交通部中央氣象局

委託研究計畫(期末)成果報告

颱風強度動力統計客觀預報方法之研發(1/2)

計畫類別：■氣象　□海象　□地震

計畫編號：MOTC-CWB-101-M-06

執行期間： 101 年 4 月 2 日至 101 年 12 月 31 日

計畫主持人：朱寶信

執行機構：Climate Systems Enterprise

中華民國　101　年　11　月　30　日

政府研究計畫(期末)報告摘要資料表

計畫中文名稱	颱風強度動力統計客觀預報方法之研發(1/2)		
計畫編號	MOTC-CWB-101-M-06		
主管機關	交通部中央氣象局		
執行機構	Climate Systems Enterprise		
年度	101 年	執行期間	101 年 4 月 2 日至 101 年 12 月 31 日
本期經費 (單位：千元)	685		
執行進度	預定（%）	實際（%）	比較（%）
	100%	100%	
經費支用	預定(千元)	實際(千元)	支用率（%）
	685	685	100%
研究人員	計畫主持人	協同主持人	研究助理
	朱寶信	Dr. X, Zhao	
報告頁數	27	使用語言	English
中英文關鍵詞	颱風強度預報, Prediction of Typhoon Intensity		
研究目的	採用先進的Bayesian Regression Model以及GFS data做颱風強度的預報		
研究成果	附件		
具體落實應用情形	可提供作業單位參考，加強颱風強度預報，以便納入全國防災預警系統		
計畫變更說明			
落後原因			
檢討與建議 (變更或落後之因應對策)			

(以下接全文報告)

An operational Statistical/Dynamical Typhoon Intensity Prediction Model

Pao-Shin Chu and XinZhao

Climate Systems Enterprise

Honolulu, Hawaii, U.S.A.

December 9, 2012

MOTC-CWB-101-M-06

1. Introduction

The proposed statistical/dynamical typhoon intensity prediction model is essentially a track pattern based Bayesian multi regression model. The data for the predictions are based on the numerical weather prediction computer model run by NOAA, namely Global Forecast System (GFS). This project shall provide an operational program, through which one can forecast the intensity of a typhoon that may affect the vicinity of Taiwan area. Given the nature of GFS data, this model shall provide forecasts up to 120 hours ahead in a 6-hour interval. The details of the proposed model are elaborated in the following context.

2. Data

The backbone of the data used for this project is the NOAA's GFS data which run from 2008 to 2011 and are provided by the Central Weather Bureau (CWB). In implementing forecast operation using our model system, the on-line GFS data are available on the website nomads.ncdc.noaa.gov. The GFS data are in 0.5 degree by 0.5

degree resolution and at 6-hour interval. Because the GFS data are generated by dynamical models and these data are used, our approach can be regarded as a statistical/dynamical method. Toward this project, we conduct the research for the typhoons occurred during the period from 2008 to 2011. Separately, the typhoon tracks over the western North Pacific in a 6-hour interval during 2005-2011 are also provided by CWB. In addition, we also used a global Sea Surface Temperature (SST) data to develop the maximum potential intensity (MPI) predictor for the forecast system. The daily SST data were downloaded from the NOAA website www.ncdc.noaa.gov and in 0.25 degree by 0.25 degree resolution. To fit the daily SST data to the 6-hour interval forecast, we perform the linear interpolation to get the 6-hour interval SST data.

3. Model Development and Predictor Selection

We adopt the statistical forecast model of Chu et al. (2010a). Essentially, the mathematical model is a Bayesian multi linear regression model. As in Knaff et al. (2005), the potential predictors used in this project can be divided into two categories: 1) those related to climatology, persistence and trends of typhoon track pattern and intensity, termed herein as “static predictors”; and 2) those related to current and future environmental conditions, termed herein as “time dependent or environmental predictors”.

In specifics, in the first part and the second part of this section, we shall respectively introduce the track pattern clustering algorithm and the Bayesian multiple regression model. Following the similar line suggested in Knaff et al. (2005), in the third part of this chapter, we shall address the development of the static part of the operational forecast model. In the last part of this section, we shall discuss the inclusion of on-line

GFS forecast data and SST data as the backbone of the time dependent part of the forecast model. The potential predictor pool is listed in Table 1-1 and 1-2. As a side note, all predictors are developed using a "perfect prog" methodology (Knaff et al., 2005) where the analyses and actual tropical cyclone best track are used.

3.1. Track Pattern Clustering

The predictor selection of the proposed forecast model is grounded on a track pattern oriented classification scheme. In this research, we shall use a finite mixture Gaussian model introduced in Camargo et al. (2007) and Chu et al. (2010b) to solve the tropical cyclone (TC) track clustering problem.

Based on the assumption that there are a few distinct types characterizing TC tracks in a basin of interest, we model each TC track path as a second-order polynomial function of the lifetime of this TC. Mathematically, for each specific track type, the set of coefficient of this polynomial function is presumably jointly Gaussian distributed. Each TC track type thus has a unique distribution parameter. The space spanned by the parameters of this track type model is a linear combination of a set of distinct Gaussian distributions.

Assuming there are n observed track records for a given TC. For each record, there are three features reported—latitude, longitude, and the time. We denote the path record of a TC and its relative observed time vector for the second order polynomial function, respectively, by

$$\mathbf{z} = [\mathbf{z}_{lat}, \mathbf{z}_{long}] = \begin{bmatrix} z_{1,lat} & z_{1,long} \\ \dots & \dots \\ z_{n,lat} & z_{n,long} \end{bmatrix}, \ \mathbf{T} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \dots & \dots & \dots \\ 1 & t_n & t_n^2 \end{bmatrix} \qquad \text{(1a)}$$

where $z_{i,lat}$ and $z_{i,long}$ for $i = 1,..,n$ represent the i-th latitude and longitude record; and t_i represents the time for the i-th records of this TC relative to the first record for $i = 1,..,n$. We further assume that there are K distinct TC track types in the basin of interest, where K is assumed to be a constant in a given hypothesis or model. With the model defined in Eq. (1a), if a TC is categorized as type k, $1 \le k \le K$, the link function between the TC track path and relative time is governed by the following formula

$$\mathbf{z} = \mathbf{T}\boldsymbol{\beta}^{\mathbf{k}} + \boldsymbol{\varepsilon}, \text{ where } \boldsymbol{\beta}^{\mathbf{k}} = \begin{bmatrix} \beta_{0,lat}^k \; \beta_{0,long}^k \\ \beta_{1,lat}^k \; \beta_{1,long}^k \\ \beta_{2,lat}^k \; \beta_{2,long}^k \end{bmatrix} \text{ and } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma}^{\mathbf{k}}),\ 1 \le k \le K. \tag{1b}$$

In model (1b), the parameter set $\boldsymbol{\beta}^{\mathbf{k}}$ is distinct for each TC clustering types and $N(\bullet,\bullet)$ denotes the normal distribution. With this model, intuitively one can see that the zero-order coefficient dual provides the mean genesis location of this clustering type; the first-order term features the characteristic linear direction of this path type; the second-order term determines the recurving shape of the typical path of this type; and the covariance matrix (Σ) determines the spread of a particular type. The noise term in model (1b), $\boldsymbol{\varepsilon}_i$, is assumed multivariate Gaussian with zero mean and a 2 by 2 covariance matrix, $\boldsymbol{\Sigma}_k$. The conditional density for the i-th cyclone, conditioned on membership in the cluster type k, is therefore defined as

$$P(\mathbf{z}_i \mid \mathbf{T}_i, \boldsymbol{\theta}_k) = (2\pi)^{-n_i} \left|\boldsymbol{\Sigma}_k\right|^{-n_i/2} \exp\left\{-\operatorname{tr}[(\mathbf{z}_i - \mathbf{T}_i\boldsymbol{\beta}_k)\boldsymbol{\Sigma}_k^{-1}(\mathbf{z}_i - \mathbf{T}_i\boldsymbol{\beta}_k)']/2\right\}. \tag{2a}$$

