交通部中央氣象局委託研究計畫成果報告

西北太平洋颱風路徑機率月與季預報

計畫類別：■國內　　□國外

計畫編號：MOTC-CWB-95-3M-06

執行期間：95 年 2 月 9 日至 95 年 12 月 31 日

計畫主持人：朱寶信(Pao-Shin Chu)

執行單位：Climate Systems Enterprise

中華民國 95 年 11 月 26 日

Climate Prediction of Tropical Cyclones Activity in the Vicinity of Taiwan Using the multivariate least absolute deviation regression method

Final Report of the Project MOTC-CWB-95-3M-06

Pao-Shin Chu

Climate Systems Enterprise

U.S.A.

November 26, 2006

Abstract

In this study, a multivariate linear regression model is applied to predict the annual tropical cyclone (TC) counts in the vicinity of Taiwan using large-scale climate variables available in May. The model is based on the least absolute deviation so that regression estimates are more resistant (i.e., not unduly influenced by outliers) than those derived from the ordinary least square method. Through correlation analysis, four variables (sea surface temperature, sea level pressure, precipitable water and low-level relative vorticity) in key locations of the tropical western North Pacific and a CLIPER (Climatology and Persistence) variable are identified as predictor data sets. Results from cross-validation suggest that the statistical model is skillful in predicting TC activity, with a correlation coefficient of 0.77 for the test period of 1970-2003 (34 years). Relative importance of each predictor variable is evaluated. For example, for predicting higher than normal typhoon activity, warmer sea surface temperatures, lower sea level pressures, and a moist troposphere in May appear to be important.

1. Introduction

Taiwan has been ravaged by typhoons with enormous property damage and loss of lives from time to time. Because of their catastrophic socio-economic consequences, there is an interest in understanding climate controls that are instrumental for the year-to-year typhoon variability and developing a basis for seasonal prediction of typhoon activity. The predictive information, if properly made, could be important for decision makers in relevant agencies to mitigate efforts of potential flooding, storm surges, and high winds resulting from typhoons.

W. Gray pioneered the seasonal hurricane prediction enterprise using a regression-based statistical model called the least absolute deviation method (e.g., Gray et al., 1992). They showed that nearly half of the interannual variability of hurricane activity in the North Atlantic could be predicted in advance. This is amazing because a hurricane is a small system, and physical mechanisms governing its formation are complicated. Also based on a regression model, Elsner and Schmertmann (1993) attempted to predict intense annual Atlantic hurricane counts with some success. Along the same line of statistical modeling, Chan et al. (1998) developed a model to predict seasonal typhoon activity over the western North Pacific and the South China Sea. When tested against a 30-yr sample through the jackknife method, skillful forecasts are noted for a suite of predictands (e.g., the annual number of typhoons and the annual number of tropical storms and typhoons). The informative results of Chan et al. (1998) are pertinent to the vast western North Pacific basin and the South China Sea. For a smaller geographic domain such as in the vicinity of Taiwan, the frequency of typhoon occurrences may be very different from that of the basin-wide numbers. For example, tropical cyclone counts over the entire western North Pacific during the developing year of the El Ni隳 do not differ much from the long-term climatology, yet the genesis location during the peak and late season of the El Ni隳 developing year is dramatically shifted east- and southward so fewer storms are found to the west of 150蚩 (Wang and Chan, 2002; Chu, 2004). Therefore, it has yet to be demonstrated that new predictive models would work for a smaller area.

In this study, we attempt to predict the annual number of tropical cyclones (TCs) in the vicinity of Taiwan area on the basis of the Least Absolute Deviation (LAD) regression method. This method has been tested for many years by Gray and his associates and is quite mature. Section 2 discusses the dataset, and section 3 outlines the LAD model. In section 4, procedures for selecting appropriate predictor variables are described. Section 5 discusses the forecast results. A summary is found in section 6.

2. Data and data processing

The tropical cyclone (tropical storms and typhoons) series in the vicinity of Taiwan from 1970 to 2003 compiled by the Central Weather Bureau is used. This series covers an area between 21祇 -26祇 and 119蚩 -125蚩 .

Monthly mean sea level pressure, wind data at 850- and 200-hPa levels, relative vorticity data at the 850 hPa level, and total precipitable water over the western North Pacific (0□-30祇) are derived from the NCEP/NCAR reanalysis dataset (Kalnay et al., 1997; Kistler et al., 2001). The horizontal resolution of the reanalysis dataset is 2.5? latitude-longitude. Tropospheric vertical wind shear is computed as the square root of the sum of the square of the difference in zonal wind component between 850- and 200-hPa levels and the square of the difference in meridional wind component between 850- and 200-hPa levels (Clark and Chu, 2002). The monthly mean sea surface temperatures, at 2? horizontal resolution, are taken from the NOAA Climate Diagnostic Center in Boulder, Colorado. As our interest is to develop a prototype model, large-scale environmental parameters only for the month of May are derived.

3. Least Absolute Deviations regression

A linear regression model can be generally written as

$$y(t) = \sum_{j=1}^{K} c_j x_j(t) + N(t) \tag{1}$$

where $y(t)$ is the desired predictive variable or predictand, $x_i(t), i = 1,...,K$ represent the predictors and $c_i, i = 1,...,K$ are the corresponding regression parameters, while $N(t)$ is a random variable and represents the regression deviations (residuals). The least square error (LSE) is probably the best known method for fitting linear regression models and by far the most widely used. However, the LSE is not necessarily the optimum fitting method if the deviation $N(t)$ is not of the Gaussian distribution. Moreover, the residuals in the LAD are computed from the median, whereas in the LSE they are derived from the mean. Because the median is a robust estimator of the location than the mean, LAD regression estimates are less sensitive to large outliers (e.g., extreme values) than the LSE method. In particular, if the deviation is double exponentially distributed, the optimum method for linear regression will be LAD.

