A STUDY OF EXTREME VALUE ANALYSIS ON TYPHOON WAVE

Tai-Wen Hsu[1,2] Muhajir Usman[2] Yuan-Jyh Lan[1] and Youe-Ping Lee[3]

Taiwan is a country with a higher number of typhoon events every year, an average of almost 4 typhoons per year attacking Taiwan based on the records of Central Weather Bureau (CWB) of Taiwan from 1958-2012, includes the year event, duration, strength level, path, pressure, wind speed, etc. Beside data from CWB, there are numerical simulation results of events by National Science and Technology Center for Disaster Reduction (NCDR) of Taiwan based on their own prediction system. However there are little has known about the relevancy between the simulated typhoon data from NCDR and recorded typhoon data from CWB. Therefore, this study will focus on comparing the typhoon data from CWB and NCDR by using extreme value analysis for several determined return periods. In general, this research generates the results (wave height) by using extreme value analysis (Gumbel distribution and Weibull distribution) with the help by numerical models (Rankine-Vortex model, Wind Wave Model). After getting the results, this research uses visual analysis and independent samples *t* test to investigate the results. Based on results analysis, it can be concluded that the simulated typhoon data from NCDR does not represent the recorded typhoon data from CWB very well.

Keywords: waves; typhoon events; wave height; extreme value analysis; independent sample t test

INTRODUCTION

Taiwan is a country with a higher number of typhoon events every year. This is caused by Taiwan is situated between the continental shelf of China and the open sea of the Pacific Ocean and it is often subject to severe sea states induced by typhoons generated during the summer in either the South China Sea or the Northwest Pacific Ocean near the Philippine islands, resulting in extensive loss of life and property (Ou et al. 2002). An averages of almost 4 typhoon per year attacking Taiwan based on the records of Central Weather Bureau (CWB) of Taiwan from 1958-2012. The information recorded includes the year event, duration, strength level, path, pressure, wind speed, etc.

Beside data from CWB, in Taiwan there are numerical simulation results of typhoon events by National Science and Technology Center for Disaster Reduction of Taiwan (NCDR) based on their own prediction system which is Quantitative Precipitation Estimation and Segregation Using Multiple Sensors (QPESUMS). QPESUMS is at total system integration incorporating data from multiple radars, numerical models, satellites, lightning and surface sensors. NCDR is Taiwan's government agency with one the purpose is to provide the government with comprehensive disaster analysis information for disaster prevention. The numerical model of NCDR is based on WRF model (Weather Research and Forecasting model) downscaling 20km wind field data (produced by Japanese Meteorological Research Institute, MRI) to 5km results. This research been generated the first edition of simulated typhoon events in 2012 for past (1979-2003) and future time (2015-2039 and 2075-2099).

However there are little has known about the relevancy between the simulated typhoon data from NCDR and recorded typhoon data from CWB. Therefore, this research will focus on comparing the typhoon data from CWB and NCDR by using extreme value analysis for determined return periods.

The sources of data for this research are from CWB and NCDR. The recorded typhoon data will use the data from CWB, while the simulated typhoon data will use the data from NCDR. For data set, this research will use 20 year data set from CWB and NCDR. The choice of 20 year data set is to create a similar year range data set from CWB and NCDR and to follow the description by Coastal Engineering Manual (2002).

Input variables for this research from the data are typhoon pressure and typhoon path line (longitude and latitude coordinate). According to CWB, typhoon paths with wind field having physically affected the Taiwan are divided into ten categories. Fig. 1 illustrates the statistics of those categorized paths from 1989 to 2011, which consist of five west bounds (paths 1, 2, 3, 4 and 5), four north bounds (paths 6, 7, 8 and 9) and one special path which is path 10. The path 10 termed as special path because its path did not fit to categorize as other nine's path, but its wind field affected the Taiwan.

[1] Research Center for Ocean Energy and Strategies, National Taiwan Ocean University, Keelung 202, TAIWAN (ROC)
[2] Dept. of Hydraulic and Ocean Engineering, National Cheng Kung University, 1 University Rd., Tainan, 701, TAIWAN (ROC)
[3] Water Resources Agency, Ministry of Economic Affairs, 501 Sec.2 Liming Rd., Taichung 408, TAIWAN (ROC)

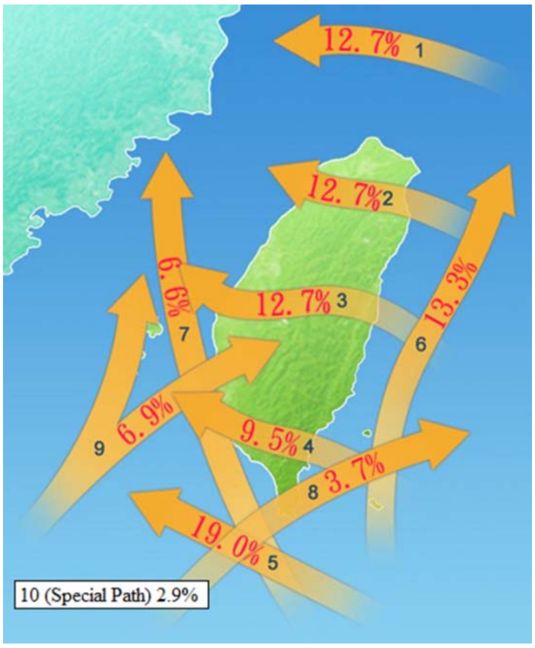

Figure 1. The categorized typhoon paths in Taiwan from 1989 to 2011 (Data source: adapted from CWB).

Meanwhile, the results after data analysis will be the wave height induced by typhoon's wind. According to Short Protection Manual (1984) wind-induced waves are surface waves that occur as the winds blow over the water surface. Wind-induced waves are generated in variety of sizes from ripples to large ocean waves as high as 30 meters.

In general, this research generates the results by using extreme value analysis with the help by numerical models. Distribution function use in extreme value analysis is Gumbel distribution or Weibull distribution, the selection of distribution function is based on evaluation parameters which are correlation coefficient (COR) and root mean square error (RMSE), the selected return periods are 5, 10, 25, 50, 100, 150, 200 and 250 year. Numerical models uses are Rankine-Vortex Model (RVM) to generate wind field, Wind Wave Model (WWM). Fig. 2 illustrates the flowchart of research approach.