In Eq. (2a), operator $\exp\{\bullet\}$ denotes an exponential function with a natural base; we adopt the notation $\boldsymbol{\theta}_k = \{\boldsymbol{\beta}_k, \boldsymbol{\Sigma}_k\}$, which is referenced in model (1); and operator tr(.) denotes the matrix operation function “trace.” By the definition of a mixture Gaussian model, Eq. (2a) leads to the marginal mixture model

$$P(\mathbf{z}_i \mid \mathbf{T}_i) = \sum_{k=1}^{K} \alpha_k P(\mathbf{z}_i \mid \mathbf{T}_i, \boldsymbol{\theta}_k) \quad \text{(2b)}$$

where, $P(\mathbf{z}_i \mid \mathbf{T}_i, \boldsymbol{\theta}_k)$ is given by (2a), and α_k is the posterior probability of cluster k, which implies $\sum_{k=1}^{K} \alpha_k = 1$. If we let $\mathbf{Z}' = [\mathbf{z}_1^{'}, \mathbf{z}_2^{'}, ..., \mathbf{z}_N^{'}]$ be the complete set of all observed TC trajectories, and $\mathbf{T}' = [\mathbf{T}_1^{'}, \mathbf{T}_2^{'}, ..., \mathbf{T}_N^{'}]$ be the associated measurement times, then the full probability density of $\mathbf{Z}$ given $\mathbf{T}$, the conditional likelihood, is formulated by

$$P(\mathbf{Z} \mid \mathbf{T}) = \prod_{i=1}^{N} \sum_{k=1}^{K} \alpha_k P(\mathbf{z}_i \mid \mathbf{T}_i, \boldsymbol{\theta}_k) \quad \text{(3)}$$

where $P(\mathbf{z}_i \mid \mathbf{T}_i, \boldsymbol{\theta}_k)$ is defined in Eq. (2a). Assume that the number of cluster type, K, is given (for a real application, we can refer to the literature to choose the proper number for this parameter). Because hypothesis selection is not the focus of this section, it is proper to choose a non-informative prior for the model coefficients; that is, $P(\boldsymbol{\theta}_k, \alpha_k) \propto 1$ for model (3). With this non-informative prior assumption, and following the basic Bayes formula given in Eq. (1), the posterior distribution for $\{\boldsymbol{\theta}_k, \alpha_k\}$ is proportional to the conditional likelihood given in Eq. (3).

In many real-world applications, only the peak areas of the posterior distribution may be of interest. An efficient approach to estimating the mode of the posterior distribution is the Expectation-Maximization (EM) algorithm. Given the likelihood model (3), in the E-step the membership probability of a TC categorized to each clustering type is calculated. In the M-step, the optimization estimation for the model parameter set of each type is calculated. These include regression parameters, the posterior probability of cluster *k*, and the covariance matrix. The maximization formula for coefficient parameter

$\hat{\boldsymbol{\beta}}^{\mathbf{k}}$ and variance parameter $\boldsymbol{\Sigma}^{\mathbf{k}}$ are derived from a linear Bayesian regression model. The details of the formula for the EM algorithm are provided in Chu et al. (2010b).

Given the number of clusters and an initial setting of the model parameters, after a few iterations, the proposed EM algorithm will converge to a fixed set of parameter estimation. Usually, the convergence of an EM algorithm is determined when the difference between two iterations is less than a sufficiently small value. Note that these convergent values are not necessarily the global optimum estimation and are determined by the initial starting values. Therefore, multiple different initial values should be selected and the set of estimation with the maximum likelihood of the observation chosen. After applying the aforementioned clustering algorithm to the track pattern data, we noticed that there are 5 distinct track pattern types that affect the western North Pacific (Fig. 1).

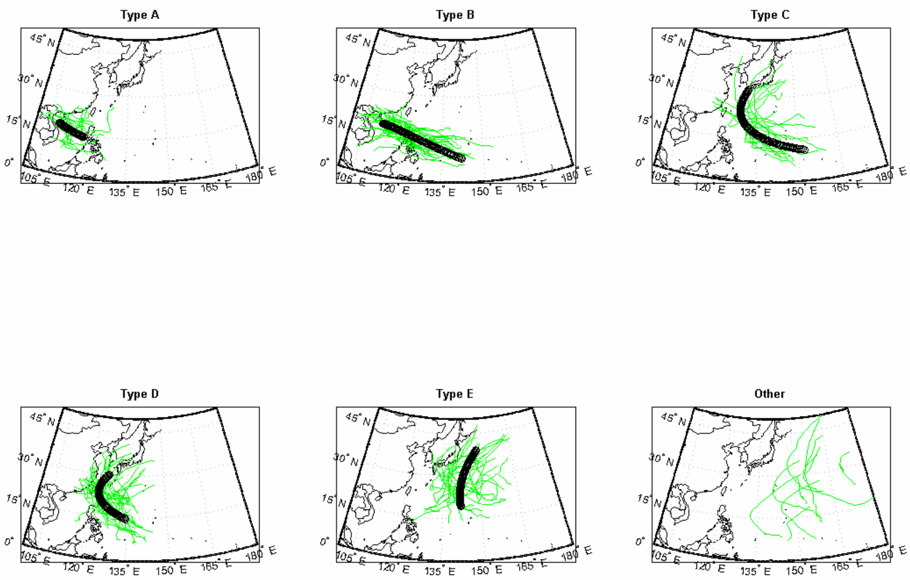

Fig. 1: Track pattern type for typhoons affect the vicinity of Taiwan (2005-2011)

In Fig.1, each reported typhoon path from 2005 to 2011 is plotted as a light green curve in its category. The black circles in each panel denote the mean track of each type. Some typhoon tracks are not distinct and are labeled as "Other". Based on the clustering results given in Fig. 1, during the 7-year period from 2005 to 2011, there are totally 148 typhoons, in which 20 Type A, 33 Type B, 17 Type C, 37 Type D, 29 Type E, and 12 other type. We thus shall first develop the static predictor set for each of these 5 types.

3.2 Bayesian Regression Model

Throughout this report, $Normal(\bullet)$ denotes the normal distribution, which is the foundation of the proposed regression model. In details, the classical Bayesian linear regression model can be formulated as follows

$\mathbf{Z} \mid \boldsymbol{\beta}, \sigma^2, \mathbf{X} \sim Normal(\mathbf{Z} \mid \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_N)$, where, specifically

$\mathbf{X}' = [\mathbf{X}_1^{'}, \mathbf{X}_2^{'}, ..., \mathbf{X}_N^{'}]$, $\mathbf{I}_N$ is the $N \times N$ identity matrix, and

$\mathbf{X}_i = [1, X_{i1}, X_{i2}, ..., X_{iK}]$ is the predictor vector for Z_i, $i = 1,2,...,N$,

$$\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, ..., \beta_K]'. \tag{4a}$$

In (4a), $\mathbf{Z}_i$ denotes the change in intensity for the i-th observation (for example, it can denote the 6-hour interval change); $\mathbf{X}_i$ denotes the selected predictor set for $\mathbf{Z}_i$; $\boldsymbol{\beta}$ is the regression coefficient vector, in which specifically β_0 is referred to as the intercept. Given the current intensity and intensity change forecast, the intensity forecast is obtained by adding the intensity change forecast and the current intensity measure.

