10,000 Year Wave Statistics for a Tropical Cyclone Region

Conference Paper · July 2005

6 authors, including:

Luciano Mason
University of Tasmania
60 PUBLICATIONS **658** CITATIONS

Jason Mcconochie
Shell Global
24 PUBLICATIONS **239** CITATIONS

Lance Bode
James Cook University
55 PUBLICATIONS **987** CITATIONS

10,000 YEAR WAVE STATISTICS FOR A TROPICAL CYCLONE REGION

Tom Hardy[1], Luciano Mason[2], Ashley Astorquia[3], Jason McConochie[4], Lance Bode[5] and Bruce Harper[6]

Abstract: A state-of-the-art modelling system required adaptation and development of a series of models including synthetic tropical cyclone, windfield, and wave models to determine wave information for significant wave heights with return periods from 100 to 10,000 years for the oil and gas producing regions of northern and northwest Australia. A very large ensemble of approximately 450,000 individual synthetic tropical cyclones (100,000 years) was created to represent the tropical cyclone population of the study area. Methods were tested in order to reduce the number of simulated storms to a feasible level to accurately define the return period above the 100 year level.

INTRODUCTION

Return periods of wave conditions during tropical cyclones are required for the design of oil and gas facilities in the North-west and Northern Australian region. However, measured wave data during tropical cyclones is not available at many of the desired locations and where available is not of sufficient duration from which to determine severe tropical cyclone wave design conditions. Consequently, modelling of tropical cyclone winds and waves is required for the determination of tropical cyclone design criteria. Historical tropical cyclone databases are routinely used to determine the 100-year return period tropical cyclone design wave criteria; however, with a move by the oil industry to design facilities for a reliability of 1 in 10,000, this historical dataset is inadequate.

Modelling of tropical cyclone waves relies on the specification of the wind fields of tropical cyclones requiring, as a minimum, specification of the cyclone position, central pressure and radius of maximum winds (McConochie et al., 2004). However, as is the case with other tropical cyclone regions, the length of the historical record in the North-west Australian region is not sufficient for the accurate determination of the winds that would cause the 1 in 10,000 year wave conditions.

---

[1] Marine Modelling Unit, Australian Maritime College, Launceston, TAS, 7250 Australia, t.hardy@amc.edu.au
[2] Marine Modelling Unit, Australian Maritime College, Launceston, TAS, 7250 Australia, l.mason@amc.edu.au
[3] School of Engineering, James Cook Uni., Townsville, QLD, 4811 Australia, ashley.astorquia@jcu.edu.au
[4] Woodside Energy Ltd, 240 St Georges Tce, Perth, W.A., 6000, Australia, jason.mcconochie@woodside.com.au
[5] School of Maths & Physics, James Cook Uni., Townsville, QLD, 4811 Australia, lance.bode@jcu.edu.au
[6] Systems Engineering Australia, 7 Mercury Court, Bridgeman Downs QLD 4035, Australia, seng@uq.net.au

As a consequence of the shortness of the record, the generation of a synthetic population of tropical cyclones is essential. This paper describes the generation of this synthetic population. The techniques presented here have been continuously evolving in a series of projects over the last several years in the Coral Sea off the north-eastern coast of Australia (James and Mason, in press; Hardy *et al.* 2004a, b; Hardy *et al.* 2003; and Hardy *et al.* 2000).

POSITION AND PRESSURE MODEL

James and Mason (in press) present an empirical approach to the modelling of the time series of position and pressures of tropical cyclones affecting the Great Barrier Reef Marine Park (GBRMP) and the Queensland coast. The approach is similar to that of Vickery *et al.* (2000), but the historical database in the Coral Sea is much smaller than of the western Atlantic, and the models are simpler. An important feature of the James and Mason model is the random interpolation similar to Scheffner *et al.* (1996) of the historical initial conditions to provide a large database from which to initiate the simulations.

Model Equations

Although based on the James and Mason (in press) model, the formulation of model developed for the current project has significant differences. These are not merely new values of coefficients for the new region, but more importantly the continued development and reformulation of the technique. The North-Western Australian region differs from the Coral Sea both in the characteristics of its tropical cyclones and in the availability and quality of the historical tropical cyclone database (Woodside Energy Tropical Cyclone Best Tracks). The different characteristics of the region required new formulations of the three governing equations (changes in time of latitude, longitude and central pressure) and the development of sub-regions each with its own unique values of coefficients.