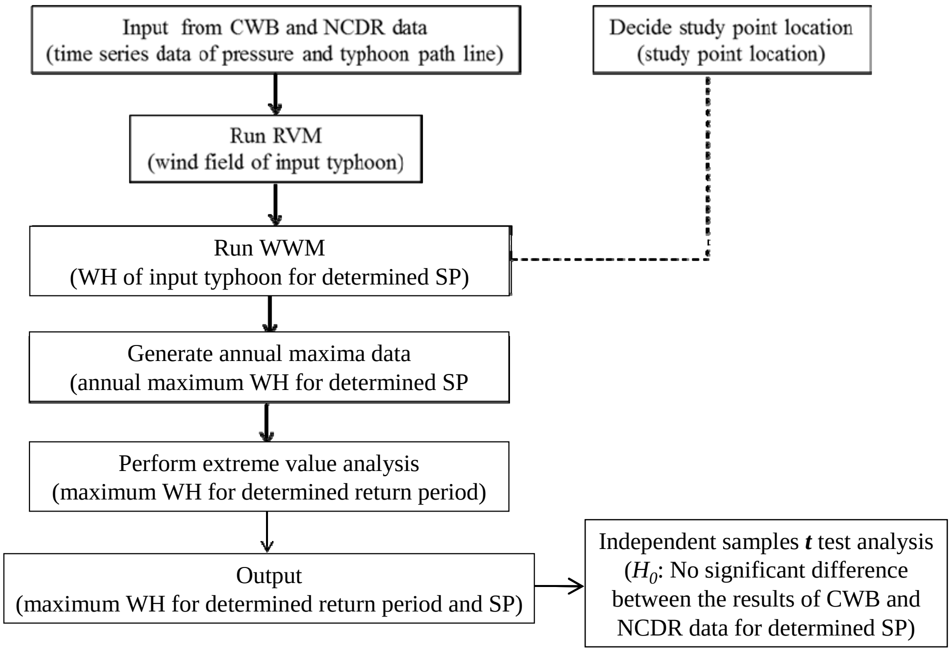

Figure 2. The flowchart of research approach.

After getting the results, this research will use visual analysis and independent samples *t* test to investigate the research objectives.

This research uses different year range for typhoon data set, where NCDR data uses data set with year range 1980-1990, while CWB data uses data set with year range 1990-2009. This is because the research approach uses time series data for the input, meanwhile typhoon data report from CWB before 1990 did not have time series data for pressure. Because of this reason, it decided to use the data set with year range 1990-2009 for this research. Fig. 3 shows the difference in CWB report page before and after year 1990.

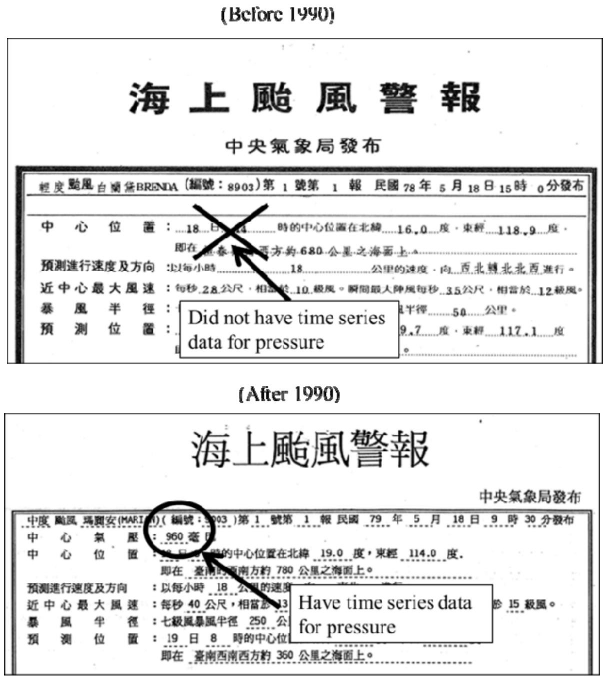

Figure 3. The different in CWB typhoon report page before and after 1990 (Data source: CWB).

In addition, for NCDR typhoon data, there are year where their typhoon prediction system estimate no typhoon risk affecting Taiwan. Because this research uses 20 years data set, so it was decided to substitute the year with no typhoon risk with another data. There are no typhoon risks for year 1982, 1983 and 1998. The data was substituted to year 1979, 2000 and 2001 respectively.

These limitations will not affect the results because Coastal Engineering Manual (2002) described when form an annual maximum series data, a data with record length of 20 years or more is needed without requirement the said data shall be in sequence. Furthermore, Goda (2000) described the individual data in a sample data set are independent to each other.

THEORETICAL BACKGROUNDS

(a) Extreme Value Analysis

Statistical method where limited length of data used to estimate the extreme event is known as extreme value statistics. Goda (2000) described extreme value statistics have the objective of estimating an expected value of extreme event which would occur once in a long period of time. For this purpose, the concept of return period is introduced. The relationship between return period of extreme event with it probability express in the following equation:

$$P = \frac{1}{T} \quad (1)$$

$$F(x) = 1 - P \quad (2)$$

where P is the probability of the extreme event occurs, T is the return period and $F(x)$ is the cumulative distribution where extreme event (x) not exceeded. Coastal Engineering Manual (2002) described there are six commonly used cumulative distribution function which are Gumbel distribution, Weibull distribution, Fisher-Tippet II distribution, Log-Normal distribution, Log Pearson Type III distribution and Pearson Type III distribution. Coastal Engineering Manual (2002) recommended for analysis of wave to use Gumbel distribution or Weibull distribution. Equation for Gumbel distribution and Weibull distribution summarized as follows:

1. Gumbel distribution

$$F(x) = e^{-e^{-\frac{x-B}{A}}}, -\infty < x < \infty \tag{3}$$

2. Weibull distribution

$$F(x) = 1 - e^{-\left(\frac{x-B}{A}\right)^k}, B \leq x < \infty \tag{4}$$

where x is stand for the extreme event, $F(x)$ is the cumulative probability function, A is the scale parameter, B is location parameter, k is shape parameter. Table 1 show the characteristics of distribution function used to determine the value of B, A, and k (Goda 2000).