Sometimes one may pay more attention on the probability if the typhoon intensity shall increase or decrease after a given time interval. Under this kind of scenario, a probit regression model is more appropriately. A probit regression model is very similar to the regression model defined in (4a). The only difference is that, the target variable $\mathbf{Z}$ defined in (4a) becomes a latent variable, such that

$$P(\mathbf{y} \mid \mathbf{Z}) = \prod_{i=1}^{N} P(y_i \mid Z_i), \text{ where } y_i = \begin{cases} 1 & if\ Z_i \geq 0 \\ 0 & if\ Z_i < 0 \end{cases}. \qquad (4b)$$

The inference solution for Model (4a) or (4b) can be found in most classical Bayesian analysis literatures. Towards the operational model and analysis results given in this report, we shall provide the technical details of the theoretical inference solution for the regression model (4a) as follows (Gelman et al. 2004).

The regression model (4a) in literatures is often referred to as "Ordinary Linear Regression". To solve this model, one usually chooses the non-informative prior for the model parameters $\{\boldsymbol{\beta}, \log(\sigma^2)\}$. That is,

$$P(\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-2} \qquad (5a)$$

With this prior, the posterior distribution for the model parameters can be derived as:

$P(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{X}, \mathbf{Z}) = P(\boldsymbol{\beta} \mid \sigma^2, \mathbf{X}, \mathbf{Z}) * P(\sigma^2 \mid \mathbf{X}, \mathbf{Z})$, in which

$P(\boldsymbol{\beta} \mid \sigma^2, \mathbf{X}, \mathbf{Z}) \sim Normal(\hat{\boldsymbol{\beta}}, V_{\boldsymbol{\beta}} \sigma^2)$,

$P(\sigma^2 \mid \mathbf{X}, \mathbf{Z}) \sim Inv - \chi^2(N-K, s^2)$, where

$V_{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1}$,

$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Z}$,

$$s^2 = \frac{1}{N-K} (\mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}}) \qquad (5b)$$

In (5b), $Inv-\chi^2(\bullet)$ denotes the inverse χ^2 distribution.

When a new predictor set $\tilde{\mathbf{X}}$ is available (such as through the GFS system), the inference for the intensity change for the next time interval is provided by:

$$P(\tilde{\mathbf{Z}} \mid \tilde{\mathbf{X}}, \mathbf{X}, \mathbf{Z}) \sim t_{N-K}(\tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}, s^2(\mathbf{I} + \tilde{\mathbf{X}} V_{\boldsymbol{\beta}} \tilde{\mathbf{X}}^T)) \tag{5c}$$

In (5c), s^2, $\hat{\boldsymbol{\beta}}$ and $V_{\boldsymbol{\beta}}$ are calculated in (5b) and the right hand side distribution $t_{N-K}(\bullet)$ denotes a multivariate t distribution with a form following standard convention. When the number of sample N is much larger than the predictor dimension K, the t distribution in (5c) can be well approximated by a normal distribution with mean $\tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}$ and variance matrix $s^2(\mathbf{I} + \tilde{\mathbf{X}} V_{\boldsymbol{\beta}} \tilde{\mathbf{X}}^T)$.

In the following context of this report, the regression model (4a) and its theoretical solution (5) for this regression model shall serve as the backbone for the layout of the analysis results and discussions of the rest of this report.

3.3 Static Predictor Selection

In the static predictor selection phase, we consider 8 potential candidates for the forecast for change in intensity after present time: 1) change in intensity before present time (DVMX); 2) current intensity (VMAX); 3) storm translational speed (SPD); 4) latitude at present time (LAT); 5) longitude at present time (LONG); 6) absolute value of yearday minus 248 (JDAY); 7) days after VMAX larger than 18m/s (RDAY); and 8) absolute value of RDAY minus one third of this typhoon type's expected life span (ARDAY, refer to Table A-1 for the expected life span used in this study). The candidate pool described above is summarized in Table 1-1.

Predictor	Description
DVMX	Intensity change before the present time
VMAX	Current intensity
SPD	Storm translational speed
LAT	Latitude
LONG	Longitude
JDAY	Absolute value of yearday minus 248
RDAY	Value of days after VMAX larger than 18 m/s
ARDAY	Absolute value of RDAY minus 1/3 of the expected life span of its type (Table A-1)

Table 1-1: The potential static predictors used in the model

3.4 Time-Dependent Predictor Selection

In this subsection, we shall focus on developing the time dependent or environmental predictors. The candidate pool of time dependent predictors is basically divided into three categories. The first one is those related to temperature field, including SST, T200 (temperature at 200 hPa level) and T925 (temperature at 925 hPa level). The second category includes those related to moisture field, including the relative humidity (RH) at high level (RHHI, which is the average of RH at 300 hPa, 350 hPa, 400 hPa and 450 hPa) and relative humidity at low level (RHLO, at 925 hPa). The third one is those related to wind and pressure fields, including the vertical wind shear (SHRD), relative vorticity (RV, at 850 hPa), and sea level pressure (SLP). Hereby, SHRD essentially measures the wind difference between low level and high level (850 hPa and 200 hPa respectively), which is calculated by the formula:

$$SHRG = 4 * \sum_{p=850}^{200} w_p * \sqrt{(u_p - \bar{u})^2 + (v_p - \bar{v})^2} \tag{6}$$

In (6), $\bar{u} = \sum_{p=850}^{200} w_p * u_p$ is the deep layer zonal wind; $\bar{v} = \sum_{p=850}^{200} w_p * v_p$ is the deep layer meridional wind; and w_p is mass weight, which is set as 0.5 for all the simulations in this report. For each of these predictors, we shall take the area-average from the GFS values in the shaded area illustrated in Fig. 2. In specific, the center of the two circles in Fig. 2 is the center of the current eye of the target typhoon. The inner circle has a radius of 200 km and the outer circle has a radius 800 km. There are three exceptions. For SST and SLP, we choose the data at the eye of a TC. For RV, we choose the area that is within a circle with radius 1000 km.