The governing equations are equations for the changes in time or *velocities* of longitude, latitude and central pressure deficit and are given as:

$$V_\lambda^{t+1} = a_0 + a_1 V_\lambda^t + a_2 \lambda^t + a_3 A_\lambda^t + a_4 \phi^t + a_5 V_\phi^t + a_6 U_\lambda + a_7 U_\phi + \varepsilon_\lambda \tag{1}$$

$$V_\phi^{t+1} = b_0 + b_1 V_\phi^t + b_2 \phi^t + b_3 A_\phi^t + b_4 V_\lambda^t + b_5 U_\phi + b_6 V_\phi^t U_\phi + b_7 V_\lambda^t U_\lambda + b_8 V_\lambda^t U_\phi + b_9 p_d^t + \varepsilon_\phi \tag{2}$$

$$V_{p_d}^{t+1} = c_0 + c_1 V_{p_d}^t + c_2 {p_d}^t + c_3 e^{c_4 (p_d^t - p_{dMPI})} + c_5 U_\lambda + c_6 U_\phi + c_7 h + \varepsilon_{p_d} \tag{3}$$

In which λ, ϕ, and p_d are the longitude, latitude and central pressure deficit; V indicates changes with time in longitude, latitude and central pressure deficit (according to the subscript) and A the acceleration of the same; U_ϕ and U_γ are the meridional (north-south) and zonal (east-west) wind speed; P_{dMPI} is the maximum potential intensity of the pressure deficit of the storm (Emanuel 1988) , h is the water depth and the superscript indicate the time step. Values of the model coefficient vectors, ***a***, ***b*** and ***c*,** were estimated by ordinary least squares multiple linear regression combined with a simple "directed search" for c_4 in exponential term of eqn (3). A large set (>2000) of residuals (ε_λ, ε_ϕ,, and ε_{pd}) was determined as the one-step-ahead prediction errors of the deterministic equations (eqns (1) to (3), without the residual term) at each time step in historical time series of position and pressure data. During the simulation of a synthetic tropical cyclone, values of the residuals, ε_λ, ε_ϕ, and ε_{pd}, were randomly selected with replacement from the appropriate residual set. These residual terms, in effect, represent as random variations, effects that are not accounted for in deterministic model equations.

Each of the terms in the model equations was selected to provide specific characteristic of tropical cyclone behaviours. For example, the first two terms on the RHS of each of eqns (1) to (3) model persistence from one time step to the next in each of the velocities; The fourth term on the RHS of (3) models the increasing tendency for changes in pressure deficit to be positive as the central pressure deficit approaches the maximum potential intensity (MPI) (see below).

Many possible terms were tested based on experience and judgement. Those that improved the comparison of model and measurements for a variety of measures (see below) were selected for inclusion. Those that offered little improvement were discarded. Linear terms composed of a coefficient times the values of latitude, longitude and central pressure deficit, as well as their velocities and accelerations were considered. These terms were the result of the more obvious enquires, such as – *Are changes in position or central pressure deficit a function of position and central pressure deficit? Are they a function of the rate of change of these quantities? Are they a function of how their rate of change is changing?* Nonlinear terms, multiplying two or more parameters, were also considered and although the physical significance may be less apparent, several were included, such as terms 7, 8 and 9 in eqn. (2). Water depth was found to be useful for the pressure deficit equation in the Darwin subregion, perhaps as an indicator of the capability of a storm to lose intensity by mixing up cooler waters in deeper regions as opposed to maintaining intensity in the more well mixed, shallower shelf regions. The value of h as a function of position was available from the bathymetric grids developed for the wave modelling.

The meridional (north-south) and zonal (east-west) wind speeds (U_ϕ and U_λ) were found to be valuable parameters. These were not in the WEL database but were obtained from processing eight years of winds from modelling by Australian Bureau of Meteorology (BOM) using the Limited Area Prediction System (LAPS, BOM http://www.bom.gov.au/nmoc/bulletins). The data were edited into eight separate years and then the eight annual time series were aligned and averaged at each time of year to create a single yearly ensemble averaged time series. This yearly signal was then low pass filtered to obtain a smoother record that was used for input into eqns. (1) to (3). Recent conversations with BOM staff (J. Callaghan, personal communication) have indicated that if these winds serve as an indication of steering, a better steering parameter may be upper level (500 hPa level) winds. However this data was not available to us for the timing of this project.