The basic idea of LAD regression problem is generally stated as below. Given a sample size of n points $\{\underline{\mathbf{x_i}}, y_i\}, \underline{\mathbf{x_i}} \in R^k, i = 1,\ldots,n$, the LAD fitting problem is to find a minimizer, $\hat{\underline{\mathbf{c}}} \in R^K$, of the distance function (absolute deviation):

$$f(\underline{\mathbf{c}}) = \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{K} c_j x_{ij} \right| = \sum_{i=1}^{n} \left| y_i - \langle \underline{\mathbf{c}}, \underline{\mathbf{x_i}} \rangle \right| = \left\| \underline{\mathbf{y}} - \mathbf{x}\underline{\mathbf{c}} \right\|_1 \qquad (2)$$

where $\underline{\mathbf{y}} = [y_1, y_2, \ldots, y_n]'$, $\mathbf{x} = [\underline{\mathbf{x}_1}, \underline{\mathbf{x}_2}, \ldots, \underline{\mathbf{x}_n}]'$. $\underline{\mathbf{c}} = [c_1, c_2, \ldots, c_K]'$

such that, $f(\hat{\underline{\mathbf{c}}}) = \min(f(\underline{\mathbf{c}}))$.

This problem is solvable because $f(c)$ is continuous and convex. Due to the nonlinearity of absolute operation, solving LAD problem is no longer a linear problem. In the past, an abundance of algorithms was developed based on the well-studied linear programming (LP) problem, since LAD and LP are similar in their very basic nature. The LP problem in standard form is to find $\hat{\underline{x}}$ which maximizes $f(\underline{x}) = \langle \underline{c}, \underline{x} \rangle$ subject to constraint $A\underline{x} \le \underline{b}$ and $\underline{x} \ge 0$ with given vector $\underline{c}$, $\underline{b}$ and matrix A. With some straightforward but tedious derivations, it can be shown that any LAD curve-fitting can be expressed as an equivalent bounded feasible LP problem. We choose the Bloomfield-Steiger algorithm to find the minimizer (1980). The basic idea of this algorithm is to find the normalized steepest direction in each iteration of the algorithm. Suppose the current fit is $\underline{c}$, and $\underline{\delta_1}, \underline{\delta_2}, \ldots, \underline{\delta_K}$ is a set of directions along which the next iteration could move, the optimum descent direction being $\underline{\delta_p}$ along which

$$\min(f(\underline{c}+t\underline{\delta_p}), t \in R) = \min[\min(f(\underline{c}+t\underline{\delta_i})), i \le K] \quad (3)$$

the inner minimization over t in R. To find this direction, the K-weighted median calculations would need to be done (one for each i in the right hand side of the equation (3)). A pseudo code for the Bloomfield-Steiger (BS) algorithm is listed in Appendix A.

4. Procedures for selecting predictor variables

Tropical cyclone activity, on the seasonal time scale, is modulated by the large-scale environmental conditions (e.g., Gray, 1977; Watterson et al., 1995). In the following, correlation analysis between the TC occurrences and the environmental parameters (e.g., SST, relative vorticity) over the tropical western North Pacific is used to identify their physical relationships. If correlations over a particular area of the western Pacific are found to be statistically significant at the 95% confidence level, the parameter over this area is retained as a potential predictor variable. For a sample size of 34, this critical value is 0.34 when a two-tailed t-test is applied.

4.1 Sea surface temperatures (SSTs)

SSTs are known to be important for TC formation and intensification. Warmer SSTs are expected to fuel the overlying atmosphere with additional warmth and moisture, thereby reducing atmospheric stability and increasing the likelihood of deep tropical convection. Deep convection such as organized thunderstorm clouds acts as the primary mechanism for the vertical coupling of the lower and upper tropospheric flow patterns for TC formation. The contour plot for the correlation between TC counts and SST is shown in Fig. 1a, in which a maximum value of 0.58 is found in the core of the warm pool (2? 146崖). For simplicity, we choose just one area where the SST is most highly (in absolute value) correlated with the number of TCs.

4.2 Sea level pressures (SLPs)

The contour plot for the correlation between the TC frequency in the vicinity of Taiwan and the May SLP is shown in Fig. 1b. The highest negative correlation (-0.48) is found at 15祇 , 130豈 , very close to Taiwan. This result is physically reasonable as lower SLPs to the southeast of Taiwan in May correspond to higher TC frequency near Taiwan, and vice versa. Dynamically, the juxtaposition of the maximum correlations found in Figs. 1a and 1b suggests a Rossby-wave type response of atmosphere to equatorial heating as demonstrated in Gill model.

4.3 Precipitable water (PW)

The entrainment of drier air in the midtropospheric results in less buoyancy for the tropical convection systems as well as diminish the upper-level warming due to decreased release of latent heat (Knaff, 1997). Consequently, drier atmosphere tends to suppress deep convection and inhibits TC activity. Positive and strong correlations between PW and TC frequency are found in the core of the tropical western North Pacific (Fig. 2a). Thus, more moisture in the atmosphere in the tropical western North Pacific in May is conducive to more TC activity near Taiwan and vice versa. Of particular note is a smaller area of statistically significant correlation (0.61) near 10祇 , 137.5豈 , and this maximum center lies between the maximum correlations found in Figs. 1a and 1b.

4.4 Low-level relative vorticity (VOR)

The monsoon trough in the western North Pacific is characterized by the strong relative cyclonic vorticity in the lower troposphere and is known to be the birthplace of typhoons. Figure 2b displays a positive and high correlation in the Philippine Sea (17.5祇 , 132.5豈) with a value of 0.47 which is statistically significant. This result is internally consistent with those in Figs. 1b and 2a, in that lower SLPs in the Philippine Sea induce stronger cyclonic vorticity near the surface and higher moisture in the atmosphere, leading to more TC frequency in the vicinity of Taiwan. Also note that the maximum correlations in Figs. 1 and 2 are statistically significant at the 95% confidence level based on a t-test when climatological persistence is taken into account (Clark and Chu, 2002).