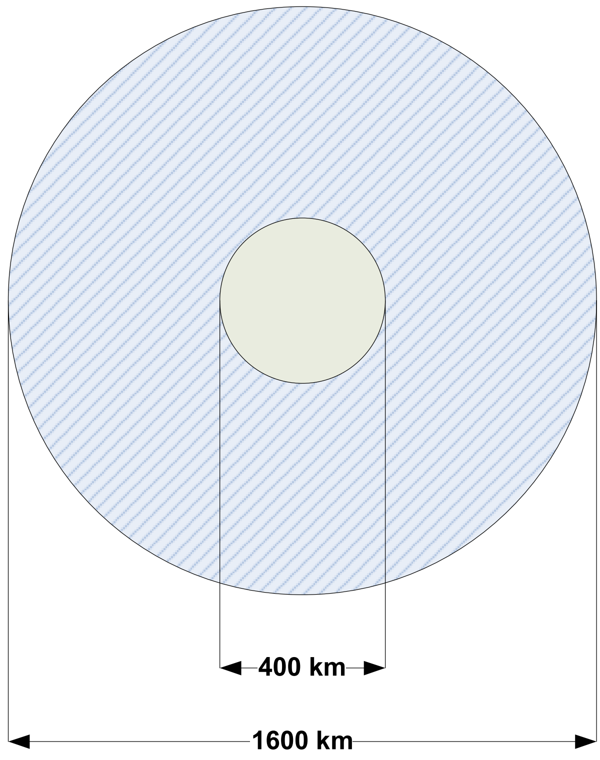

Fig. 2 The area for potential track based time dependent predictors

Predictor	Description
MPI	Maximum potential intensity defined in Eq. (7)
MPI2	MPI squared
T200	Area average temperature at 200 hPa
T925	Area average temperature at 925 hPa
RHLO	Area average relative humidity at 925 hPa
RHHI	Area average relative humidity at 300, 350, 400, 450, 500 hPa (average)
SHRG	Generalized 200-850 hPa vertical wind shear defined in Eq. (6)
RV850	Area average (0-1000km) 850 hPa relative vorticity
SLP	Sea level pressure

Table 1-2: The potential environmental predictors used in the model

The primary use of the SST is to determine the upper bound of tropical cyclone intensity as a function of SST, which is commonly referred to as the maximum potential intensity (MPI). The SST values are determined at the typhoon center and we adopt the same procedure to develop the MPI as in Knaff et al. (2005). The formula of MPI given SST is:

$$MPI = A + B * \exp(C * (T - T_0)) \quad (7)$$

In (7), A = 38.21 knots (or 19.66 m/s), B = 170.72 knots (or 87.82 m/s), C $= 0.1909^{o}C^{-1}$, $T_0 = 30.0^{o}C$ and T is in the unit of ^{o}C. The candidate time dependent predictor pool is summarized in Table 1-2, in which the “area average” without specification denotes the average over the area defined in Fig. 2 centered at the eye of the TC.

In both Section 3.3 and Section 3.4, upon the predictors are chosen, the regression model is set, which applies a stepwise procedure to select variables from the predictor pool at each forecast time. The significance of each candidate predictor is based on a standard two tail F-test and the operational threshold is set as 99%. Once a predictor is chosen, we shall normalize this variable with respect to its time span.

4. Results and Discussions

In this section, we shall provide the simulation results and discussions with applying the proposed forecast framework to the data described in Section 2. To evaluate the model performance, we shall perform a leave-one-out cross-validation (LOOCV). The results thus shall be compared with the climatology and persistence (CLIPER) and the benchmark STIPS model developed in Knaff et al. (2005). We also assess the model performance using the same model development approach without using the track pattern classification (that is, no TC track type is involved). Throughout this report, we denote the track pattern classification based model as “Model Cluster” and the one without as “Model General”.

We applied the proposed forecast model to the typhoons occurred in the WNP during the period 2008 to 2011 (totally 76 typhoons). To compare with non-Bayesian benchmark forecast models “CLIPER” and “STIPS”, we shall use the mean output as the forecast of the models developed in this research (“Model Cluster” and “Model General”). The potential forecast capability of model is measured by the conventional percent variance explained (R^2), mean absolute error (MAE), and prediction standard error (SE, the standard deviation of forecast error) in this study. In specific, R^2 is positive oriented, which ranges from 0 to 1. The closer is R^2 to 1, the better skill a forecast model is. On the other hand, MAE and SE are very similar and both of them are negative-oriented and ranges from 0 to infinity. That is, the closer is MAE or SE to 0, the better a forecast model is. Due to the non-linearity characterized by TC intensity change, we deem MAE is a more robust statistics than SE as a forecast skill measure for this study.

TC Type	Model	Measure	6-hour	12-hour	24-hour	36-hour	48-hour
A	Model Cluster	R2 (%)	27.4	53.4	81.3	68.1	73.6
		MAE	1.57	2.15	2.70	4.65	4.83
		SE	2.23	2.87	3.50	6.90	6.14
	Model General	R2 (%)	31.1	45.8	61.6	60.4	46.6
		MAE	1.55	2.20	4.00	5.64	6.71
		SE	2.17	3.12	5.02	7.16	8.74
	CLIPER	MAE	1.54	2.89	6.28	9.19	10.10
		SE	2.59	4.16	8.05	11.35	11.79
B	Model Cluster	R2 (%)	28.7	39.5	57.2	72.1	80.4
		MAE	1.80	2.84	4.36	4.47	4.13
		SE	2.60	3.90	5.57	5.68	5.25
	Model General	R2 (%)	30.2	37.1	47.4	55.8	64.4
		MAE	1.77	2.9	4.82	5.71	5.85
		SE	2.57	3.97	6.19	7.22	7.13
	CLIPER	MAE	1.72	3.23	5.94	7.76	9.40
		SE	3.07	4.99	8.49	10.73	11.85
C	Model Cluster	R2 (%)	19.7	43.9	65.0	77.2	88.4
		MAE	2.14	3.08	4.17	4.08	3.38
		SE	2.95	4.02	5.68	5.8	4.74
	Model General	R2 (%)	30.8	46.1	66.5	80.2	87.0
		MAE	1.96	2.88	4.32	4.45	4.30
		SE	2.68	3.92	5.48	5.37	5.07
	CLIPER	MAE	1.88	3.59	7.15	9.87	11.74
		SE	3.21	5.34	9.45	12.02	13.85
D	Model Cluster	R2 (%)	21.5	45.7	61.3	72.7	80.2
		MAE	2.16	3.17	4.97	5.88	5.52
		SE	3.08	4.19	6.35	7.27	7.25
	Model General	R2 (%)	26.5	41.0	55.8	68.8	76.2
		MAE	2.07	3.38	5.59	6.54	6.55
		SE	2.95	4.36	6.82	8.00	8.20
	CLIPER	MAE	2.23	4.20	8.14	11.48	13.95
		SE	3.45	5.67	10.17	13.87	16.28
E	Model Cluster	R2 (%)	28.4	40.4	65.7	72.2	81.0
		MAE	1.70	2.53	3.25	3.47	3.63
		SE	2.41	3.46	4.20	4.53	4.90
	Model General	R2 (%)	32.5	36.9	47.6	48.3	55.0
		MAE	1.67	2.68	4.32	5.27	5.86
		SE	2.32	3.55	5.29	6.50	7.08
	CLIPER	MAE	1.45	2.69	4.90	6.32	9.03
		SE	2.82	4.44	7.14	8.54	11.18

Table 2: Static model forecast performance for each TC type. “MAE” denotes “mean absolute error”; “SE” denotes “standard error”, both in the unit “m/s”.

Based on the predictor selection procedure described in Section 3.3 and 3.4 (in significance level 0.01), we generate our regression model for each typhoon type respectively. For each simulation, we also generate another independent model, which adopts the same predictor selection procedure without using the cluster classification described in Section 3.1. For the model evaluation, we apply a strict LOOCV procedure for each of the 76 typhoons occurred in the western North Pacific (WNP) from 2008 to 2011. The results are provided in Table 2 (along with CLIPER). To be noted, due to the limited number of samples for some TC types because of the short time period of the GFS data (i.e., only 4-yr), we only provide up to 48 hours lead forecast for the "Model Cluster" (Table 2 and 3) to avoid the potential model evaluation bias. For 12-hr forecast, the "Model Cluster" generally has smaller MAE relative to the "Model General" and "CLIPER" (Table 2). This is most clearly for Type D which affects the Taiwan area the most (Fig. 1). For longer lead times (24- to 48-hr), the forecast skill of "Model Cluster" distinguishes itself even more from the other two benchmark systems.