The present project has employed a value of MPI that varies in two dimensions spatially as well as in time. An internal report (McGuffie, 2002) commissioned by WEL performed an analysis of MPI in the study region (based on Holland, 1997) and provided a spatial and varying monthly mean MPI for the years 1948-2001. Spatial plots of the monthly mean MPI by averaging over the 53 years of data at a spatial resolution of 2.5° × 2.5° were presented. We obtained the 53 years of monthly means from that analysis. These monthly means were aligned at the centre of each month and a daily signal was interpolated and converted from MPI to MPI deficit. The time series of MPI and MPI deficit (P_{dMPI}) at a location show an annual sinusoidal signal caused by the change in solar energy with seasons. The minimum potential central pressure approaches or goes below 860 hPa in several of the years. Examples of the temporal and spatial variations are shown in Fig. 1.

When fitting the model coefficients using data of historical tropical cyclones, the value of the MPI deficit for the time and position of the historical storm was used. When the synthetic ensemble was generated, a year from 1948 to 2001 was randomly selected and the day of the year and starting position are determined in the initialisation process (see below).

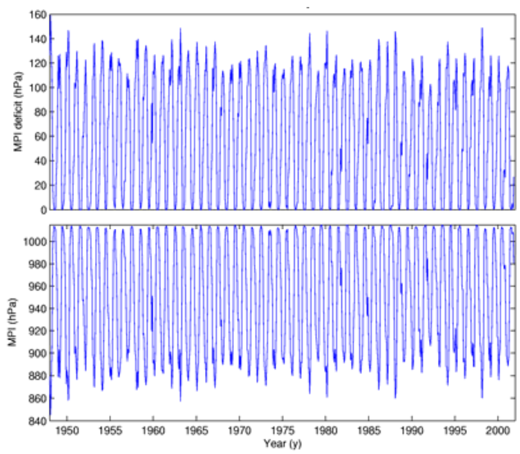

Fig. 1. Time series of mean monthly MPI (top) and mean monthly MPI deficit (bottom) at a single unidentified point for years 1948-2001.

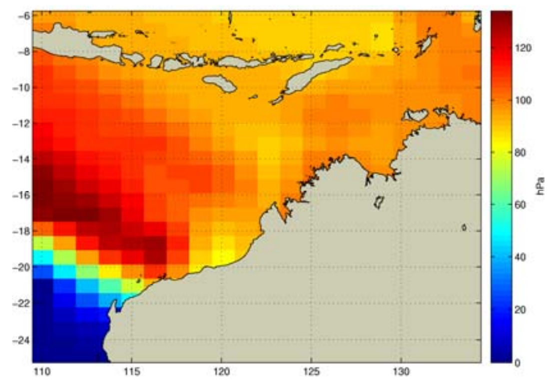

Fig. 2. Example of spatial variation in MPI deficit at an unspecified day

Model sub-regions

A major innovation in the application of track and pressure model to the present project was the inclusion of multiple sub-regions, each with its own set of coefficients. If the processes that control the genesis, development, movement and decay of tropical cyclones were completely understood and data of these were available, then one set of model equations would suffice for all regions. Since this is not the case we found that subregions were necessary. Motivation for the use of subregions was three fold: (1) Some regional changes are obvious, e.g. the differences in over-land and over-water tropical cyclones. Other regional differences may occur that are not included in the formulation of equations.; (2) Tropical cyclones in the North-Western Australian region appeared to display a greater spatial change in characteristics than those in the Coral Sea; (3) The more frequent rate of occurrence of tropical cyclones and the higher quality data set allowed the analysis of subregions in North-Western Australia compared with the Coral Sea where data in sub-region is too sparse.