4.5 Vertical wind shear (VWS)

Strong VWS disrupts the organized deep convection (the so-called ventilation effect) which inhibits intensification of the TCs. Negative but weak correlations exist in the low latitudes with a center near 5祇 , 155崖 (Fig. 3a). This correlation (-0.31) is, however, statistically insignificant. A positive and strong correlation is noted near 5祇 , 120崖 but that location is too close to the western boundary of the domain and there is a lack of physical explanation for the VWS and TC frequency. As a result, VWS is not used in the subsequent forecasting simulation.

4.6 CLIPER

The variance analysis method is a statistical technique to test the existence of hidden periods in a time series. The basic idea of this method is found in Appendix B. Suppose the hidden periods for the series are given, we can use CLIPER method to find the periodical oscillation. The basic idea for CLIPER prediction is very simple: for the first significant period, say p , we can re-group the data as defined in formula (B1). Then we can calculate each group mean, which will be the basis for prediction for this group. Figure 3b displays the variance analysis method. The significant period at the 95% confidence level is 16 years.

5. Prediction results

With the various predictor variables selected through correlation analysis, we then use a cross-validation method to establish the overall forecasting ability of LAD model. The approach is as follows (Yu et al., 1997). The predictor and predictand data set of T time points are divided into L segments. A model is then developed using the data of L-1 segments. This model is then used to predict TC frequency in the remaining segment. This process is repeated by changing the segment that has been excluded from the model development. In this study, we remove only one observation at a time for each case. By doing this, we obtain N predictions. These predicted values are then correlated with N observations and the overall forecast skill can be determined. The cross-validation results are shown in Fig. 4 and a reasonably skillful forecast is seen. In some years, forecast values are smaller than actual observations (e.g., 1982) but in other years they are larger than observations (e.g., 1998). Thus, there is no systematic bias revealed in the prediction scheme. The linear correlation between the cross-validation results and the raw TC data is 0.77.

It is also interesting to compare the correlation skill from the regression based forecast scheme with that obtained from a random forecast. A random sample is generated from the 34 observed TC values ranging from one to seven using a random number generator, and a correlation coefficient between the randomly generated time series and the observed data is computed. This procedure is repeated 10,000 times and a histogram of the correlation coefficient can be constructed. Results from the Monte Carlo simulation reveal the correlation coefficient at the 95% (99%) confidence level is

0.33 (0.42). Accordingly, the correlation coefficient (0.77) from the LAD model in this study is deemed to be skillful relative to the benchmark random samples.

6. Relative contribution of predictor variables

This section examines the relative importance of the predictor variables, which is reflected in the contingency relationships between the predictand and the predictors. The predictors and predictands are each classified into three categories using data from 1970-2003. The predictor categories are based on the 30:40:30 proportion of the probability distribution. However, a similar proportion cannot be followed for the predictand because of the limited data length and the repeated nature of typhoon counts. In this study, we set the elow? and bove? thresholds from the ormal? category as two and five, respectively. That is, a year is classified as elow? when the annual typhoon count equals or less than two. Conversely, a year is chosen as bove? when the count equals or greater than five. As a result, eight years are chosen as elow? than normal years. They are 1972, 1973, 1983, 1988, 1989, 1993, 1997, and 1999. For the bove? normal years, they are 1978, 1982, 1985, 1986, 1987, 1990, 1994, 1998, 2000, 2001, and 2003.

Table 1a illustrates that during the elow? typhoon years, SLP tends to be normal or above normal, while PW and vorticity (VOR) tend to be normal or below normal. During the bove? typhoon years (Table 1b), SST shows a tendency to be normal or above normal. There is a preference for VOR to be in the normal or above normal category, and the PW varies in a similar manner as VOR. Also note that in Table 1a, there are five years (1973, 1983, 1989, 1993 and 1997) in which the signs for both VOR and PW are all negative when the typhoon count is below normal. When typhoon count is above normal (Table 1b), there are seven years (1978, 1985, 1990, 1994, 2000, 2001, and 2003) in which a positive signal appears in PW, and a negative signal appears in SLP. Based on the composite of the above and below normal typhoon years, one can further clarify the category relationship between the predictors and the predictand by computing the category percentage for each predictor. In Figure 5a, during the elow□ typhoon years the percentages of below normal PW and VOR are very high, reaching 75%, while a lack of above normal PW and VOR is conspicuous. Interestingly, during the bove? typhoon years (Fig. 5b), the percentage of below normal SLP is rather high (73%). Because the frequency of below PW is very low when typhoon count is low, a strong linear relationship between PW and typhoon numbers is noted. For other variables, nonlinear relationships are more evident. For example, SLP appears to be a good predictor during active typhoon years but less so for low typhoon years. The low level relative vorticity is important only during the elow? typhoon years.

7. Summary

Climate prediction of tropical cyclone activity has traditionally been carried out for the North Atlantic and the western North Pacific by various research teams. Because of the vast expanse of ocean basins and pronounced interannual climate variations in the tropics, there is no guarantee that such basin-wide prediction is also applicable to smaller regions within a basin. In this study, a multivariate least-absolute-deviation regression method is adopted to predict the annual tropical cyclone frequency in the vicinity of Taiwan using large-scale climate information available in May. Through correlation analysis between TC frequency and each individual climate variables (e.g., SST, SLP) over the tropical western North Pacific, we identified key locations to be used as the predictor data sets. We then used the leave-one-out cross-validation technique to test the predictability of TC frequency. The cross-validation provides a nearly unbiased estimate of true forecast skill.

The linear correlation between the cross-validation predictions and the corresponding actual observations for the test period 1970 -2003 is 0.77, which is significant at the 95% confidence level. This result implies that it would be possible to predict the annual TC counts for a small area with reasonable skill using a physically based regression model. To evaluate the relative contribution of each predictor, the predictors and predictand are each classified into three categories. Warmer SST, lower SLP, and higher precipitable water are the precursors to higher than normal typhoon activity. On the other hand, lower precipitable water and lower low-level relative vorticity are characteristics of lower than normal typhoon activity. In the future, it would be of interest to determine the predictability of TC frequency when climate variables chosen are for months earlier than May. Apart from pure scientific inquiry, if good skills can be obtained, say, when the April predictors are chosen, it would allow decision makers more lead time to respond.

