best position.

Finally, we determine the degree of similarity between historical and current typhoons by comparing the values of the similarity indices of the best positions for the analog typhoons. The analog with the minimum similarity index is the most similar typhoon, rank number 1 is assigned to it, then 2 is assigned to the second, etc.

e. Adjustment of Analog Track for Position Prediction

After similar typhoons are chosen by the analog selection, both "best position" and "rank number" are decided. The track translation is then carried out. The method of simple translation of the analog track to the current position is used by HURRAN and has been illustrated in Figure 2. A more advanced method considers the possibility that if there is a significant difference in direction and velocity of movement between current and similar typhoon at the beginning, then the difference will increase more and more with time. Therefore, an additional adjustment of the translated track should make a better forecast available (Jarrell and Somervell, 1970). We follow the same approach, that is, the vector displacements of the current and analog typhoons over the preceding 12 hours are obtained, and the difference is the adjustment vector. Then 2, 4, 6 times this particular vector will be added to the positions of the translated track at 24, 48, 72 hours from the present position as shown in Figure 3. These positions form the adjusted analog track.

f. Weighted Composition of Predictions for Final Prediction

After all the tracks of analog typhoons are translated and adjusted, all the positions after the "best position" of each analog typhoon will be obtained. Composition of the positions for a given future time then gives the predicted track position.

Methods of composition include arithmetic averaging, weighted averaging and persistence modified averaging, and a combination of the last two referred to as doble weighting averaging.

1) Arithmetic Average Method

An arithmetic mean value of the corresponding adjusted analog positions is used as the predicted position.

2) Weighted Average Method (Ranking)

This method uses the ranking of the analogs to weight the more similar typhoons more heavily.

$$LAT_p = \frac{2}{N(N+1)} \sum_{i=1}^{N} (N-i+1)\,(\text{Latitude})_i$$

$$LON_p = \frac{2}{N(N+1)} \sum_{i=1}^{N} (N-i+1)\,(\text{Longitude})_i$$

Here, p is the prediction interval.

N is the total number of analog typhoons.

i is the rank number of analog typhoon (i=1 for most similar to i=N for least similar).

$(\text{Latitude})_i$ is latitude of ith analog typhoon at p hours.

$(\text{Longitude})_i$ is longitude of ith analog typhoon at p hours.

LAT_p is forecasting latitude at p hours.

LON_p is forecasting longitude at p hours.

3) Persistence Modified Arithmetic Average Method

This method is used by HURRAN. As discussed in section 4, the analog tracks are adjusted by terms referring to the persistence of movement of the current typhoon (section 4.1.3) and then a direct arithmetic average is taken. This process is equivalent to taking a direct arithmetic average of the analog positions and adding the persistence correction.