During the simulation of a synthetic tropical cyclone its change in position and change in central pressure deficit is calculated using the coefficients determined for that region. During its life a synthetic storm might transit through several subregions and thus be synthesised by several sets of coefficients. The eight geographic subregions (Fig. 3), inside which different sets of coefficients are applied, were selected after extensive testing. In Fig. 3, the red polygon indicates the initialisation area. Synthetic storms are started either inside or on the boundary of this area. If a storm exits across the red boundary, results outside the red boundary are only stored if the storm re-enters the polygon later in its life. The outer green rectangle in Fig. 3 indicates the outer limits of the modelling region. If a synthetic tropical cyclone crosses this boundary, then its simulation ceased.

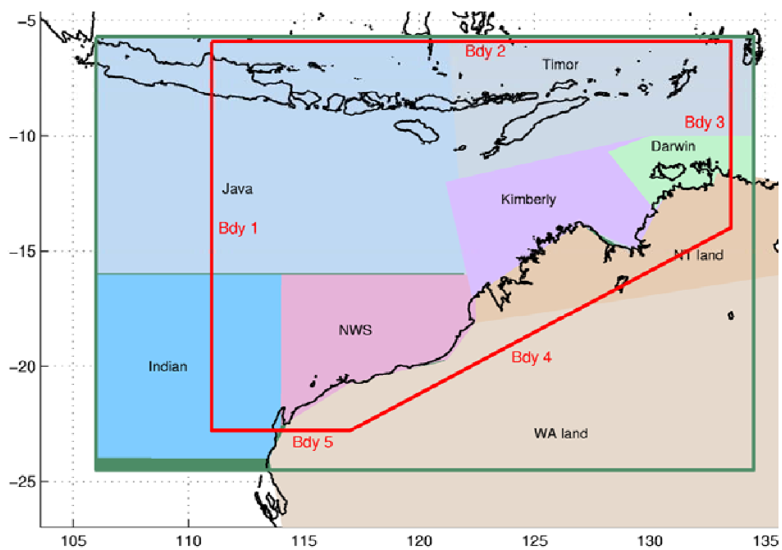

Fig. 3. The eight geographical subregions each with a unique set of model coefficients

Model Initiation

The model equations for the synthesising the database require specification of values for 17 parameters (λ, ϕ, P_d, V_λ,, V_ϕ,, V_{pd}, A_λ, A_ϕ, U_λ, U_ϕ, P_{dMPI} , h, ε_λ, ε_ϕ , ε_{pd}, Y, T_A), where Y is randomly selected year from 1948 to 2001 and T_A is the time of year. The deterministic form of the equations used to fit model coefficients require 14 parameter values (the residuals, $\boldsymbol{\varepsilon}$, are excluded). The values of U_λ, U_ϕ, P_{dMPI}, h, and Y are deterministic. The initial values of A_λ, A_ϕ, were set to zero. The initial values of parameter values of λ, ϕ, P_d, V_λ,, V_ϕ,, V_{pd}, and T_A were obtained from a random selection from within a seven-dimensional hypersphere space centred on a randomly selected historical starting point. For initialisation in the current project, the radius of the hypersphere was the 7th norm distance that was determined by

$$r_{init} = \sqrt[7]{\sum_{i=1}^{7} d_i^2} \tag{4}$$

in which d_i are the one-dimensional distances to the nearest neighbour after the values of the seven parameters have been standardised by subtracting the mean and dividing by the standard deviation. The 7th norm distance was used to prevent the Euclidian distance of the multiple dimensional hyperspace from becoming too large.

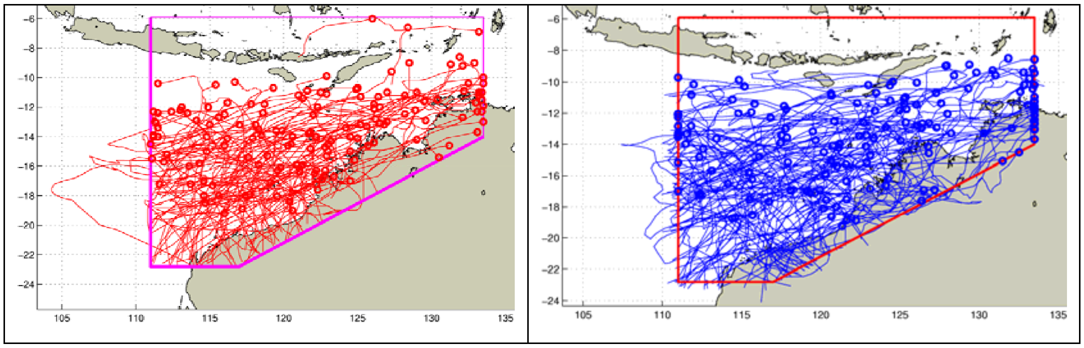

Figure 4. Starting positions (circles) and tracks of the 150 historical (red) and 150 synthetic (blue) tropical cyclones that started inside or crossed into the initialisation zone. Storms that crossed into the zone were started on the boundary. Those tracks that appear outside the zone are from storms that left and re-entered.