Table 1. Characteristics of distribution function for extreme analysis.

Dist. Function	Mode	Mean	Standard Deviation
Gumbel	B	$B + \gamma A$	$\frac{\pi A}{\sqrt{6}}$
Weibull	$B + A\left(1 - \frac{1}{k}\right)^{1/k}, k > 1$	$B + A\Gamma\left(1 + \frac{1}{k}\right)$	$A\left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right]^{1/2}$

Note: $\Gamma(\,)$ is the gamma function and γ is Euler's constant ($= 0.5772\,...$).

Goda (2000) described there are several methods of fitting a theoretical distribution function to a sample of extreme data. They are graphical fitting method, least square method, method of moments, maximum likehood method, and few others. The least square method is recommended to use because of its simplicity in algorithm and applications. The unbiased plotting position equation can be expressed in the following general form:

$$Fm = 1 - \frac{m-\alpha}{N+\beta} \qquad ; m = 1,2, \dots, N \tag{5}$$

Eq. 5 is derived as the expected probability of the mth ordered variate in the population. The value of constant α and β are given in Table 2.

Table 2. Characteristics of distribution function for extreme analysis.

Dist. Function	α	β
Gumbel	0.44	0.12
Weibull	$0.20 + 0.27/\sqrt{k}$	$0.20 + 0.23/\sqrt{k}$

There are various evaluation factor can be used as parameter when selecting the distribution function. Two commonly used parameters are root mean square error and correlation coefficient. Equation for root mean square error is as follows

$$MSE = \frac{1}{n}\sum(F_i - F_m)^2 \tag{6}$$

$$RMSE = \sqrt{MSE} \tag{7}$$

where MSE is mean square error, Fi is cumulative probability from sample data, and n is number of population. Equation for correlation coefficient can be expressed as

$$COR = \frac{\sum(cov\, F_i)(cov\, F_m)}{\sum \sigma F_i \sum \sigma F_m} \tag{8}$$

where cov is stand for covariance, and σ is notation for standard deviation.

(b) Typhoon Wind Field

Ou et al. (2002) described typhoon wind fields are usually intense, spatially inhomogeneous and directionally varying. The large gradients in wind speed and the rapidly varying wind directions of the typhoon vortex can generate very complex ocean wave fields, but for practical applications the wind fields are always represented in terms of relatively simple parametric models. They used modified Rankin-Vortex Model (Holland 1980) to generate typhoon wind field, and its pressure distribution form is assumed to be an exponential relation that can be expressed as (Myers 1954)

$$P_r = P_0 + \Delta P e^{-r_0/r} \tag{9}$$

where P_0 is the central low pressure, ΔP is the atmospheric pressure depression ($\Delta P = P_\infty - P_0 = 1013.3 - P0$, $P\infty$ is the atmospheric pressure in the far field of the typhoon and has a theoretical constant value of 1013.3 hPa, P_r is the ambient atmospheric pressure at radius *r* from the low-pressure center, and r_0 is the radius of the maximum wind speed. Following Graham and Nunn (1959), r_0 is represented by

$$r_0 = 28.52\, tanh[0.0873(\phi - 28)] + 12.22/exp\left[\frac{(1013.3 - P_0)}{33.86}\right] + 0.2V_f + 37.22 \tag{10}$$

where ϕ is the latitude while V_f is the typhoon's forward speed. Following Eq. 9, the pressure gradient induces the cyclostrophic wind U_r at a distance *r* from low-pressure center, which can be expressed as

$$U_r^2 = \frac{1}{\rho_a} r \frac{dP_r}{dr} = \frac{1}{\rho_a} \Delta P \left(\frac{r_0}{r}\right) e^{-r_0/r} \tag{11}$$

where ρ_a is the density of air. The relationship between the cyclostrophic wind and the gradient wind yields:

$$V_r^2 + frV_r = \frac{1}{\rho_a} r \frac{dP_r}{dr} \tag{12}$$

$$V_r^2 + frV_r = U_r^2 \tag{13}$$

$$V_r = -0.5fr + \sqrt{(0.5fr)^2 + U_r^2} \tag{14}$$

where V_r is the gradient wind speed at position *r* from a typhoon's low-pressure center, *f* is the Coriolis parameter ($f = 2\omega sin\phi = 0.525 sin\phi$). The parameters ω and ϕ respectively represent the angular velocity of the earth's rotation and the latitude. The gradient wind velocity V_r can be reduced to V_{rs} (at 10 m above the sea level) by the application of a factor of 0.8 (Powell 1980) as

$$V_{rs} = 0.8V_r \tag{15}$$

Adding the typhoon's forward speed V_f to the sea wind the corrected velocity of the typhoon yields

$$V_{rs}^* = V_{rs} + 0.5V_f cos\theta_c \tag{16}$$

where V_{rs}^* is the composite wind velocity by the typhoon forward speed and the gradient wind speed, and θ_c is the angle between the composite wind velocity and the typhoon forward velocity. Fig. 4 demonstrates the relationship between V_{rs}^* and V_f, where point O is the center of the typhoon, OM is the radial direction of the maximum wind speed, α is the included angle of the composite wind velocity and the isobar of 25^o (Shea and Gray 1973), δ is the included angle between the typhoon forward velocity and the *x*-coordinate, and γ is the included angle between OG and the *x*-coordinate. Thus, θ_c, can be expressed by $\theta_c = 90^o + \alpha - \delta + \gamma$.