Model	**Measure**	**6-hour**	**12-hour**	**24-hour**	**36-hour**	**48-hour**	**60-hour**	**72-hour**
Model Cluster	R2 (%)	24.8	43.3	62.6	72.9	81.4	NA	NA
	MAE	1.89	2.83	4.15	4.61	4.36	NA	NA
	SE	2.71	3.82	5.43	6.05	5.77	NA	NA
Model General	R2 (%)	29.4	39.8	53.3	62.1	69.8	78.3	78.3
	MAE	1.83	2.92	4.79	5.66	5.83	5.58	5.79
	SE	2.60	3.92	6.05	7.12	7.33	6.99	7.39
CLIPER	MAE	1.80	3.40	6.54	8.93	10.80	12.94	14.62
	SE	3.10	5.05	8.84	11.56	13.34	14.99	15.86
STIPS	R2 (%)	NA	40.0	49.4	54.6	57.7	59.6	61.2
	MAE	NA	2.88	4.78	6.22	7.56	8.75	9.57

Table 3: Static model forecast performance comparison over all TCs through 2008-2011. "MAE" denotes "mean absolute error"; "SE" denotes "standard error", both in the unit "m/s". The model performance measure of "STIPS" is from Table 6 of Knaff et al. (2005).

We also summarize the model performance statistics over all 76 typhoons in Table 3, in which we also include the performance of the benchmark forecast model STIPS developed in Knaff et al. (2005). The performance results in Table 3 can be further visualized in Figure 3, from which we can see that, for short time lead forecast (6-hr), the benchmark model "STIPS", the "Model General" and the proposed track pattern classification based "Model Cluster" deliver very similar performance. However, for longer lead times, the "Model Cluster" unanimously outperforms the "Model General". Throughout the lead time span we tested, "Model General" always performs better than the benchmark "STIPS" and much better than the simpler "CLIPER" approach. The similar performance comparison conclusion between "Model Cluster" and "Model General" can also be drawn from Table 2.

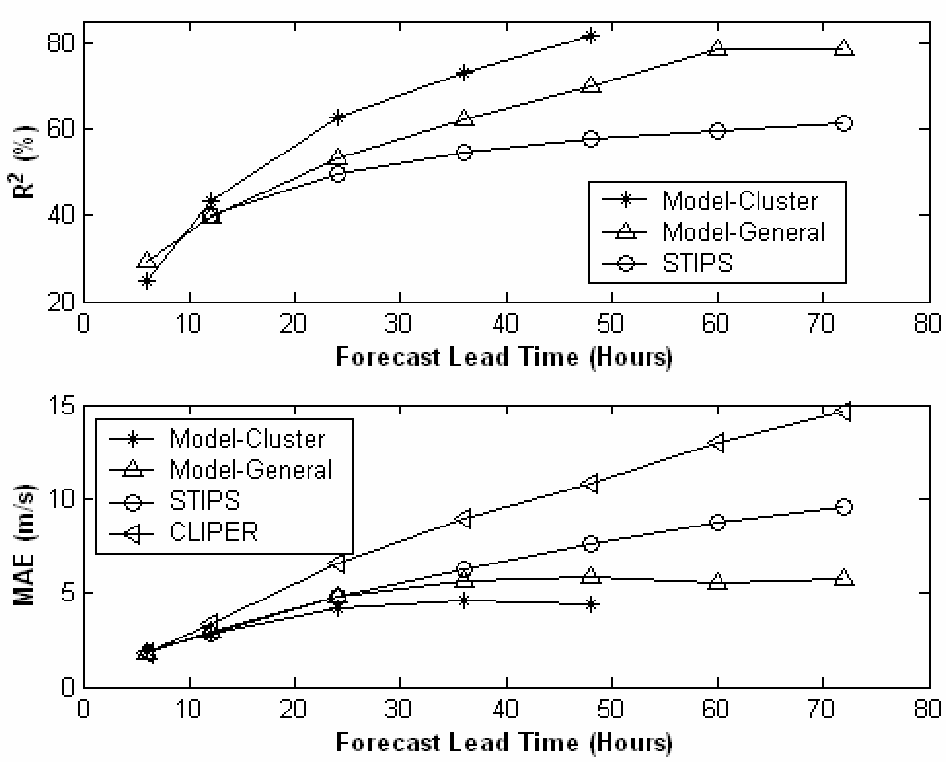

Fig. 3 Model performance comparison

TC Type	6-hour	12-hour	24-hour	36-hour	48-hour
A	MPI (0.45)	SHRG (-0.59)	SHRG (-0.62)	RHLO (-0.55)	VMAX(-0.60)
B	MPI (0.44)	LONG (0.49)	LONG (0.57)	LONG (0.56)	VMAX (-0.62)
C	MPI (0.50)	MPI2 (0.60)	MPI2 (0.67)	LAT (-0.74)	LAT (-0.78)
D	MPI2 (0.47)	MPI2 (0.54)	MPI2 (0.57)	MPI2 (0.57)	VMAX (-0.64)
E	MPI (0.51)	MPI (0.54)	MPI (0.58)	SHRG (-0.57)	VMAX (-0.64)

Table 4: The most significant predictor for each forecast model ("Model Cluster") in Table 2. The number after each predictor in the table is its respective correlation coefficient.

In Table 4, we provide a list of the most important (significant) predictors for each lead time forecast model of each track type. Not surprisingly, the SST related MPI and the vertical wind shear (SHRG) are almost always important for the forecast.

Three examples

In the following context, we shall provide three examples as the case study to illustrate the details of the proposed statistical forecast model, "Model Cluster". In all three examples, we shall respectively provide a probabilistic forecast. We also want to discuss a few key points of using this forecast model with the aid of the explicit examples. In the first step of each example, we assume that we have run the track pattern based clustering algorithm over all the typhoons occurred in the past (e.g., 2005-2011). Therefore, we have had the 0^{th}-order, the 1^{st}-order and the 2^{nd}-order coefficients (Eq. 1b) for each typhoon type. In this study, we adopt the results obtained in Chu et al. (2010b). The details of the key results are given in the Appendix and the detail procedure is described in Section 3.1. With the track coefficient of each TC typhoon, for a given TC track record, one can measure its distance to the mean or regressed track for each type. In the end, choose the cluster type that is closest to the track pattern of this TC.

Based on the coefficient table given in Table A-2, we decide Typhoon Melor in 2009 belongs to Type C. In an operational setting, the belonging of an on-going typhoon to a particular path type is not known a priori. However, users may first consult the predicted typhoon trajectory and compare it to the mean track of each typhoon track cluster (Fig. 1 and Table A-2) as a basis for decision. Following the procedure detailed in Section 3, we have the model for 6-hour lead forecast for this typhoon. And the result is given in Figure 4. In Fig. (4a), we first show the mean of the probabilistic forecast of the proposed model (solid line). Based on the distribution derived from (5c), we also provide the upper quartile (P(intensity) < 0.75) and the lower quartile (P(intensity) < 0.25) of the forecast (both in dotted lines). The distance between these two bounds shows the variation of forecast. Conceptually, the smaller the SE of a model forecast is, the better (or accurate) the model is. Similarly, the detailed forecast results for Typhoon Morakot are shown in Fig. 5 and the results for Typhoon Nanmadol are shown in Fig. 6.