The tracks and the starting locations for the historical tropical cyclones used to formulate and test the model are shown in Fig. 4a. A total of 150 historical tropical cyclones entered or started inside the red polygon. A set of randomly selected synthetic tropical cyclones tracks equal in number to this historical set is presented in Fig. 4b. Numerous visual tests (not shown here) were conducted in the process of determining both the form of the equations and the values of the model coefficients before the production model was finalised.

GOODNESS-OF-FIT TESTS FOR POSITION AND PRESSURE MODEL

Both numerical and visual goodness-of-fit tests were conducted for the iterative process of model development and to validate the final model. Examples of some the comparisons of model and measurements used in model development and validation are contained in Fig**.** 5. The green polygon in Fig**.** 5(a) shows the geographical area in which the comparison is conducted. This will be called the *comparison polygon*. Comparisons are between historical data and model results are tallied only when historical and synthetic storms are inside the polygon. In this example, n_h = 27 historical and n_s = 8541 synthetic tropical cyclones entered or originated in the polygon of Fig**.** 5a. In all graphs in Fig**.** 5 the red curve is historical data and the *thick* blue curve is model results. Both model results and historical data will have uncertainties and random variations and a measure of confidence that the model results and historical data are from the same population is needed.

Cumulative distribution functions (CDF) were calculated and plotted for (lettered as the corresponding sub-graph of Fig**.** 5): (d) maximum central pressure deficit; (f) speed of storm; (h) east-west component of the speed of the storm; (i) north/south component of speed; (j) heading of storm; (l) total time in hours; and (m) mean annual time. All of these (both historical and synthetic) were calculated using data when the storm was inside the comparison polygon.

The two *thin* blue curves in the graphs of Fig**.** 5 are analogous to 90% confidence limits. They are calculated with a bootstrap-like process in which 27 storms (equal in number to the number of *historical* storms) were randomly selected with replacement from the whole set of *synthetic* storms that were inside the green polygon sometime during their lives. Statistics and the corresponding plots were generated from each of these samples. This process was repeated 5000 times and 5000 CDF curves were generated. As an example of this process of generating these confidence curves we will examine, Fig**.** 5d which is the CDF of maximum intensity. This data is the maximum central pressure deficit that occurred while the tropical cyclone was inside the comparison polygon. All of the 5000 bootstrapped curves would have the same set of 27 values of ordinate (cumulative probability). At each of these ordinates, there would be a value of maximum intensity (maximum central pressure deficit) for each of the 5000 curves. The two *thin* blue curves have values of the maximum intensity (abscissa) for which 5% of the 5000 values of maximum intensity are less (left most curve) and 5% greater (right most curve). Thus 90% of the values lie between these bounds at each of the 27 values of the ordinate.

An indication of the confidence bounds of *whole* curves in addition to that of individual ordinates is valuable. A technique was developed (James and Mason, in press) to give a visual estimate of the region in which the 90% of bootstrapped curves would lie. The two *medium* thickness blue curves in the plots of Fig**.** 5 contain between them 90% of the 5000 bootstrapped curves.

Fig**.** 5b and 5c proved to be the most valuable indicators of model performance during model development. The frequency of cyclone occurrence close to the important locations is of obvious interest, and the frequency as a function of storm intensity is crucial to the determination of the 1 in 10,000 year wave conditions. However, simulated storms, initiated over a very large domain, often travelled considerable distances to affect the locations of interest. It is important that numbers of simulated cyclones reach each of the locations of interest and have a distribution of intensity at these locations that are consistent with the historical rate of occurrence. For Fig**.** 5b and 5c the maximum intensity (maximum central pressure deficit) of each storm is determined during the time that it is in the comparison polygon. These are then converted to λ_{cPd}, the number of cyclones per year that have a value of maximum central pressure deficit equal to or greater than the stated value (Fig**.** 5b). Another way of viewing this information is by creating a return period ($R = 1/\lambda_{cPd}$) curve of maximum central pressure deficit inside the area of the green polygon (Fig**.** 5c).