(c) Wind Wave Model

Hsu et al. (2011) explained in the Wind Wave Model, the evolution of the wave spectrum is described by the spectral action balance equation, which is expressed in Cartesian coordinates as follows (Hasselmann et al., 1973):

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x}(C_x N) + \frac{\partial}{\partial y}(C_y N) + \frac{\partial}{\partial \sigma}(C_\sigma N) + \frac{\partial}{\partial \theta}(C_\theta N) = S_{total} \tag{17}$$

where $N = N(t, x, y, \sigma, \theta)$ is the action density function, *t* is the time, $C_x, C_y, C_\sigma, C_\theta$ are the propagating velocities in the *x*, *y*, σ and θspaces respectively, σ is the relative frequency, θ is the wave propagation direction. The term $S_{total} = S_{tot}/\sigma$ on the right-hand side of the equation is the source term, which is in terms of energy density, representing the combined effects of energy generation S_{in}, dissipation S_{ds}, bottom friction S_{bs} and also non-linear wave interactions S_{nl}. S_{tot} is the spectral density of the wave energy, they adopted the formulation for the energy generation and dissipation proposed by Makin and Kudryavtsev (1999). The discrete interaction approximation (Hasselmann et al. 1985) scheme is

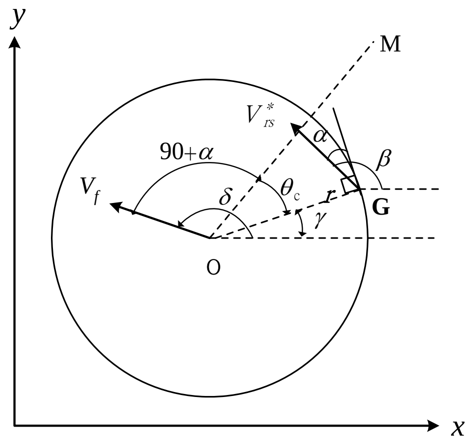

Figure 4. A definitions sketch of the velocity fields for a moving cyclone (Data source: Ou et al. (2002)).

applied for the description of the non-linear term S_{nl}, and the bottom friction term is in terms of the empirical formula proposed by Hasselmann et al. (1973).

For application to large-scale oceanic regions, Eq. 17 needs to be integrated in terms of the spherical coordinates

$$\frac{\partial \hat{N}}{\partial t} + (cos\phi)^{-1}\frac{\partial(\dot{\phi}cos\phi\hat{N})}{\partial\phi} + \frac{\partial(\dot{\lambda}\hat{N})}{\partial\lambda} + \frac{\partial(\dot{\sigma}\hat{N})}{\partial\sigma} + \frac{\partial(\dot{\theta}\hat{N})}{\partial\theta} = S_{total} \quad (18)$$

where $\hat{N} = \hat{N}(\lambda, \phi, \sigma, \theta, t)$ is the wave action density spectrum for spherical coordinates, ϕ is the latitude, λ is the longitude and $\dot{\phi}$, $\dot{\lambda}$, $\dot{\sigma}$ and $\dot{\theta}$ denote the time rates of change of ϕ, λ, σ and θ respectively. Equation for $\hat{N}$, $\dot{\phi}$, $\dot{\lambda}$, $\dot{\sigma}$ and $\dot{\theta}$ are given by the following equations:

$$\hat{N} = NR^2cos\phi \quad (19)$$

$$\dot{\phi} = \left(C_g cos\theta + \boldsymbol{U}/north\right)R^{-1} \quad (20)$$

$$\dot{\lambda} = \left(C_g sin\theta + \boldsymbol{U}/east\right)(Rcos\theta)^{-1} \quad (21)$$

$$\dot{\theta} = C_g sin\theta\ tan\phi\ R^{-1} + (\boldsymbol{k}x\boldsymbol{k})k^{-2} \quad (22)$$

$$\dot{\sigma} = \partial\sigma/\partial t \quad (23)$$

where R is the radius of the earth, C_g is the group velocity, $\boldsymbol{U}$ is the current velocity vector, $\boldsymbol{k}$ is the wavenumber vector and $k = |\boldsymbol{k}|$ is the wavenumber. In Eqs. 20 and 21, north and east represent latitude and longitude of the earth respectively. Eq. 18 will be used as the transport equation of the WWM. Detail of the numerical scheme and boundary conditions in WWM are given in Hsu et al. (2005). For the verification of the WWM, Hsu et al. (2005) described the numerical results obtained from the WWM are in fairly good agreement with measurements.

Fig. 5 shows the mesh configuration of the WWM boundary condition use in this research. It contains 2454 nodes and 4542 elements.

(d) Measuring Statistical Significance of Independent Samples

Buckingham and Saunders (2004) described independent samples t test is used to test the significance of a difference between the mean values recorded for two samples on a dependent variable. Assumptions for independent samples t test:

1. The two samples are independent of one another.
2. The dependent variable is normally distributed.
3. It requires that the variance within each of the groups being compared should not be too dissimilar.

Jackson (2006) expressed the equation for determining *t* is

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (24)$$

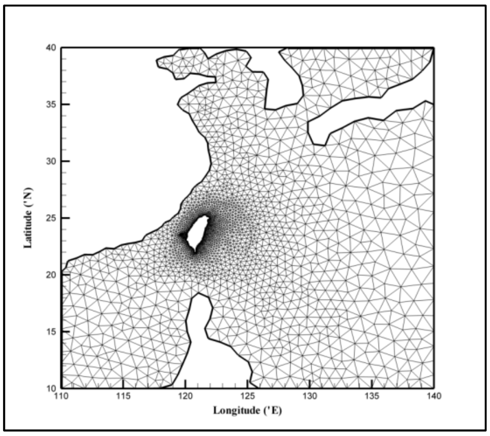

Figure 5. Mesh configurations of WWM.

where t is the t value obtained from calculation, $\overline{X}_1$ and $\overline{X}_2$ are the mean of the two samples, s_1^2 and s_2^2 are the variances of the two samples, n_1 and n_2 are the number of data in each sample.

Before determine the t test value, there are three parameters need to decide first which are the hypothesis direction, degree of freedom and $\propto$ level.

There are two ways in forming hypothesis for t test. Jackson (2006) described the first one is called a directional hypothesis. In this case, the researcher predicted the direction of the difference. Thus, the hypothesis for directional test might more appropriately be written as

$$H_0: \mu_1 \leq \mu_2 \quad ; \quad H_a: \mu_1 > \mu_2 \qquad (25)$$

The other way is called non-directional hypothesis, in which the researcher expects to find differences between the groups but is unsure what the difference will be. The statistical notation for non-directional test is

$$H_0: \mu_1 = \mu_2 \quad ; \quad H_a: \mu_1 \neq \mu_2 \qquad (26)$$

where H_0 is the hypothesis that will be tested and H_a is alternative hypothesis that will be accepted if the H_0 rejected, μ_1 and μ_1 are the population mean of data 1 and data 2.