Acknowledgments

This project has been funded by the Central Weather Bureau, CWB 95-3M-06.

Appendix:

A. **Pseudo Code for the BS Algorithm**:

[1] Accept $\{\underline{\mathbf{x}}_{\mathbf{i}}, y_i\}, \underline{\mathbf{x}}_{\mathbf{i}} \in R^k, i = 1,...,n$

[2] $A \leftarrow \begin{pmatrix} \mathbf{X} & \underline{\mathbf{y}} \\ \mathbf{I}_k & \underline{\mathbf{0}} \end{pmatrix}$, where $\mathbf{X} = [\underline{\mathbf{x}}_1,...,\underline{\mathbf{x}}_{\mathbf{n}}]'$, $\underline{\mathbf{y}} = [y_1,..., y_n]'$, $\mathbf{I}_k$ is an $k \times k$ identity matrix and $\underline{\mathbf{0}} \in R^k$

[3] $j \leftarrow 0$

[4] $\mathbf{Z} \leftarrow \{i, 1 \le i \le n : A_{\mathrm{i,k+1}} = 0\}$

[5] for m = 1 to k

$$g_m = -\sum_{\mathbf{Z}} |A_{i,m}|$$

$$h_m = \sum_{\mathbf{Z'}} A_{i,m} sign(A_{i,k+1})$$

$$f_m = \max(g_m - h_m, g_m + h_m) / \sum_{i=1}^{n} |A_{i,m}|$$

end

[6] $p \leftarrow \arg\min_i \{f_i, i = 1,..., n\}$

if $f_p \le 0$ go to [10]

[7] $t \leftarrow A_{q,k+1} / A_{qp}$, the weighted median of $\{A_{i,k+1} / A_{i,p}, \mathrm{i} = 1,...\mathrm{n}\}$ with weights $|A_{i,p}|$

[8] for matrix $\mathbf{A}$:

column $p \leftarrow$ column p / A_{qp}

column $i \leftarrow$ column $i - A_{qi}$ * column p

[9] $j \leftarrow j + 1$ go to [4]

[10] $\underline{\mathbf{c}}_i \leftarrow -A_{n+i,k+1}, \mathrm{i} = 1,..., \mathrm{k}$

B. Variance Analysis

In order to testify that there is a hidden period p existing in a given series $x_i, i = 1,..., N$, where $2 \le p \le \left\lfloor \frac{N}{2} \right\rfloor$, one can divide the series into p groups $y_j, j = 1,..., p$, where the k-th element of the sub-group y_j is defined as

$$y_{jk} = x_{(k-1)*p+j}, j = 1,..., p, k = \begin{cases} \left\lfloor \frac{N}{p} \right\rfloor, if\ j > \mathrm{mod}(N,p) \\ \left\lceil \frac{N}{p} \right\rceil, otherwise \end{cases} \tag{B1}$$

Then, one calculates the within-group variance and among-group variance given by

$$S_{within} = \sum (y_{jk} - \bar{y}_j)^2 / (N - p) \tag{B2}$$

$$S_{among} = \sum_{j=1}^{p} n_j (\bar{y}_j - \bar{x})^2 / (p - 1) \tag{B3}$$

where n_j represents the number of elements in j-th sub-group

Note that these variances should be normalized by their respective degrees of freedom. Last, we calculate the ratio S_{among} / S_{within} and compare it to the critical value determined by given confidence α (for example 95%), that is $F(\alpha, p-1, N-p)$ where F represents F-distribution. If this ratio is bigger than the critical value, it will suggest that the original series has a significant period of p.

References:

Bloomfield, P., and W. L. Steiger, 1980: Least absolute deviations curve-fitting. *J. Sci. Stat. Comput.*, **1**, 290-301.

Chan, J.C.L., J.-S. Shi, and C.-M. Lam, 1998: Seasonal forecasting of tropical cyclone activity over the western North Pacific and the South China Sea. *Wea. Forecasting*, **13**, 997-1004.

Chu, P.-S., 2004: ENSO and tropical cyclone activity. *Hurricanes and Typhoons: Past, Present, and Future*, R.J. Murnane and K.-B Liu, Eds., Columbia University Press, 297-332.

Clark, J.D., and P.-S. Chu, 2002: Interannual variation of tropical cyclone activity over the central North Pacific. *J. Meteor. Soc. Japan*, **80**, 403-418.

Elsner, J.B., and C.P. Schmertmann, 1993: Improving extended-range seasonal predictions of intense Atlantic hurricane activity. Wea. Forecasting, 8, 345-351.

Gray, W.M., 1977: Tropical cyclone genesis in the western North Pacific. *J. Meteor. Soc. Japan*, **55**, 465-482.

Gray, W.M., C.W. Landsea, P.W. Mielke, and K.J. Berry, 1992: Predicting Atlantic seasonal hurricane activity 6-11 months in advance. *Wea. Forecasting*, **7**, 440-455.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc*., **77**, 437-471.

Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc*., **82**, 247-267.

Knaff, J.A., 1997: Implications of summertime sea level pressure anomalies in the tropical Atlantic region. *J. Climate,* **10**, 789-804.

Wang, B., and J.C.L. Chan, 2002: How strong ENSO affect tropical storm activity over the western North Pacific. *J. Climate*, **15**, 1643-1658.

Watterson, I.G., J.L. Evans, and B.F. Ryan, 1995: Seasonal and interannual variability of tropical cyclogenesis: diagnostics from large-scale fields. *J. Climate*, **8**, 3052-3066.

Table 1 Predictor anomaly categories during the (a) below and (b) above typhoon (TY) years. Symbol – represents the below normal category, + above normal category, and O normal category.