4) Double Weighting Average Method

This method combines both the rank

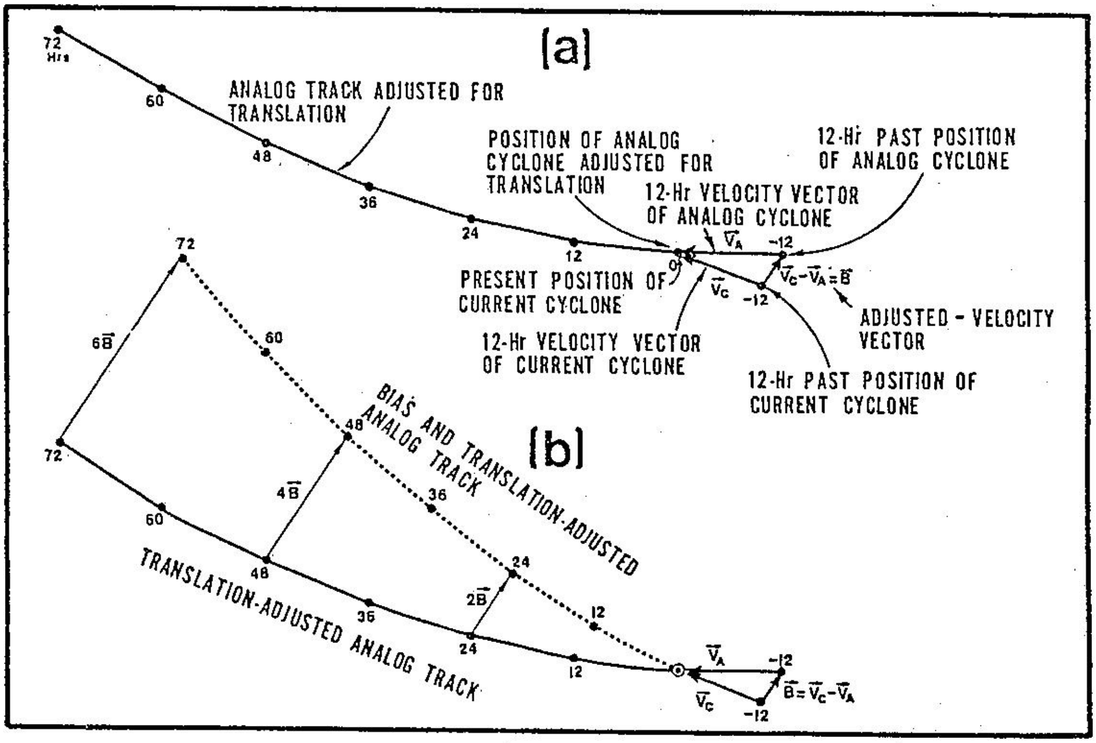

Figure 3. Velocity-difference vector (bias) adjustment: (a) Obtaining the "adjustment-velocity vector" B; and (b) Obtaining the biasadjusted position for different times. (After Jarrell and Wagoner, 1973)

weighting of tracks and the inclusion of persistence. The adjusted analog positions are first adjusted further by a persistency correction, as in section 4.1.3, and these corrected positions are then combined by the weighted average method above. This combination is loosely referred to as "double weighting".

6. ERROR STATISTICS FOR THE WEIGHTED ANALOG METHOD

The weighted analog method was applied to the sample years 1976 and 1978 with the same error statistics calculated as for the HURRAN and least squares methods. Both the single and double weighting variations were examined. Results are summarized in Table 11-18. As in the HURRAN case, the data base used to obtain analogs is the file of Western Pacific typhoons over 1959-1978 with 1976 and 1978 omitted.

We first consider the single weighting analog method with error statistics and forecast errors in Tables 11-14.

The single weighting predictor shows large variations in accuracy (50% probability ellipse) between 1976 and 1978 and is therefore unstable. The variations in accuracy are about the same as the best least-squares predictor (linear 3-point).

The single weighting predictor also shows accuracies that are approximately the same as the best least squares predictor and not as good as HURRAN.

The bias of the single weighting pre-

Table 11. Error statistics for 12 to 72 hour prediction by Single Weighting method (1976 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	-0.241	-0.562	-0.795	-1.508	-2.154	-2.357
$\overline{Dx}$	-0.085	-0.109	-0.272	-0.493	0.084	1.057
A	0.625	1.206	2.124	2.999	4.304	5.220
B	0.809	1.717	2.757	4.674	6.448	8.218
Corr. coef.	0.087	0.063	0.077	0.270	0.109	0.103
Total point	189	170	151	131	113	86

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse
Total point = number of predicted points

Table 12. Error statistics for 12 to 72 hour prediction by Single Weighting method (1978 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	-0.201	0.418	-0.776	-1.092	-1.286	-1.903
$\overline{Dx}$	-0.100	0.081	-0.197	-0.028	0.599	1.158
A	0.711	1.477	2.537	3.644	4.896	6.282
B	1.111	2.205	3.474	5.109	7.229	9.442
Corr. coef.	-0.048	-0.071	-0.081	-0.160	-0.156	-0.085
Total point	257	233	206	176	159	136

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse
Total point = number of predicted points

Table 13. Mean forecast errors for 12 to 72 hour prediction by Single Weighting method (1976 data)

	Hours of prediction					
	12	24	36	48	60	72
Vector	76.148	166.551	280.900	450.428	653.948	819.732
Right angle	50.054	112.579	173.792	271.172	437.585	565.864
Direction	24.161	26.050	32.821	37.426	43.306	47.981
Negative angles	68	65	60	55	53	44
Zero angles	0	1	0	0	0	0
Positive angles	122	106	92	77	60	43
Total point	190	172	152	132	113	87

Vector error = mean radial error.
Right angle error = mean distance from predicted position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

Table 14. Mean forecast errors for 12 to 72 hour prediction by Single Weighting method (1978 data)