One of the key concepts in hypothesis testing of statistical significance test is determine the $\propto$ level which defined as the probability rejecting hypothesis when in fact it was true, this is same as controlling the Type I error to a specified level. Jackson (2006) explained the Type I error occur when the hypothesis is rejected when in reality the hypothesis is true. There is another error that called as Type II error, this type of error occur when not rejecting the hypothesis when it is false. Fisher (1990) explained based on past experience, he suggested that a $\propto$ level of 0.05 is a good compromise between the likelihoods of making Type I and Type II errors. Anytime a decision made using statistics, there are four possible outcomes. Two of the outcomes represent correct decisions, whereas two represent errors. Table 3 adapted from Kreyszig (1999) to illustrates the possible outcomes of decision taking use statistical method.

Table 3. Four possible outcomes in statistical decision making.

	Unknown truth to the researcher	
	H₀ is True	*H₀* is False
Reject H_0	Type I error	True decision
Accept H_0	True decision	Type II error

Next on is about the degrees of freedom. Iversen and Norpoth (1987) explained the degrees of freedom appear any time when compute a sum of squares. One way to define the concept is to say that the degrees of freedom for a particular sum of squares is equal to the smallest number of term in the sum need to know in order to find the remaining terms and thereby compute the sum. Jackson (2006) expressed the equation for determining degree of freedom for independent samples *t* test is

$$d_f = n_1 + n_2 - 2 \qquad (27)$$

Degrees of freedom use to decide the critical value of t (t_{crit}) for selected ∝ level to determine rejection region where the hypothesis in not true. Fig. 6 illustrates hypothesis testing related to t_{crit}.

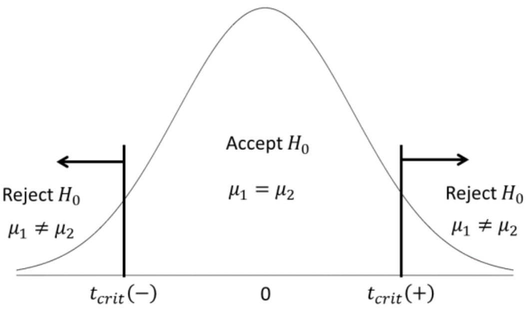

Figure 6. Hypothesis testing related to the critical value of *t* (t_{crit}) for non-directional hypothesis.

Jackson (2006) reproduced critical value of t for Student's t distribution with the n-degree of freedom shows in Table 4. The t value from Eq. 24 then compared with t_{crit} to determine if the hypothesis is accepted or rejected.

Table 4. Critical value of *t* for the Student's *t* Distribution.

d_f	Level of Significance for Directional Test (t_{crit})					
	∝=0.10	∝=0.05	∝=0.025	∝=0.01	∝=0.005	∝=0.0005
	Level of Significance for Non-directional Test (t_{crit})					
	∝=0.20	∝=0.10	∝=0.05	∝=0.02	∝=0.01	∝=0.001
1	3.0780	6.3140	12.7100	31.8200	63.6600	636.6000
2	1.8860	2.9200	4.3030	6.9650	9.9250	31.6000
3	1.6380	2.3530	3.1820	4.5410	5.8410	12.9200
4	1.5330	2.1320	2.7760	3.7470	4.6040	8.6100
5	1.4760	2.0150	2.5710	3.3650	4.0320	6.8690
6	1.4400	1.9430	2.4470	3.1430	3.7070	5.9590
7	1.4150	1.8950	2.3650	2.9980	3.4990	5.4080
8	1.3970	1.8600	2.3060	2.8960	3.3550	5.0410
9	1.3830	1.8330	2.2620	2.8210	3.2500	4.7810
10	1.3720	1.8120	2.2280	2.7640	3.1690	4.5870
11	1.3630	1.7960	2.2010	2.7180	3.1060	4.4370
12	1.3560	1.7820	2.1790	2.6810	3.0550	4.3180
13	1.3500	1.7710	2.1600	2.6500	3.0120	4.2210
14	1.3450	1.7610	2.1450	2.6240	2.9770	4.1400
15	1.3410	1.7530	2.1310	2.6020	2.9470	4.0730

Table 5. Study Points Information (Data source: extracted from Google Earth).

Study Point	Location Coordinate		Approx. Water Depth (m)	Approx. Distance from Shoreline (Km)
	Longitude (˚E)	Latitude (˚N)		
SP1	121.9855	25.1061	110	6.0
SP2	121.0237	25.0868	50	5.0
SP3	120.8179	24.7667	40	5.0
SP4	120.4989	24.3702	50	5.0
SP5	120.2100	23.9899	30	10.0
SP6	119.9716	23.0986	75	5.0
SP7	120.2065	22.5982	75	4.5
SP8	120.8953	21.9705	280	5.0
SP9	121.1781	22.6910	120	5.0
SP10	121.4526	23.1562	280	5.0
SP11	121.6771	24.0534	350	5.0
SP12	121.9072	24.6407	225	5.0

DATA ANALYSIS

Study points are determined before data analysis. The purpose of study points is to be the location where the wave height results from numerical models will be investigated. There are twelve study points selected which spread evenly between Western and Eastern area of Taiwan. The study points information shown in Table 5.

The study points are selected to cover all of Taiwan area with the distance from the shoreline are approximately 5.0 to 10.0 Km. Fig. 7 illustrates the location of selected study points.

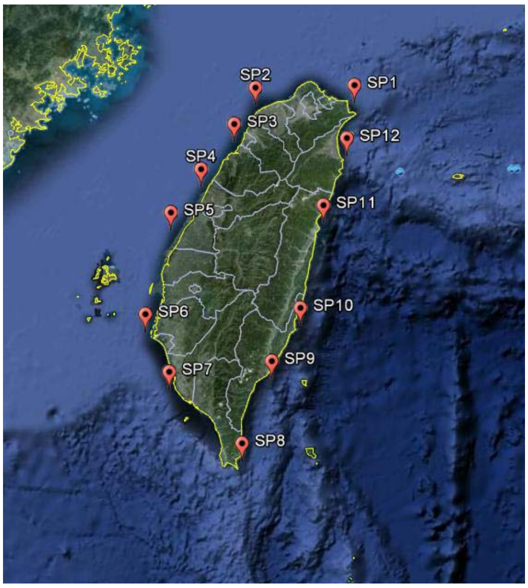

Figure 7. Study points location (Data source: adapted from Google Earth).