(a) Below typhoon years

Year \ Var	TY	SST	SLP	PW	VOR
1972	-	-	**0**	-	**0**
1973	-	**0**	+	-	-
1983	-	-	+	-	-
1988	-	**0**	**0**	**0**	**0**
1989	-	**0**	**0**	-	-
1993	-	-	+	-	-
1997	-	**0**	+	-	-

(b) Above typhoon years

Year \ Var	TY	SST	SLP	PW	VOR
1978	+	-	-	+	**0**
1982	+	**0**	**0**	**0**	+
1985	+	**0**	-	+	+
1986	+	+	-	**0**	+
1987	+	**0**	+	-	**0**
1990	+	+	-	+	+
1994	+	+	-	+	**0**
1998	+	**0**	+	**0**	**0**
2000	+	+	-	+	+
2001	+	+	-	+	-
2003	+	+	-	+	**0**

Figures

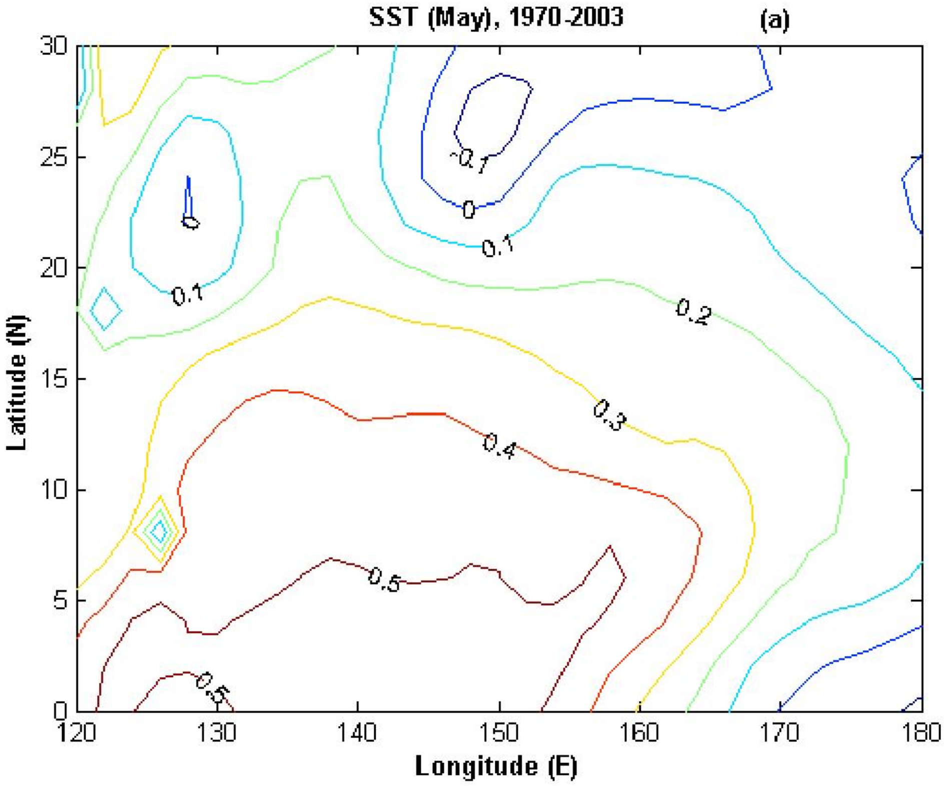
Fig. 1. (a) Correlation map between tropical cyclone count series in the vicinity of Taiwan and the sea surface temperatures over the tropical western North Pacific in May. (b) Same as (a) except for the sea level pressures in May.

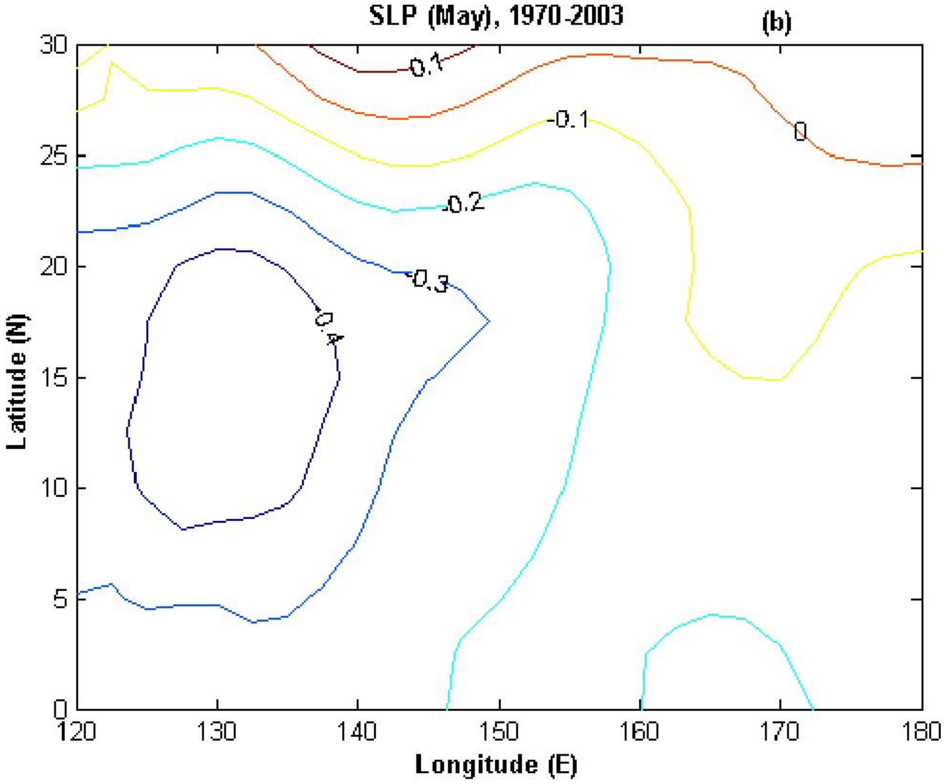
Fig. 2. (a) Same as Fig. 2(a), except for the precipitable water in May. (b) Same as Fig. 2(a), except for the relative vorticity at 850 hPa in May.