	Hours of prediction					
	12	24	36	48	60	72
Vector	96.368	198.913	337.117	489.159	676.432	879.128
Right angle	59.009	121.073	198.641	296.349	420.637	567.963
Direction	23.505	26.731	33.129	37.187	42.479	46.262
Negative angles	89	92	68	66	69	56
Zero angles	4	4	0	0	0	0
Positive angles	164	137	139	111	90	80
Total point	257	233	207	177	159	136

Vector error = mean radial error.
Right angle error = mean distance from predicted position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

Table 15. Error statistics for 12 to 72 hour prediction by Double Weighting method (1976 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	0.591	0.767	0.560	-0.061	-0.675	-0.883
$\overline{Dx}$	0.028	0.148	-0.044	-0.504	-0.055	0.774
A	0.743	1.181	1.530	2.326	3.581	4.236
B	0.983	1.759	2.773	4.521	6.134	7.942
Corr. coef.	0.110	0.241	0.443	0.482	0.302	0.298
Total point	189	170	151	131	113	86

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse
Total point = number of predicted points

Table 16. Error statistics for 12 to 72 hour prediction by Double Weighting method (1978 data).

	Hours of prediction					
	12	24	36	48	60	72
$\overline{Dy}$	0.382	0.486	0.198	-0.093	-0.359	-0.970
$\overline{Dx}$	-0.352	-0.468	-0.589	-0.489	0.087	0.583
A	0.796	1.364	2.123	3.107	4.296	5.572
B	1.173	2.234	3.251	4.702	6.754	8.945
Corr. coef.	0.035	0.045	-0.037	-0.151	-0.139	-0.076
Total point	257	233	206	176	159	136

$\overline{Dy}$, $\overline{Dx}$ = mean errors (degree)
A, B = semi minor, semi major axes (degree) for 50% probability ellipse.
Total point = number of predicted points

Table 17. Mean forecast errors for 12 to 72 hour prediction by Double Weighting method (1976 data)

	Hours of prediction					
	12	24	36	48	60	72
Vector	112.425	184.345	257.774	403.232	568.755	720.990
Right angle	36.121	74.523	124.580	191.800	329.619	440.499
Direction	28.828	28.224	37.001	38.652	43.317	48.574
Negative angles	95	81	73	55	50	42
Zero angles	1	0	0	0	0	0
Positive angles	94	91	79	77	63	45
Total point	190	172	152	132	113	87

Vector error = mean radial error.
Right angle error = mean distance from predicted position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

Table 18. Mean forecast errors for 12 to 72 hour prediction by Double Weighting method (1978 data)

	Hours of prediction					
	12	24	36	48	60	72
Vector	118.044	206.152	290.906	419.566	599.447	789.165
Right angle	31.101	68.244	133.512	225.778	337.353	474.531
Direction	24.781	26.323	33.232	37.023	42.411	45.908
Negative angles	111	108	70	72	74	56
Zero angles	7	2	1	0	0	0
Positive angles	139	123	136	105	85	80
Total point	257	233	207	177	159	136

Vector error = mean radial error.
Right angle error = mean distance from predicted position to closest point on true path.
Direction error = mean of absolute value of heading difference to true and predicted positions.
Negative/Zero/Positive angles = number of angles with indicated sign.

dictor is very consistent over two years, and it may be that this predictor is stable with respect to bias. Examination of other years would be necessary to confirm this point. Notice that the bias is the reverse of that observed for HURRAN and the linear 3-point predictors: here the major bias is in the latitude.

The double weighting predictor also shows large variations in accuracy between 1976 and 1978 and is unstable.

The accuracy of the double weighting predictor cannot conclusively be compared with the single weighting. Double weighting is less accurate for 12 hours, about the same accuracy at 24 hours, and more accurate at 36 hours. In fact, at 36 hours the accuracy is better than the linear 3-point predictor, but overall double weighting is less accurate than

HURRAN.

The bias shows quite erratic behavior between 1976 and 1978, and double weighting therefore shows instability in both accuracy and bias.

7. CONCLUSION AND SUGGESTIONS

This research studies four methods for typhoon track prediction: least squares curve fitting, the HURRAN analog method, and single and double weighting analog methods. Error statistics are compiled using Western Pacific typhoons of 1976 and 1978. The data base used as a source of analog typhoons consisted of all Western Pacific typhoons over the period 1959-1978, omitting the sample years 1976 and 1978.

The three major ways in which predictors are compared are:

1) Stability – error statistics which show only mild fluctuation from year to year.

2) Accuracy – the size of the 50% probability ellipse for the errors in position prediction (semiaxes are represented by A and B, the area is π AB).