In all three examples, we can see most of the true observed intensity measures are well bounded by their relative intensity forecast upper quartile and lower quartile (Fig. 4a, 5a, 6a). It's also worth noting that, in all Fig. 4b, 5b and 6b,, the forecast model indeed provides right intensity change directions in most of the time, however it does not perform well when the typhoon intensity has drastic change (say larger than 10 m/s in 6 hours). All these three examples illustrate a shortcoming of the proposed model. That is, a linear regression model is only optimum when observation noise is normally distributed.

However, the abrupt surge or drop of TC intensity behaves more like an extreme distribution with a very long tail. Also, the proposed model use a macro-scale (mostly within 200-800 km) area average, which would smooth out the regional weather signal. Nevertheless, the forecast skill of our approach is still better than the CLIPER (Table 5).

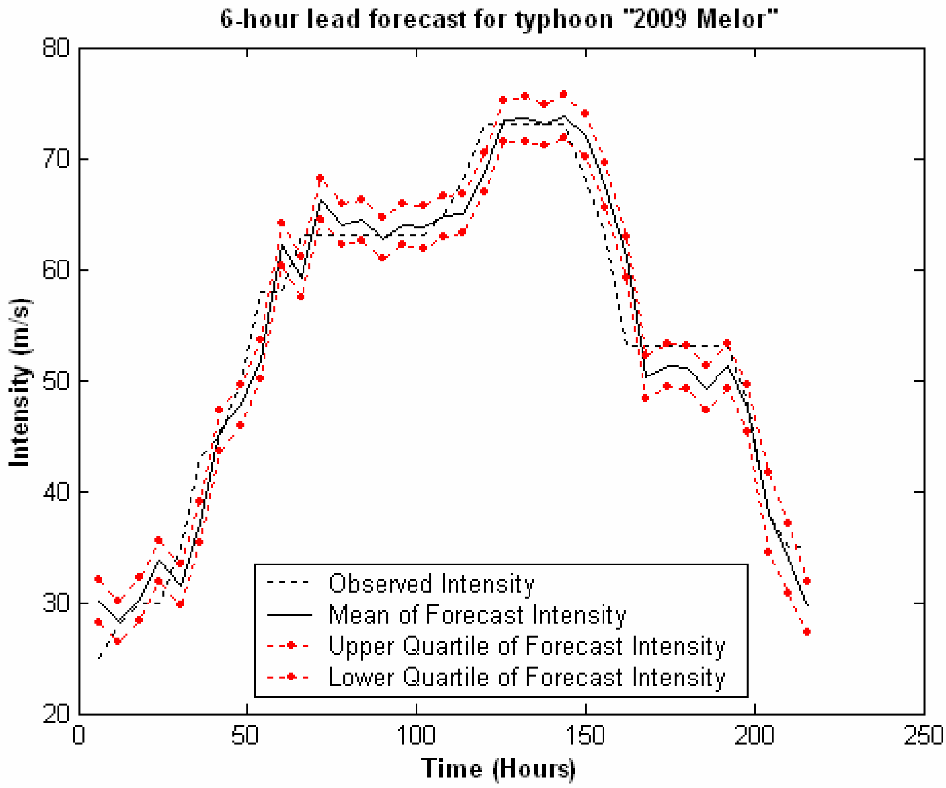

Fig. 4a: 6-hour lead forecast for the intensity of typhoon "2009 Melor." The upper and lower quartiles of the forecasts are indicated by red broken lines.

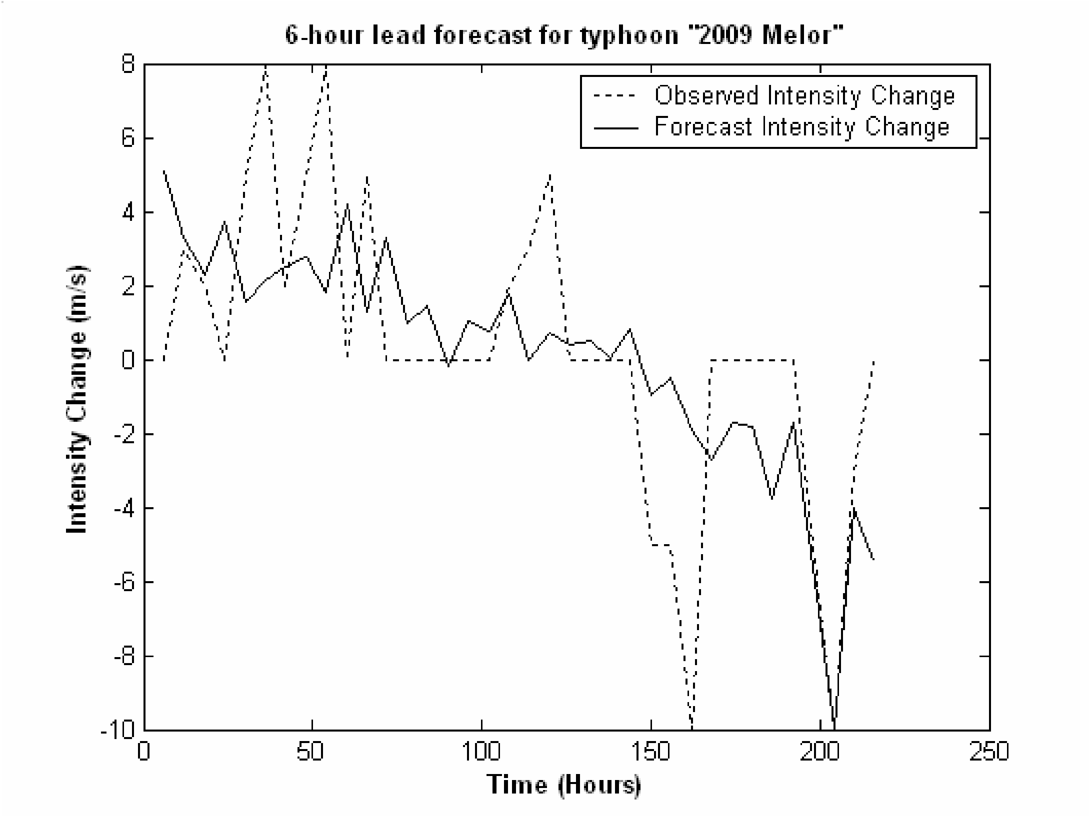

Fig. 4b: 6-hour lead forecast for the intensity change of typhoon "2009 Melor."