The input data for this research are time series data of typhoon pressure and typhoon path line from CWB and NCDR. Not all typhoon data is taken into consideration, only the typhoons with wind field having physically affected the Taiwan are selected.

The first step is running the numerical models by using time series data of typhoon pressure and typhoon path line. Run RVM to generate the wind field, after that run WWM to generate the wave height for selected study points. The results are the wave height of the typhoon. Then, generate the annual maximum data of wave height for selected study points.

The next step is performing extreme value analysis by using the annual maximum data of wave height for selected study points. The results are wave height for selected study points and determined return periods. Then compare the evaluation parameters RMSE and COR. Table 6 illustrates the evaluation parameters of CWB and NCDR data for wave height from Gumbel distribution and Weibull distribution.

Table 6. Evaluation parameters results of extreme value analysis for wave height.

a. CWB 1990-2009							b. NCDR 1980-1999					
Study Point	Gumbel Dist.		Study Point	Weibull Dist.			Study Point	Gumbel Dist.		Study Point	Weibull Dist.	
	RMSE	COR		RMSE	COR			RMSE	COR		RMSE	COR
SP1	0.0653	0.9808	SP1	0.0459	0.9877		SP1	0.0363	0.9920	SP1	0.0374	0.9913
SP2	0.0727	0.9781	SP2	0.0402	0.9911		SP2	0.0800	0.9662	SP2	0.0590	0.9828
SP3	0.0854	0.9686	SP3	0.0419	0.9897		SP3	0.0768	0.9685	SP3	0.0576	0.9825
SP4	0.0774	0.9741	SP4	0.0351	0.9926		SP4	0.0568	0.9813	SP4	0.0530	0.9831
SP5	0.0771	0.9773	SP5	0.0500	0.9875		SP5	0.0542	0.9883	SP5	0.0253	0.9964
SP6	0.0509	0.9850	SP6	0.0485	0.9857		SP6	0.0542	0.9887	SP6	0.0261	0.9960
SP7	0.0667	0.9801	SP7	0.0609	0.9795		SP7	0.0786	0.9721	SP7	0.0441	0.9879
SP8	0.0662	0.9826	SP8	0.0591	0.9821		SP8	0.0411	0.9925	SP8	0.0262	0.9958
SP9	0.0685	0.9782	SP9	0.0453	0.9873		SP9	0.0821	0.9717	SP9	0.0575	0.9815
SP10	0.0648	0.9817	SP10	0.0305	0.9942		SP10	0.0590	0.9875	SP10	0.0391	0.9926
SP11	0.0653	0.9838	SP11	0.0435	0.9904		SP11	0.0730	0.9768	SP11	0.0400	0.9905
SP12	0.0941	0.9613	SP12	0.0553	0.9820		SP12	0.0452	0.9902	SP12	0.0322	0.9936

Based on Table 6, more than 90% of evaluation parameters results shows the Weibull distribution are better than Gumbel distribution if compare between each study point. So the results from Weibull distribution are selected. In the end, the results were the maximum wave height for selected study points and determined return periods. Table 7 and Fig. 8 represent the wave height results.

Table 7. Wave height results.

Study Point	Wave height result CWB 1990-2009 (m)							
	5	10	25	50	100	150	200	250
SP1	7.27	8.39	9.55	10.28	10.92	11.26	11.50	11.67
SP2	4.86	5.58	6.32	6.77	7.17	7.38	7.52	7.63
SP3	4.69	5.33	5.98	6.39	6.74	6.92	7.05	7.14
SP4	4.53	5.11	5.70	6.07	6.39	6.56	6.67	6.76
SP5	4.27	4.74	5.22	5.52	5.78	5.92	6.01	6.08
SP6	4.54	5.13	5.82	6.30	6.74	6.99	7.17	7.30
SP7	4.12	4.56	5.05	5.37	5.66	5.82	5.93	6.01
SP8	5.75	6.58	7.46	8.03	8.53	8.80	8.98	9.12
SP9	5.88	6.51	7.15	7.54	7.89	8.07	8.20	8.29
SP10	6.76	7.52	8.29	8.77	9.18	9.41	9.55	9.67
SP11	8.24	9.49	10.78	11.59	12.31	12.70	12.96	13.15
SP12	7.38	8.42	9.49	10.16	10.75	11.06	11.27	11.43
Study Point	Wave height result NCDR 1980-1999 (m)							
	5	10	25	50	100	150	200	250
SP1	5.55	7.08	8.88	10.13	11.31	11.97	12.43	12.78
SP2	3.63	4.94	6.61	7.83	9.02	9.71	10.20	10.57
SP3	3.27	4.37	5.75	6.76	7.74	8.31	8.70	9.01
SP4	3.31	4.15	5.13	5.81	6.46	6.82	7.07	7.26
SP5	3.62	4.14	4.68	5.03	5.34	5.50	5.62	5.70
SP6	5.24	5.83	6.42	6.79	7.11	7.28	7.39	7.48
SP7	5.52	6.14	6.77	7.15	7.48	7.65	7.77	7.86
SP8	7.05	8.46	9.99	11.00	11.92	12.42	12.77	13.03
SP9	7.68	8.96	10.29	11.14	11.89	12.30	12.57	12.77
SP10	7.49	8.70	9.95	10.74	11.44	11.82	12.07	12.26
SP11	5.85	6.65	7.47	7.98	8.42	8.66	8.82	8.94
SP12	5.77	6.88	8.08	8.88	9.59	9.98	10.25	10.45

Based on visual analysis of Fig. 8, there are differences in wave height results investigated by comparing the results based on the source of data. For this research, the results of CWB and NCDR are quite different. It is estimated because the inputs are not controlled, the inputs for this research are time series data of typhoon path line and pressure.

To further study the differences between the results, this research is conducting the independent samples *t* test to investigate the results based on statistical method that will be explained in next the part of the report.