3) Bias – The mean values ($\overline{Dy}$, $\overline{Dx}$= latitude, longitude) of the error, in position prediction. The error is the vector difference $(X_{true}, Y_{true}) - (X_{pred}, Y_{pred})$.

The error statistics from the previous chapters are summarize in Table 19 and 20. The following conclusions were reached:

1) The simplest least squares predictor (linear 3-point) was consistently the most accurate of all the least squares predictors examined.

2) With respect to accuracy, HURRAN shows fairly stable behavior over 1976 and 1978, the 24 hour and 36 hour accuracies are very close although the 12 hour accuracies show a large difference. The other three predictors show large variations between the two years with the linear 3-point showing the worst variation: the areas of the 50% ellipse change by a factor of 2.

3) With respect to bias, HURRAN and the single weighting method show stable behavior. The linear 3-point shows large variation in longitudinal bias; and double weighting shows large variations in both latitudinal and longitudinal bias.

4) HURRAN is the most accurate predictor. The linear 3-point and single weighting show equal accuracies. Double weighting is somewhat ambiguous: it is less accurate than single weighting at 12 hours and more accurate at 36 hours.

5) Finally, HURRAN and the single weighting method are the predictors most appropriate for actual use. Their stable behavior with respect to bias indicates that specific values for the bias applicable to all years can be obtained, and these values subtracted from the position predictions to yield relatively unbiased predictions. The HURRAN predictions will be more accurate than single weighting in its present form.

The basic idea of the single and double weighting methods is to improve the HURRAN analog method by ranking analogs by their similarity to the current typhoon track and than giving more weight to the more similar analogs in forming a final position prediction. This concept is reasonable but did not work in this particular study. We now consider possibility for further work on this concept by comparing the HURRAN and weighted analog procedures.

Table 19. Summary of Accuracy Statistics

		Least Squares*		HURRAN		SW		DW	
		76	78	76	78	76	78	76	78
12 hrs	A	0.507	0.760	0.503	0.593	0.625	0.711	0.743	0.796
	B	0.847	1.019	0.613	1.004	0.809	1.111	0.983	1.173
	πAB	1.349	2.433	0.969	1.870	1.588	2.482	2.295	2.933
24 hrs	A	1.032	1.608	1.064	1.167	1.206	1.477	1.181	1.364
	B	1.882	2.144	1.548	1.784	1.717	2.205	1.759	2.234
	πAB	6.102	10.831	5.174	6.541	6.505	10.231	6.526	9.573
36 hrs	A	1.519	2.614	1.840	1.797	2.124	2.537	1.530	2.123
	B	2.942	3.486	2.906	2.800	2.757	3.474	2.773	3.251
	πAB	14.039	28.627	16.798	15.807	18.397	27.689	13.329	21.683

*LINEAR 3-POINT.
A, B = semi minor, semi major axes (degree) for 50% probability ellipse.
πAB = area of 50% ellipse.

Table 20. Summary of Bias

		Least Squares*		HURRAN		SW		DW	
		76	78	76	78	76	78	76	78
12 hrs	$\overline{Dy}$	0.029	-0.010	-0.033	-0.043	-0.241	-0.201	0.591	0.382
	$\overline{Dx}$	0.271	0.179	0.248	0.164	-0.085	-0.100	0.028	-0.352
24 hrs	$\overline{Dy}$	0.032	-0.022	-0.155	-0.170	-0.562	-0.418	0.767	0.486
	$\overline{Dx}$	0.926	0.586	0.730	0.499	-0.109	-0.081	0.148	-0.468
36 hrs	$\overline{Dy}$	0.103	-0.054	-0.506	-0.486	-0.795	-0.776	0.560	0.198
	$\overline{Dx}$	1.938	1.216	0.859	0.729	-0.272	-0.197	-0.044	-0.589

*Linear 3-point.
$\overline{Dy}$, $\overline{Dx}$ - mean errors (degree)

HURRAN	Weighted Analog
Step 1. Choose analog typhoons.	Step 1. Choose analog typhoons.
Step 2. On each analog find the point closest to the current position.	Step 2a. On each analog find the point at which the analog track is most similar to the current track.
	Step 2b. Rank the resulting analog tracks from most similar to least similar.
Step 3. Adjust the analog track to the current position. Simple translation plus the "persistence/climatology" adjustment is used.	Step 3. Adjust the analog track to the current position. Simple translation and an adjustment based on the difference in velocities of the current and analog tracks are used.
Step 4. Take appropriate future positions on each adjusted analog and obtain a predicted portion by averaging.	Step 4. Single weighting method. Take appropriate future positions on each adjusted analog and obtain a predicted position by a weighted average. A linear weighting from least to most similar is used. Double weighting method. Apply the "persistence/climatology adjustment to each adjusted track, then proceed as in the single weighting method.