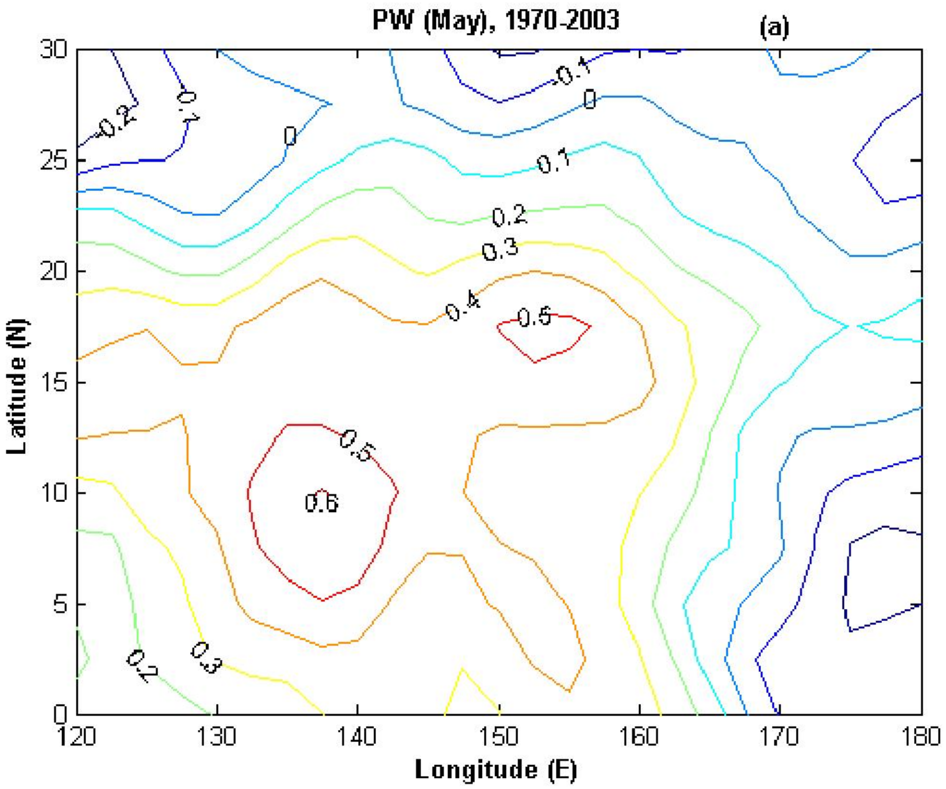
Fig. 3. (a) Same as Fig. 2(a), except for the vertical wind shear in May. (b) Variance analysis of cyclone count series. The x-axis is the years and y-axis is the F-score.

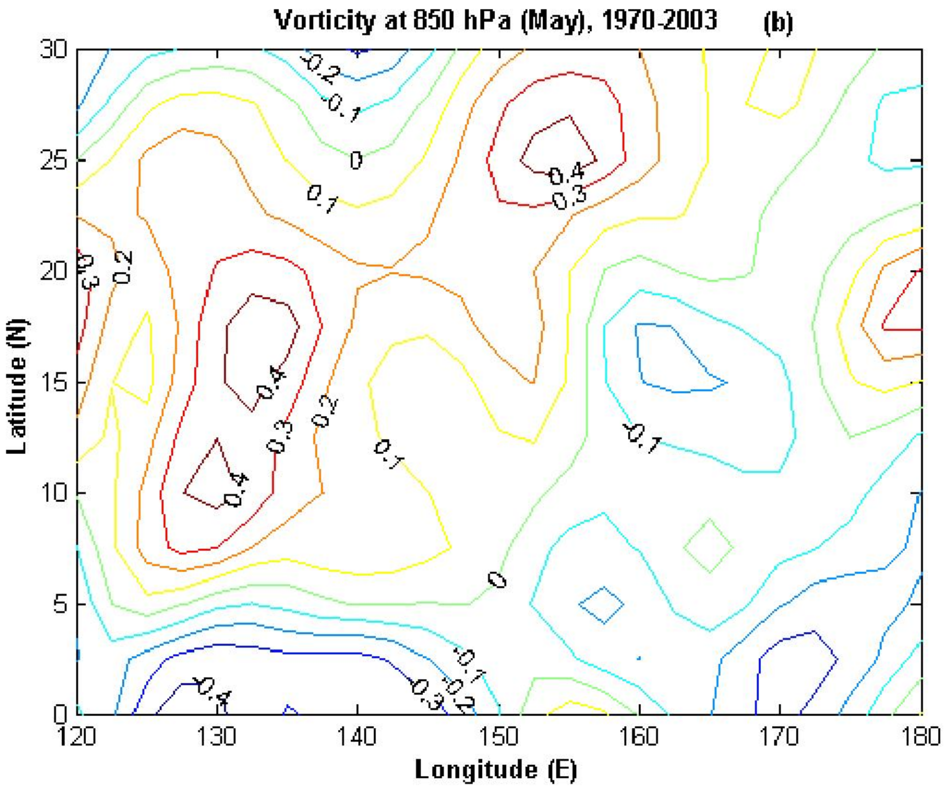
Fig. 4. Time series of observed (broken line) and cross-validated forecasts (solid line) of tropical cyclone counts.