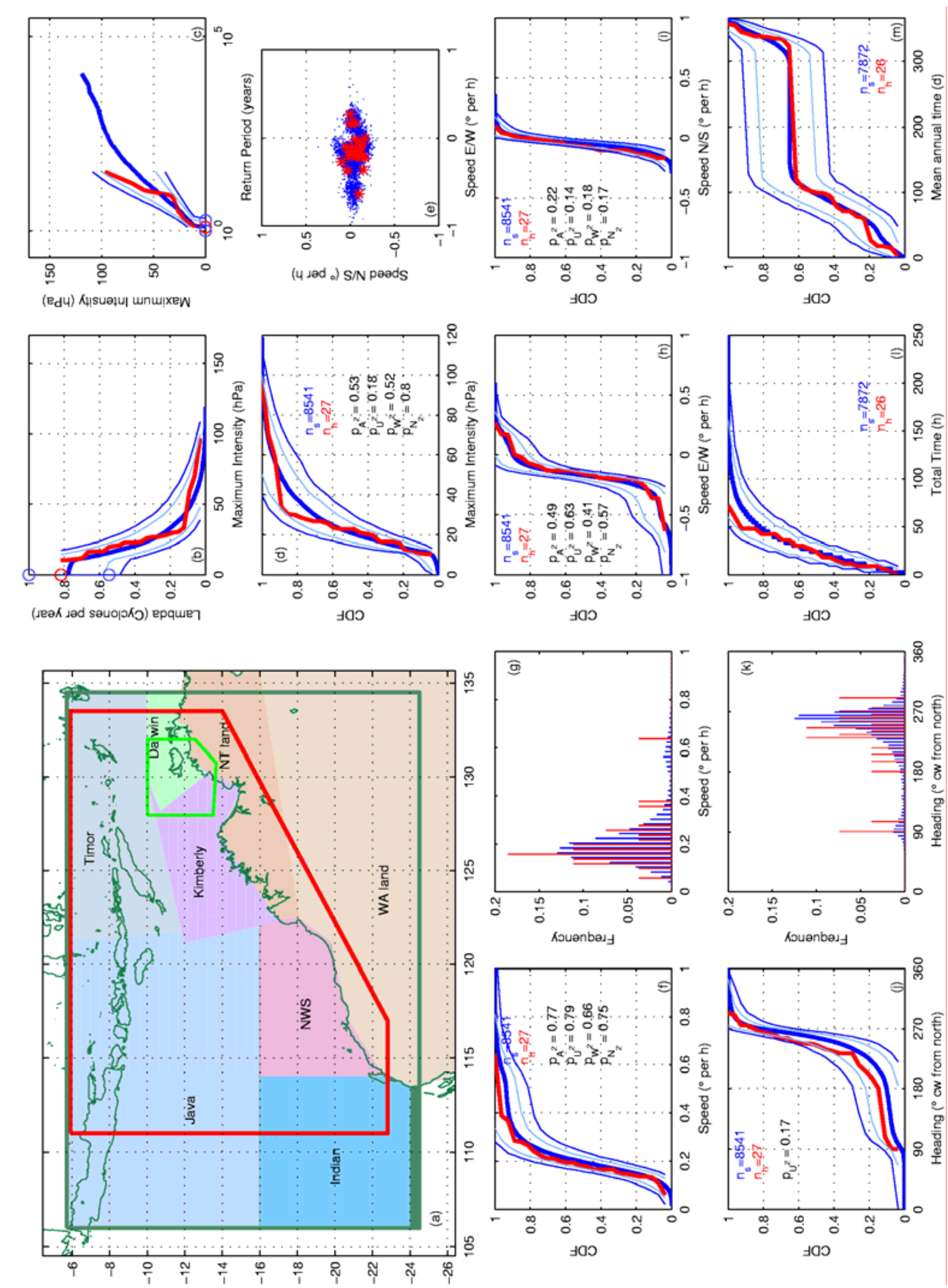

Fig. 5. Example comparison between historical data (red) and model results (blue collected inside the green polygon in (a). Thin blue curves in graphs are 90% confidence bands about individual ordinates. The medium blue curves bound 90% of a set of 5000 curves each with the same number of points as the historical set, which were obtained from a bootstrap-like analysis of the synthetic results. Goodness of fit statistics (*P* value) for W^2 (Cramer-von Mises), A^2 (Anderson-Darling), U^2 (Watson), N_2 (Neyman) are included on the CDF figures.