The first point to be considered is the choice of analogs in Step 1. The selection criteria used in the weighted analog method was broader than HURRAN and permitted a larger number of analogs to be chosen. This may have led to cases where no analogs were selected by HURRAN and no position prediction was made, while the weighted analog method did select analogs and make a prediction but the analogs were very poor. This possibility is supported by the fact that the HURRAN statistics (Table 7 and 8) show smaller sample sizes than the single and double weighting methods (Tables 11, 12, 15, 16). Thus, further investigation should examine the effects of narrowing the selection criteria for the weighted analog approach.

The weighted analog method differs from HURRAN in the following respects:

1) The choice of best point on the analog track (based on the similarity index of different points calculated from weighting factors for 28 track parameters).

2) the ranking of analogs from most similar to least similar (also based on the similarity index).

3) the weighted average used in obtaining the final position prediction.

4) the inclusion of current track velocity in adjusting the analog tracks (note this is done twice in the double weighting method).

Consequently, our basic suggestion for further work on the weighted analog method is the systematic investigation of the effect of each of these areas on the predictor accuracy. Specifically:

1) restrict analog track adjustment to simple translation and the "persistence/climatology" adjustment as in HURRAN. Track adjustment is independent of the concept of a weighted average of ranked position estimates, and such adjustments should be omitted until the weighted analog concept is studied.

2) investigating different ways of choosing the best point and ranking the analogs. For exqmple, the simplest change would be to retain HURRAN's choice of the best point as the analog point closest to the current position and then rank the analogs according the similarity of the shapes of the analog track immediately preceding the best point to the current track.

3) investigate different weighting distributions for track composition in addition to the linear distribution used here, for example, geometric weighting distribution.

Some further points on the weighted analog approach are:

1) It has a simple and clearcut phsyical basis and is objective in nature.

2) The input data of the model are very simple and computation of the results is very fast (2 minutes execution in the minicomputer).

3) The double weighting method has smaller mean vector and right angle errors at and after 36 hours than the single weighting method has. Within 36 hours, the former reduces the right angle errors of the latter. The reduced percentages are 27.8-47.3%, 33.8-43.6%, 28.3-32.8% in 12, 24, and 36 hours respectively. Also, the double weighting method reduces the systematic right deviation of prediction of the single weighting method.

4) In general, the double weighting method gives the smallest right angle error over all the predicted intervals, that is, the method is characterized by "direction superior to speed". This is very helpful to a forecaster because the direction of a typhoon is more difficult to predict than the speed.

The following recommendations would apply to all analog methods:

1) The historical file only contains typhoon data over the past 18 years in the Western North Pacific Ocean, so there is insufficient information on irregular typhoons. If it were possible to increase the historical file to 30 years, it would increase the ability to forecast the irregular typhoon. Again, if it were possible to file the typhoons together with the historical weather patterns, then the analog models could further improve the forecasting of typhoon movement near ridge and trough patterns.

2) If it is possible to divide the historical data files into westward and curved typhoons, it could reinforce the accuracy of the mdoel forecasting by taking into account the moving characteristics of the typhoon.

Appendix A
STATISTICAL FORMULAS

The error statistics for the prediction errors were calculated using the following formulas. Let (x_i, y_i) be a set of data points. The standard deviations S_x, S_y and correlation coefficient R_{xy} are:

$$S_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \overline{Y})^2}$$

$$S_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \overline{X})^2}$$

$$R_{xy} = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N} (X_i - \bar{X})^2 \sum_{i=1}^{N} (Y_i - \bar{Y})^2}}$$

If the points (x_i, y_i) belong to a bivariate normal distribution, then the associated probability ellipses have semiaxes determined by the roots of the determinantal equation:

$$\begin{vmatrix} S_x^2 - K^2 & R_{xy} S_x S_y \\ R_{xy} S_x S_y & S_y^2 - K^2 \end{vmatrix} = 0$$