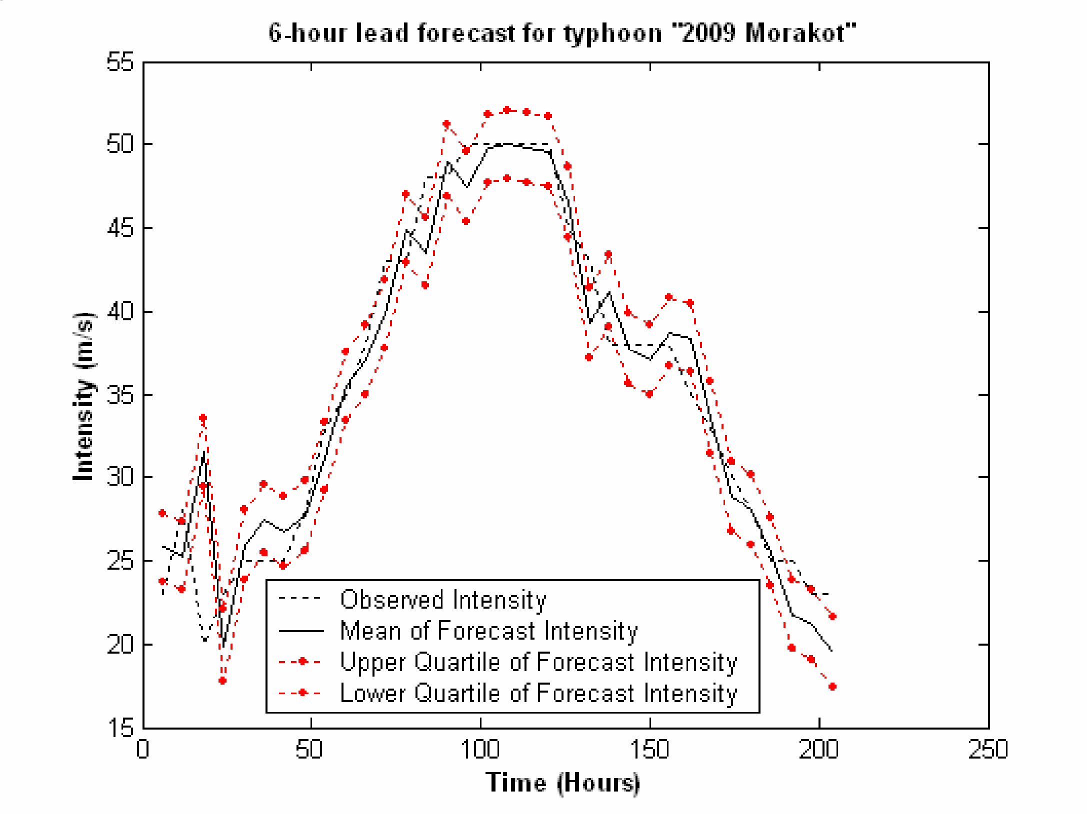

Fig. 5a: 6-hour lead forecast for the intensity of typhoon "2009 Morakot." The upper and lower quartiles of the forecasts are indicated by red broken lines.

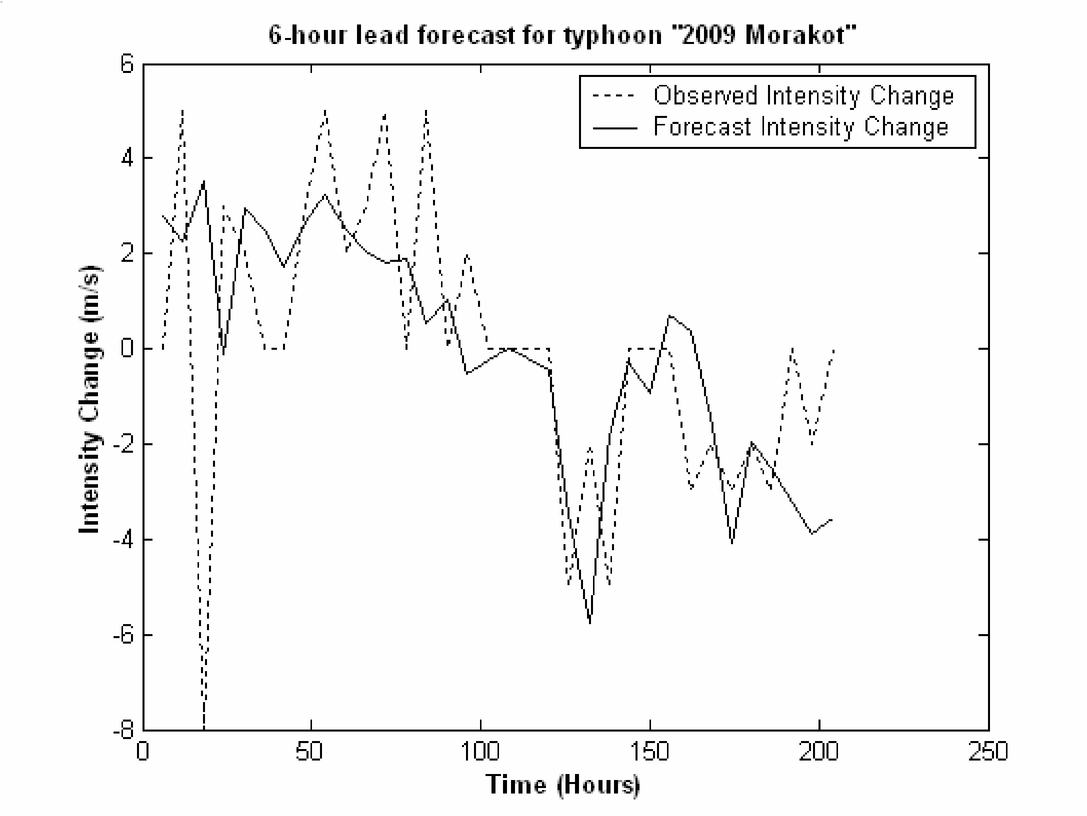

Fig. 5b: 6-hour lead forecast for the intensity change of typhoon “2009 Morakot.”

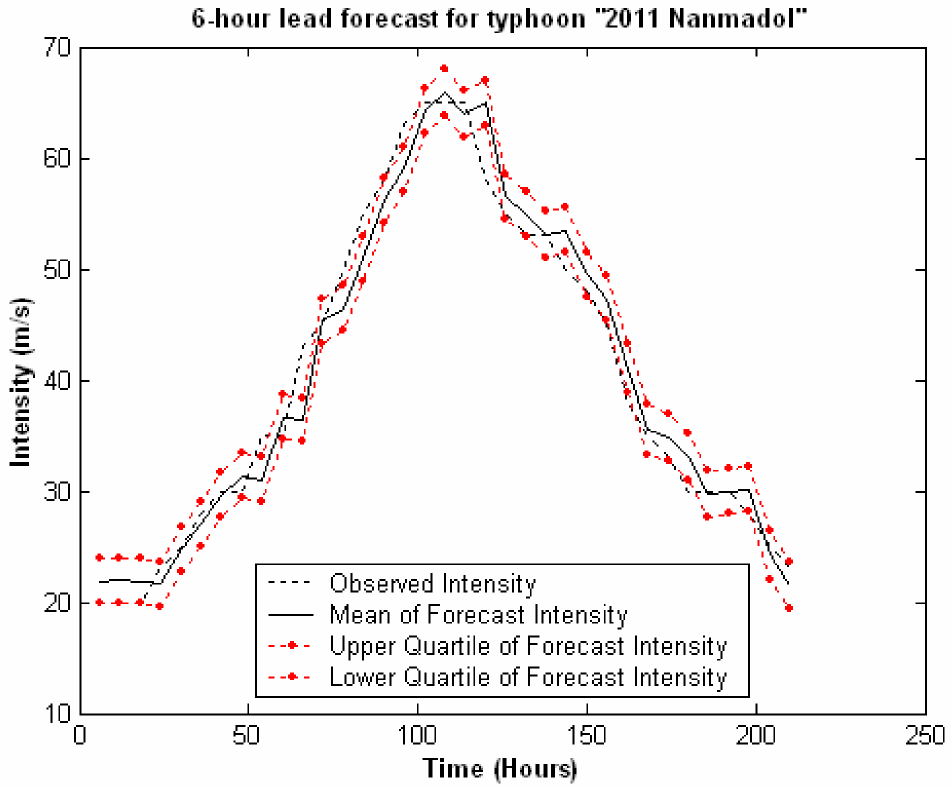

Fig. 6a: 6-hour lead forecast for the intensity of typhoon “2011 Nanmadol.” The upper and lower quartiles of the forecasts are indicated by red broken lines.