RESULTS AND DISCUSSION

The main objective of this research is to investigate the relevancy of the simulated typhoon data from NCDR to recorded typhoon data from CWB. From visual analysis, the wave height results from data analysis show there is a difference.

To investigate if the difference is significance according to statistical method, this research is conducting the *t* test analysis of the results based on the source of data. The results being compared are from CWB 1990-2009 data and NCDR 1980-1999 data. Calculation of the wave height results for study point 2 shown as an example (Table 8).

There are 8 data for each sample. The next step is defining the hypothesis (H_0) and alternative hypothesis (H_a) which are:

H_0 : No significant difference between the SP2 wave height results of CWB 1990-2009 and NCDR 1980-1999 ($\mu_1 = \mu_2$).

H_a : There is significant difference between the SP2 wave height results of CWB 1990-2009 and NCDR 1980-1999 ($\mu_1 \neq \mu_2$).

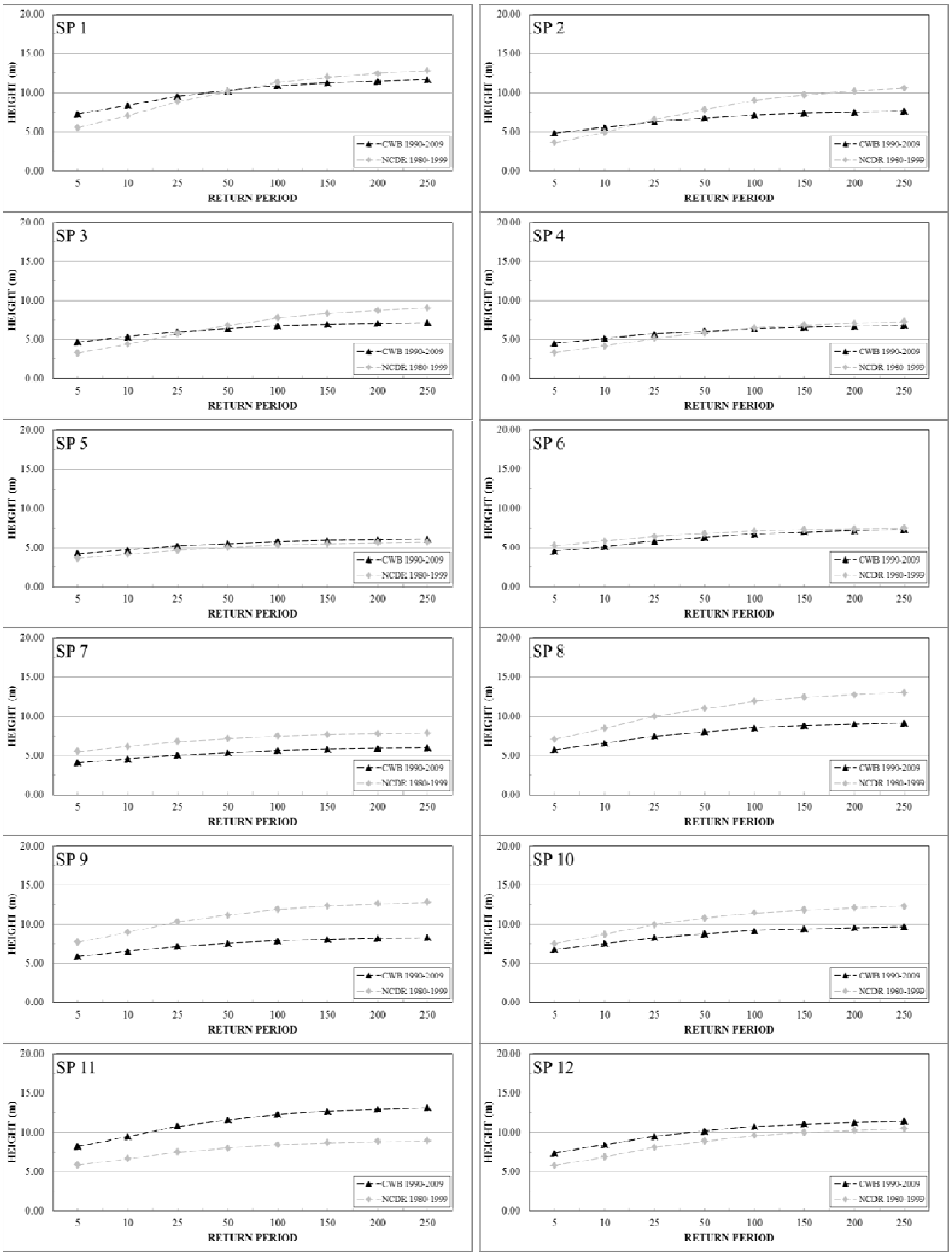

Figure 8. Wave height comparison for SP1 - SP12.

Table 8. SP2 sample data of CWB 1990-2009 and NCDR 1980-1999.

Sample	Data	Wave height result SP2 (m)							
		5	10	25	50	100	150	200	250
1	CWB 1990-2009	4.86	5.58	6.32	6.77	7.17	7.38	7.52	7.63
2	NCDR 1980-1999	3.63	4.94	6.61	7.83	9.02	9.71	10.20	10.57

The hypothesis is non-directional hypothesis. The degrees of freedom is 14 $\left(d_f = n_1 + n_2 - 2 = 14\right)$. The $\propto$ level selected is 0.05, so the value of t_{crit} based on Table 4 is ± 2.145. From the calculation, the t_{data} value is -1.1972. So the t_{data} falls under the t_{crit}. Thus, the H_0 is accepted, while H_a is rejected. Fig. 9 represent the obtained t_{data} in relation to the t_{crit}.

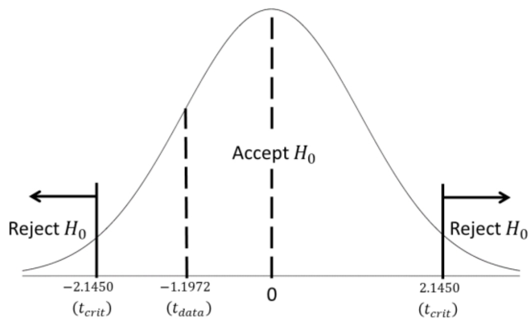

Figure 9. The obtained t_{data} in relation to the t_{crit} from SP2 wave height results analysis.