Let K_a and K_b be the larger and smaller positive roots respectively. Then the major axis (A_p) and minor axis (B_p) of the ellipse with probability p are:

$$A_p = K_a(-2 \ln (1-p))^{1/2}$$

$$B_p = K_b(-2 \ln (1-p))^{1/2}$$

Specifically, the semiaxes A, B for the 50% ellipse are:

$$A = A_{50\%} = 1.1774\, K_a$$

$$B = B_{50\%} = 1.1774\, K_b$$

The angle between major axis and latitude is:

$$Phi = \frac{\pi}{2} + Tan^{-1} \left(\frac{2R_{xy}S_xS_y}{S_x^2 - S_y^2} \right)$$

Here, $\frac{\pi}{2} < Phi <= \frac{\pi}{2}$

Note: program uses precise inverse tangent ATAN2(X1, X2), which gives the exact inverse $Tan^{-1}(X1/X2)$ with angle $-\pi <$ phi $<= \pi$.

Appendix B
THE WEIGHTING PARAMETERS AND THEIR WEIGHTS

B.1 Definition of Parameters

The 28 parameters of the weighted analog method are:

1) Displacement in latitude of past 12 hours:

$C(1) = LAT_{00} - LAT_{-12}$

2) Displacement in latitude of past 24 hours:

$C(2) = LAT_{00} - LAT_{-24}$

3) Displacement in latitude of past 36 hours:

$C(3) = LAT_{00} - LAT_{-36}$

4) Displacement in latitude of past 48 hours:

$C(4) = LAT_{00} - LAT_{-48}$

5) Displacement in latitude of past 48 to 24 hours:

$C(5) = LAT_{-24} - LAT_{-48}$

6) Acceleration in latitude of past 24 hours:

$C(6) = LAT_{00} + LAT_{-24} - 2LAT_{-12}$

7) Acceleration in latitude of past 48 hours:

$C(7) = LAT_{00} + LAT_{-48} - 2LAT_{-24}$

8) Displacement in longitude of past 12 hours:

$C(8) = LON_{00} - LON_{-12}$

9) Displacement in longitude of past 24 hours:

$C(9) = LON_{00} - LON_{-24}$

10) Displacement in longitude of past 36 hours:

$C(10) = LON_{00} - LON_{-36}$

11) Displacement in longitude of past 48 hours:

$$C(11)= LON_{00}-LON_{-48}$$

12) Displacement in longitude of past 48 to 24 hours:

$$C(12)= LON_{-24}-LON_{-48}$$

13) Acceleration in longitude of past 24 hours:

$$C(13)= LON_{00}+LON_{-24}-2LON_{-12}$$

14) Acceleration in longitude of past 48 hours:

$$C(14)= LON_{00}+LON_{-48}-2LON_{-24}$$

15) Direction of movement in past 12 hours:

$$C(15)= ARC\ TAN(C(1)/C(8))$$

16) Direction of movement in past 24 hours:

$$C(16)= ARC\ TAN(C(2)/C(9))$$

17) Direction of movement in past 36 hours:

$$C(17)= ARC\ TAN(C(3)/C(10))$$

18) Direction of movement in past 48 hours:

$$C(18)= ARC\ TAN(C(4)/C(11))$$

19) Direction of movement in past 48 to 24 hours:

$$C(19)= ARC\ TAN(C(5)/C(12))$$

20) Direction of acceleration in past 24 hours:

$$C(20)= ARC\ TAN(C(6)/C(13))$$

21) Direction of acceleration in past 48 hours:

$$C(21)= ARC\ TAN(C(7)/C(14))$$

22) Speed of movement in past 12 hours:

$$C(22)= C(1)^2+C(8)^2)^{1/2}$$

23) Speed of movement in past 24 hours:

$$C(23)= (C(2)^2+C(9)^2)^{1/2}$$

24) Speed of movement in past 36 hours:

$$C(24)= (C(3)^2+C(10)^2)^{1/2}$$

25) Speed of movement in past 48 hours:

$$C(25)= (C(4)^2+C(11)^2)^{1/2}$$

26) Speed of movement in past 48 to 24 hours:

$$C(26)= (C(5)^2+C(12)^2)^{1/2}$$

27) Magnitude of acceleration in past 24 hours:

$$C(27)= (C(6)^2+C(13)^2)^{1/2}$$

28) Magnitude of acceleration in past 48 hours:

$$C(28)= (C(7)^2+C(14)^2)^{1/2}$$

Here LON_{00}, LAT_{00} are the current longitude and latitude of typhoon center; while LON_{-12}, LAT_{-12}, LON_{-24}, LAT_{-24}, LON_{-36}, LAT_{-36}, LON_{-48}, LAT_{-48} are the longitude and latitude of typhoon center in past 12, 24, 36 and 48 hours respectively.