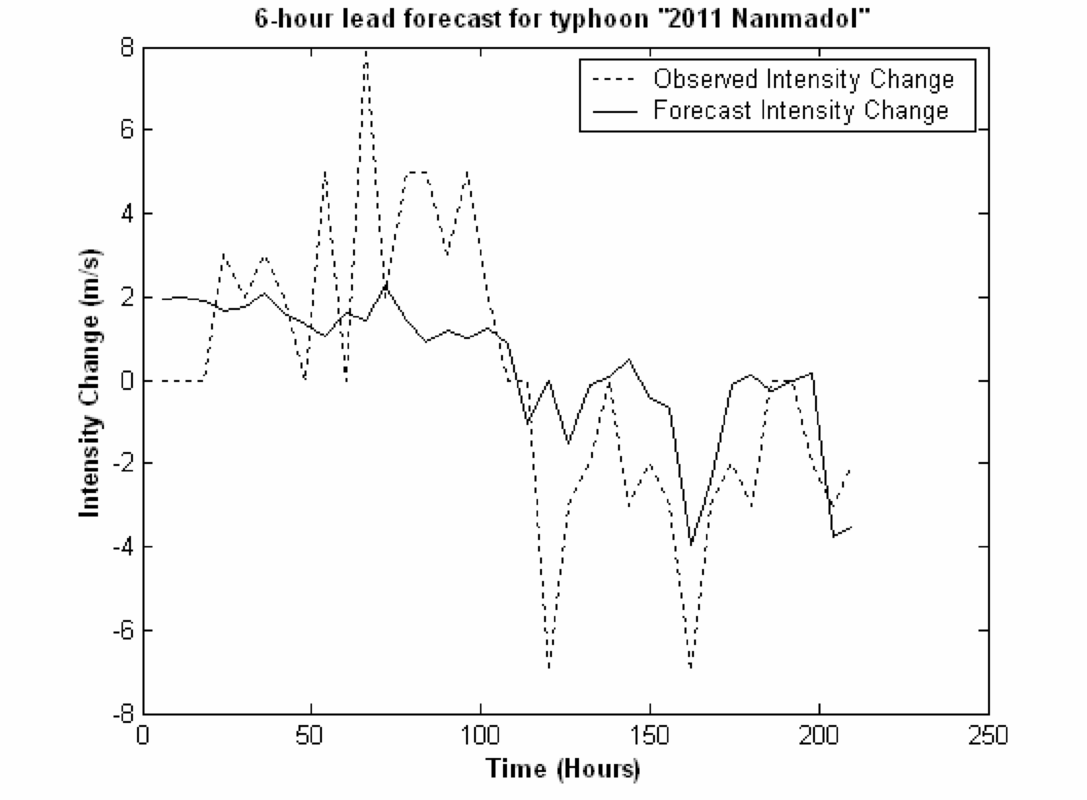

Fig. 6b: 6-hour lead forecast for the intensity change of typhoon “2011 Nanmadol.”

Typhoon Name	Model Cluster			CLIPER	
	R2 (%)	MAE	SE	MAE	SE
2009 Melor	35.6	2.44	3.23	2.65	3.93
2009 Morakot	19.8	2.01	2.92	2.06	2.99
2011 Nanmadol	40.7	1.98	2.61	2.49	3.34

Table 5: Summary of the performance measure for the case studies

In summary, we have developed a statistical/dynamical approach for forecasting typhoon intensity's probabilistic distribution. We have proven the mean output of the proposed model outperforms the benchmark forecast model CLIPER and STIPS. The model is featured with three highlights. The first one is that it is grounded on a track pattern based classification. Second, it involves both static predictors, such as current location, current intensity, and time dependent environmental variable related to temperature, wind, and moisture fields. Third, it provides a probabilistic forecast of the typhoon intensity, which provides a variation of the forecast rather than a single as do in most benchmark models.

With all the tests, we can see the proposed model always provides reliable intensity change direction. Due to the nature of linear regression model and our time dependent predictor selection procedure, the model does not perform particularly well for the extreme scenarios. In the future study, we shall explore the nonlinear regression models fitting extreme distribution such as the probit or logistic regression (Chu et al., 2010a) and polish up the time dependent predictor generation or selection procedure.

Appendix: Characteristic model parameter for Typhoon Type A, B, C, D and E

As follows, we provide the simulation results with running the clustering algorithm to the track records of all typhoons occurred in the vicinity of Taiwan from 1951 to 2009 (the data are from JTWC). Table 1 lists the basic statistics for each TC type and Table 2 provides the linear coefficients for each TC type as defined in Eq. (1b).

TC Type	A	B	C	D	E
Probability	0.1757	0.133	0.2106	0.2263	0.2544
Average Lifespan (Days)	3.4082	4.4912	5.3816	9.1173	8.3875
Average ACE	3.3025	6.1507	9.2638	22.9041	22.6494

Table A-1: The general statistics for each TC type.

Type	Coefficient	Latitude	Longitude
A	Beta_0	14.3107	118.8524
	Beta_1	0.7008	-2.4477
	Beta_2	-0.0192	0.1282
B	Beta_0	8.2155	142.2821
	Beta_1	2.0115	-4.8121
	Beta_2	-0.1019	0.158
C	Beta_0	10.8923	152.1922
	Beta_1	1.1943	-5.7349
	Beta_2	0.0577	0.3343
D	Beta_0	14.7819	134.485
	Beta_1	1.6428	-4.3084
	Beta_2	0.0288	0.4147
E	Beta_0	19.4289	141.9543
	Beta_1	3.1795	-0.5394
	Beta_2	-0.0367	0.2523

Table A-2: The coefficients for each TC type (refer to Eq. (1b))

References:

Camargo, S. J, A.W. Robertson, S.J. Gaffney, P. Smyth, and M. Ghil, 2007: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, 20, 3635-3653.

Chu, P.-S., X. Zhao, C.-H. Ho, H.-Y. Kim, and M.-M. Lu, 2010a: Bayesian Forecasting of Seasonal Typhoon Activity: A Track-Pattern-Oriented Categorization Approach for Taiwan, *J. Climate*, 23, 6654-6668.

Chu, P.-S., X. Zhao, and J.-H. Kim, 2010b: Regional typhoon activity as revealed by track Patterns and climate change. *Hurricanes and Climate Change.* Vol. **2**, J. Elsner et al., Eds., Springer, Chapter 8, 137–148.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin, 2004: Bayesian data analysis, 2nd Edition, Chapman & Hall/CRC, 668 pp.

Knaff, J.A., C.R. Sampson, and M. DeMaria, 2005: An Operational Statistical Typhoon Intensity Prediction Scheme for the Western North Pacific. *Weather and Forecasting*, 20, 699-699.