For the other study points, the calculation processes are similar. Table 9 illustrates the summary of *t* test for the rest results comparison.

Table 9. Summary of *t* test from wave height results comparison between CWB 1990-2009 and NCDR 1980-1999 data.

H_0 Statement	Study Point	t_{data}	t_{crit}	H_0
No significant difference between the wave height results of CWB 1999-2009 and NCDE 1980-1999 ($\mu_1 = \mu_2$)	SP1	0.0800	± 2.1450	ACCEPT
	SP2	-1.1972	± 2.1450	ACCEPT
	SP3	-0.5687	± 2.1450	ACCEPT
	SP4	0.3756	± 2.1450	ACCEPT
	SP5	1.3962	± 2.1450	ACCEPT
	SP6	-0.9674	± 2.1450	ACCEPT
	SP7	-4.5007	± 2.1450	REJECT
	SP8	-3.3204	± 2.1450	REJECT
	SP9	-4.8734	± 2.1450	REJECT
	SP10	-2.6854	± 2.1450	REJECT
	SP11	4.7925	± 2.1450	REJECT
	SP12	1.5904	± 2.1450	ACCEPT

From *t* test of the results, there are significant differences between CWB 1990-2009 and NCDR 1980-1999 data in five study points for the wave height results. As stated in previous section, this is predicted because this research uses time series data of typhoon path line and pressure for the input.

CONCLUSION AND SUGGESTION

The proposed methodology can verify whether the typhoon characters predicted by the dynamic simulation model can represent the actual typhoon characteristics on local coastal areas. According to result analysis, it can be assured that the input data (the pressure and the path line) hold an important role. If the inputs are not controlled like this research method approach, it shows there are significance differences between the results of CWB and NCDR data for some study points. Based on results analysis, it can be concluded that the simulated typhoon data from NCDR does not represent the recorded typhoon data from CWB very well. Visual analysis showed there are differences in the results. Furthermore, based on *t* test analysis there are significant difference between CWB and NCDR data in five study points for the wave height results in this research.

To strengthen the analysis about the relevancy of simulated typhoon data from NCDR to recorded typhoon data from CWB, extend the study by investigating other wave parameters, such as wave period and wave length.

ACKNOWLEDGMENTS

This study was financially supported by the research project of Water Resources Agency, Ministry of Economic Affairs of Taiwan under grant number MOEAWRA1020048 and MOEAWRA1030316. The support of the National Science Council of Taiwan through grant NSC101-2911-I-006-515 is also acknowledged.

REFERENCES

Buckingham, A., and P. Saunders. 2004. *The Survey Methods Workbook: From Design to Analysis*, Polity Press, Cambridge, UK.

Coastal Engineering Manual. 2002. *Coastal Engineering Manual*, Engineer Manual 1110-2-1100. U.S. Army Corps of Engineers, Washington DC, USA.

Fisher, R.A.. 1990. *Statistical Methods, Experimental Design, and Scientific Inference*, Oxford University Press, Oxford, UK.

Goda, Y.. 2000. *Random Seas and Design of Maritime Structures*, Advanced Series on Ocean Engineering - Vol. 15. World Scientific Publishing, Singapore.

Graham, H.E., and D.E. Nunn. 1959. *Meteorological Conditions Pertinent to Standard Project Hurricane, Atlantic and Gulf Coasts of United States*, National Hurricane Reserach Project, 3. US Weather Service.

Hasselmann, K., T.P. Barnett, E. Bouws, H. Carlson, D.E. Cartwright, K. Enke, J.A. Ewing, H. Gienapp, D.E. Hasselmann, P. Kruseman, A. Meerburg, P. Müller, D.J. Olbers, K. Richter, W. Sell, and H. Walden. 1973. Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), *Deutschen Hydrographischen Zeitschrift,* A8 (12), 95.

Hasselmann, S., K. Hasselmann, J.H. Allender, and T.P. Barnett. 1985. Computations and parameterizations of the linear energy transfer in a gravity wave spectrum. Part II: Parameterizations of the nonlinear transfer for application in wave models, *Physical Oceanography*, 15, 1378-1391.

Holland, G.J.. 1980. An analytical model of the wind and pressure profiles in hurricanes, *Monthly Weather Review*, 108, 1212-1218.

Hsu, T.W., J.M. Liau, J.G. Lin, J. Zheng, and S.H. Ou. 2011. Sequential assimilation in the wind wave model for simulations of typhoon events around Taiwan Island, *Ocean Engineering,* 38(2–3), 456-467.

Hsu, T.W., S.H. Ou, and J.M. Liau. 2005. Hindcasting nearshore wind waves using a FEM code for SWAN, *Coastal Engineering*, 52(2), 177-195.

Iversen, G.R., and H. Norpoth. 1987. *Analysis of variance*, Sage Publications, California, USA.

Jackson, S.L.. 2006. *Research Methods and Statistics: A Critical Thinking Approach*, Thomson Wadsworth, California, USA.

Kreyszig, E.. 1999. *Advanced Engineering Mathematics*, John Wiley & Sons, Singapore.

Makin, V.K., and V.N. Kudryavtsev. 1999. Coupled sea surface-atmosphere model, 1, Wind over waves coupling, *Geophysic Resources*, 104(C4), 7613-7623.

Myers, V.A.. 1954. *Characteristics of United States Hurricanes Pertinent to Levee Design for Lake Okechobee, Florida*, Goverment Printing Office (No. C30.70:32), Washington, D.C, pp. 126.

Ou, S.H., J.M. Liau, T.W. Hsu, and S.Y. Tzang. 2002. Simulating typhoon waves by SWAN wave model in coastal waters of Taiwan, *Ocean Engineering*, 29(8), 947-971.

Powell, M.D.. 1980. Evaluations of diagnostic marine boundary layer models applied to hurricanes, *Monthly Weather Review*, 108, 758-766.

Shea, D.J., and W.M. Gray. 1973. The hurricane's inner core region, I: Symmetric and asymmetric structure, *Atmospheric Sciences*, 30, 1544-1564.

Shore Protection Manual. 1984. *Shore Protection Manual, 1*, U.S. Army Corps of Engineers, Washington DC, USA.