B.2 The Weighting of Analog Parameters

Figure 4 shows the different areas used in determining weighting of the analog parameters. The following list gives the weight W(i) for analog parameter i, i=1 to 28. Weights equal to one are not listed. In each area the weights also vary with the season.

A) North area:

current typhoon locate between 20 and 35 degree latitude north, 120 and 150

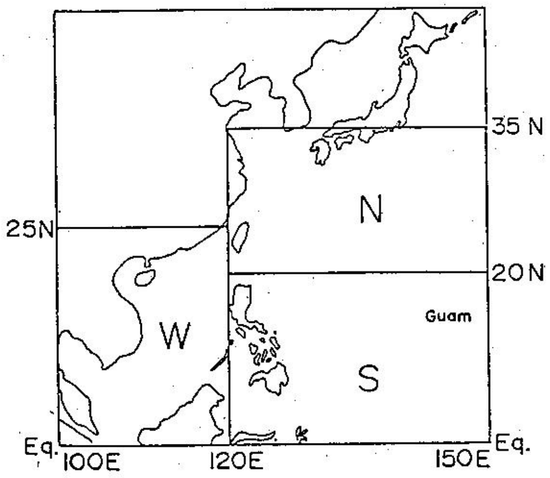

Figure 4: Areas Used in Determining Weighting (After Aoki, T., 1979)

degree longitude east.

1) January to June:

W(1)=W(10)=W(16)=W(17)=
=W(18)=W(21)=W(24)=W(25)
=W(26)=W(28)=2
W(13)=W(15)=W(19)=W(27)=3
W(2)=W(3)=W(8)=W(22)=4

2) July:

W(15)=W(20)=W(27)=W(28)=2
W(2)=W(4)=W(6)=W(13)=W(16)=3
W(1)=W(8)=4

3) August:

W(2)=W(3)=W(4)=W(5)=W(9)
=W(12)=W(15)=2
W(1)=W(8)=W(13)=4

4) September:

W(2)=W(5)=W(10)=W(20)=W(22)
=W(28)=2
W(13)=3
W(1)=W(8)=4

5) October to December:

W(12)=W(14)=W(16)=W(17)
=W(19)=2
W(2)=W(7)=W(13)=W(22)=W(23)=
=W(25)=W(27)=W(28)=3
W(1)=W(8)=4

B) South area:

current typhoon locate between 0 to 20 degree latitude north, and 120 to 150 degree longitude east.

1) January to August:

W(3)=W(8)=W(13)=W(16)=W(17)
=W(18)=W(23)=W(25)=W(27)
=W(28)=2
W(9)=W(20)=3
W(1)=4
W(11)=5

2) September to December:

W(10)=W(19)=2
W(4)=W(22)=W(24)=3
W(1)=4
W(8)=5

C) West area:

current typhoon locate at west of 120 degree longitude east.

1) January to August:

W(3)=W(4)=W(5)=W(11)=W(16)
=W(18)=W(24)=W(25)=2
W(14)=W(17)=W(22)=W(27)=3
W(1)=W(8)=W(16)=4

2) September to December:

W(7)=W(13)=W(16)=W(20)=W(21)
=W(24)=W(25)=W(27)=W(28)
=2
W(1)=W(2)=4
W(8)=5

References

Annual Typhoon Report, Fleet Weather Central/Joint Typhoon Warning Center U.S.A. 1959-1978.

Aoki, T., 1979: A Statistical Prediction of the Tropical Cyclone Position Based on Persistence and Climatological Factor in the Western North Pacific (the PC Method), Geophysical Magazine vol. 38, No. 4, pp. 17-28.

Chen, Shi-Yang, 1980: A Study on Prediction of Typhoon tracks with Analog Method. Mete. Bulletin vol. 26, No. 4, pp. 1-13. (Publish in Chinese with English abstract)

Chen, Shi-Yang, 1983: Comparison of Different Track Composite Method in CWB-80 Typhoon Analog Model. Mete. Bulletin vol. 29 No. 1, pp. 1-11.