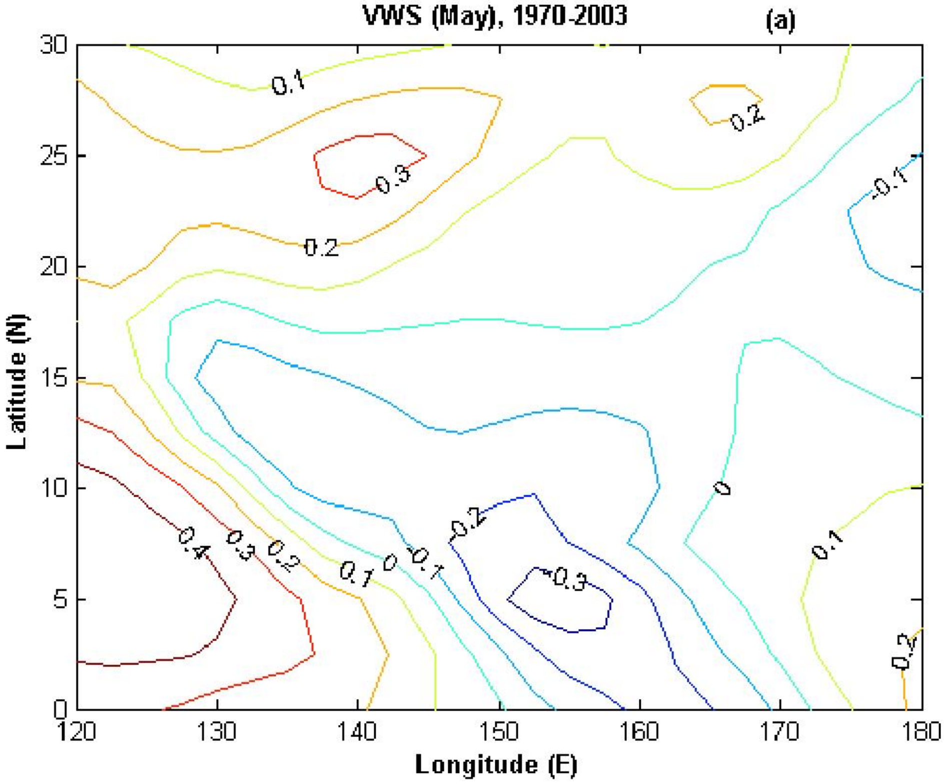
Fig. 5. The category percentage of each predictor during years of (a) below and (b) above typhoon counts.

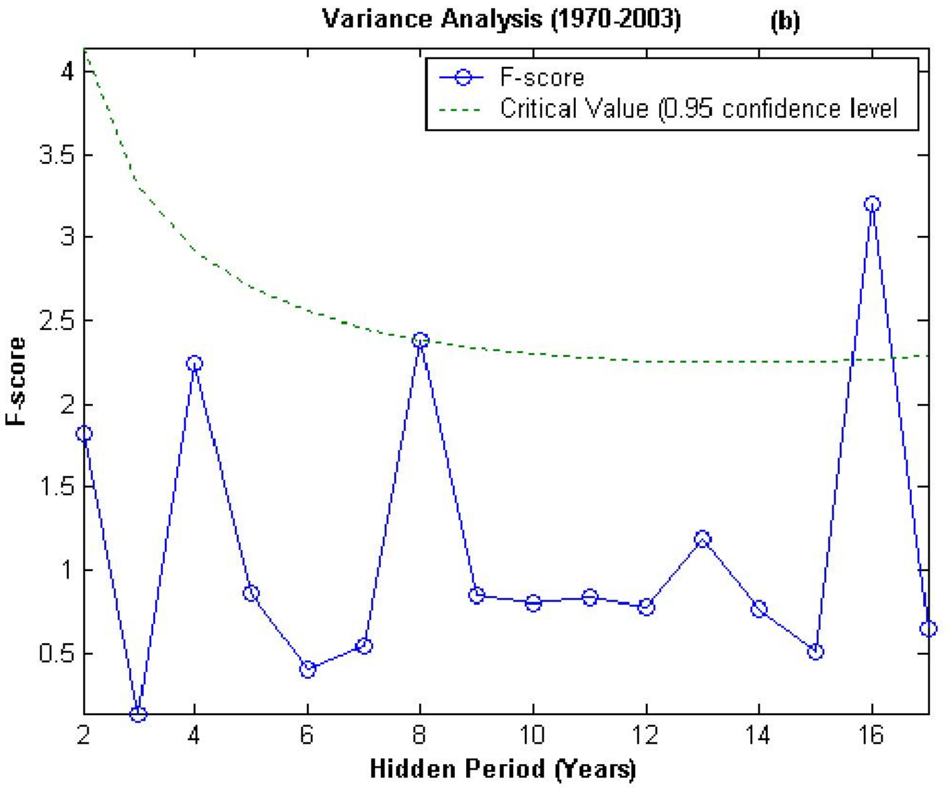

SST (May), 1970-2003
(a)
Latitude (N)
Longitude (E)
30
25
20
15
10
5
0
120
130
140
150
160
170
180
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.5
0.1
0

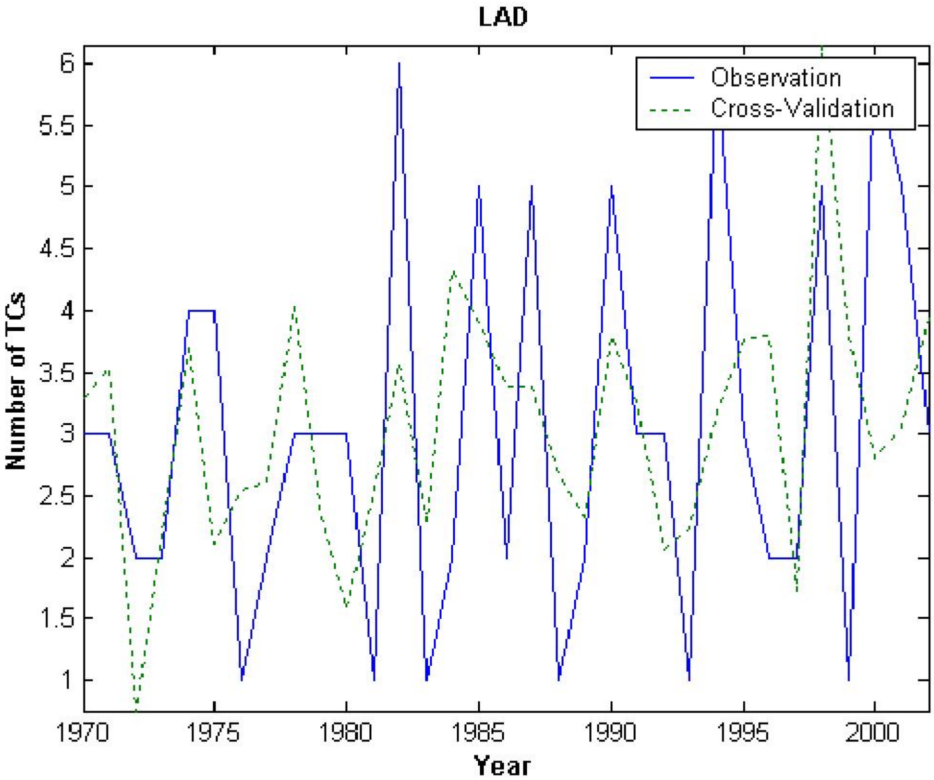

SLP (May), 1970-2003
(b)
Latitude (N)
Longitude (E)
30
25
20
15
10
5
0
120
130
140
150
160
170
180
0.1
-0.1
0
-0.2
-0.3
-0.4