Hu, Chung-Ying, and Shi-Yang, Chen, 1976: A Study of Typhoon Track Forecast with Objective Analog Method. Mete. Bulletin vol. 22, pp. 8-16.
(publish in Chinese with English abstract).

Hope, J. R, and Neumann, C, J, 1970: An Operational Technique for Relating the Movement of Existing Tropical Cyclones to Past Track, Monthly Weather Review, vol. 98, no. 12, pp. 925-933.

Hope, J. R, and Neumann, C,. J, 1977: A Survey of Worldwide Tropical Cyclone Prediction Models, post print vol. American Meteorology Society 11th Tech. Conf. on Hurricane and Tropical Meteo., Miami, Dec. 13-16, 1977, 367 pp.

Jarrell, J. D, and Somervell, Jr. W. L, 1970: A Computer Technique for Using Typhoon Analogs as A Forecast Aid, Naval Weather Research Facility, Tech. paper No. 6-70.

Jarrell, J. D, and Wagoner, R. A, 1973: The 1972 Typhoon Analog Program (TYFOON-72), Enviromental Prediction Research Facility, Tech. Paper No. 1-73.

Jarrell, J. D, C. J. Mauck and R. J. Renard, 1975: Forecasting Tropical Cyclone Motion Over The Northeastern Pacific Ocean By An Analog Scheme, Monthly Weather Review, vol. 103, pp. 674-684.

Neumann, C. J, and Hope, J. R, 1972: Performance Analysis of the HURRAN Tropical Cyclone Forecasting System, Monthly Weather Review, vol. 100, no. 4, pp. 245-255.

颱風類比模式之路徑預測誤差的統計比較

陳　熙　揚[1]　*Davis K. Cope*[2]

摘　要

本文旨在探討利用類比法製作颱風路徑預報，並提出未來研究之方向。

作者首先利用3至13點，每點間隔6小時的過去颱風中心位置，以 1 至 4 階曲線擬合 (Curve Fitting) 再加以外延的方式，預測12至36小時的颱風路徑，再以此預測路徑爲基準，比較颶風模擬 (HURRAN) ，單式權重 (Single Weighting) 及雙式權重 (Double Weighting) 等颱風類比模式的優劣。

文內除了比較各模式的預測穩定性及預測偏誤外，並以誤差的50%概率橢圓爲準，比較模式的預測準確性。經以1976及1978年西北太平洋颱風最佳路徑資料校驗的結果，發現曲線擬合法的穩定性不佳，兩年間變化的幅度約爲兩倍左右；以整體而言，3點線性外延有最準確的預測結果，但其預測偏誤却較3點2階多項式外延爲大，並且有明顯的經向偏誤。

在類比模式方面：以預測穩定性而言，三種模式均較3點線性外延爲佳，其中又以颶風模擬法最優，兩年間的預測誤差較爲接近；次就預測準確性而言，颶風模擬法有最準確的預測結果，12至36小時預測誤差的 50%概率橢圓面積分別爲 1.515 、6.01及16.187平方度；而單式權重法的預測則與3點線性外延相若；至於雙式權重法則12小時的預測較單式權重法爲差，但36小時的預測却有較好的結果；再就預測偏誤而言，颶風模擬法及單式權重法有較穩定的預測偏誤，以兩年預測的平均而言，前者 12 至 36 小時預測之 (緯度，經度) 偏値爲 (−0.04,0.20)，(−0.16,0.59) 及 (−0.49,0.87)，而後者則分別爲 (−0.22,−0.09) ，(−0.48,−0.09)及(−0.78,−0.23)；至於雙式權重法，則在經度及緯度方向均有大偏誤。

作者認爲颶風模擬法及單式權重法在目前最適合實際應用，因爲此兩法均有穩定的預測偏誤，經由偏誤的調整將可得到無偏誤的預報。

檢討颶風模擬法之所以有較準確的結果，部份原因乃在於其對類似颱風有較嚴格的篩選條件，尤其是當颱風路徑非常不規則時，此法每因無法選到類似颱風而無預報，但此時單式權重與雙式權重法由於篩選條件過寬，往往提供甚差的預報結果。故縮緊對類似颱風的篩選條件，乃是今後單式權重與雙式權重法必須採取的措施。此外，對於類似颱風最佳點的選取，類似路徑之調整及其優先順序與權重分佈等對預測準確性的影響，都將是日後研究之重點所在。

1. 中央氣象局資訊測站副主任。
2. 美國北達科塔州立大學數學部副教授。