PW (May), 1970-2003
(a)
Latitude (N)
Longitude (E)

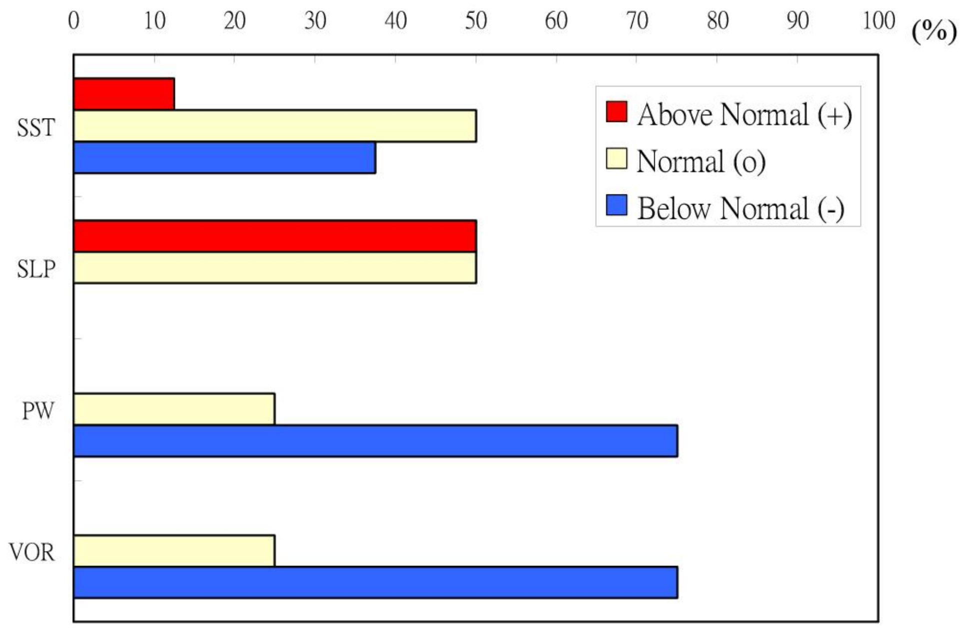

Vorticity at 850 hPa (May), 1970-2003
(b)
Latitude (N)
Longitude (E)


VWS (May), 1970-2003
(a)
Latitude (N)
Longitude (E)
0.1
0.2
0.3
0.2
-0.1
0
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0.1
0.2
30
25
20
15
10
5
0
120
130
140
150
160
170
180


Variance Analysis (1970-2003)
(b)
F-score
Critical Value (0.95 confidence level
F-score
Hidden Period (Years)
4
3.5
3
2.5
2
1.5
1
0.5
2
4
6
8
10
12
14
16

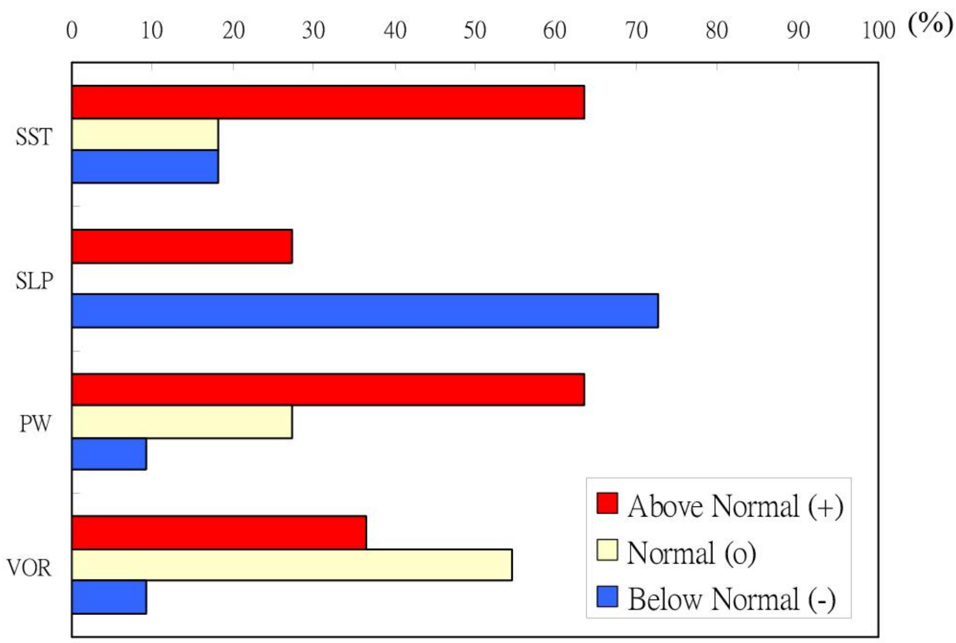

LAD
Observation
Cross-Validation
Number of TCs
Year
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
1970
1975
1980
1985
1990
1995
2000

(a)



0
10
20
30
40
50
60
70
80
90
100
(%)
SST
SLP
PW
VOR
Above Normal (+)
Normal (o)
Below Normal (-)

(b)



0
10
20
30
40
50
60
70
80
90
100
(%)
SST
SLP
PW
VOR
Above Normal (+)
Normal (o)
Below Normal (-)