The thick blue curve in Fig**.** 5b indicates the rate of occurrence of tropical cyclones in the whole of the synthetic data set which have a maximum central pressure deficit while inside the comparison polygon that is equal to or greater than then value of the abscissa. The thin and medium blue lines indicate 90% bounds of individual ordinates and whole curves, respectively. These bounds are determined in a similar manner to that described above for the seven CDF curves.

In addition to the visual test just described, a variety of tests are available for the statistical comparison of simulated and historical cyclone parameters. These test the hypothesis that the historical values come from the same statistical population as the simulated values. Stephens (1974) evaluated several goodness of fit statistics and found that the statistics D (Kolmogorov-Smirnov), W^2 (Cramer-von Mises) and A^2 (Anderson-Darling) are relatively powerful for errors in the mean; and V (Kuiper) and U^2 (Watson) are better for errors in the variance. Results also show that overall, D and the *chi-square* statistic are significantly less powerful than the others. He suggests that W^2, U^2 and A^2 should generally be used, although of these, only U^2 should be used for observations recorded on the circle (such as the direction of forward movement or heading) (D'Agostini and Stephens, 1986). Also, the Neyman statistic N_2 combines the more common sample mean and sample variance, and provides an effective test for uniformity against a wide range of alternatives (D'Agostini and Stephens, 1986).

The P values of these statistics (P_{A2}, P_{U2}, P_{W2} and P_{N2}) are included on plots in the comparison Fig.5. A P value of $P = 0.1$ is analogous to our visual measure that the historical CDF in the plots falls within the region which bounds 90% of the bootstrapped CDFs. The A^2 statistic is the closest to our visual measure. The higher the p value the better the fit. We considered that $p \geq 0.1$ to indicate an acceptable fit between modeled results and historical data.

CONCLUSION

A total of 454,546 tropical cyclones was simulated in order to produce 100,000 years of data for the current population of tropical cyclones that threatens the study area. This large number creates ten renditions of 10,000 years in order to increase the confidence of the 10,000 year wave conditions. The method to determine the population of tropical cyclones that threaten the study area has been discussed. Methods to cull this very large ensemble to a size that is feasible to model and the modelling by numerical wave generation models will be discussed at the conference.

REFERENCES

D'Agostini, R.B. and Stephens, M.A. (eds.) (1986). "Goodness-of-Fit Tests." Marcel Dekker, Inc., New York.

Emanuel, K.A. (1988). "The maximum intensity of hurricanes." *J. Atmospheric Sciences*, 45, 1143-1155.

Hardy**,** T.A.**,** Mason, L.B. and Astorquia, A. 2004. *Queensland Climate Change and Community Vulnerability to Tropical Cyclones: Ocean Hazards Assessment Stage 3, The Frequency of Surge Plus Tide During Tropical Cyclones for Selected Open Coast Locations Along the Queensland East Coast*, Marine Modelling Unit report, prepared for Bureau of Meteorology and Environmental Protection Agency, School of Engineering, James Cook University.

Hardy, T.A., Mason, L.B., Astorquia, A., and Harper B.A. 2004. *Queensland Climate Change and Community Vulnerability to Tropical Cyclones: Ocean Hazards Assessment Stage 2, Tropical Cyclone-Induced Water Levels and Waves: Hervey Bay and Sunshine Coast*, Marine Modelling Unit report prepared for Bureau of Meteorology and Environmental Protection Agency, School of Engineering, James Cook University.

Hardy, T.A., McConochie, J.D., and Mason, L.B. (2003). "Modelling tropical cyclone wave population of the Great Barrier Reef." *J. of Waterway, Port, Coastal and Ocean Engineering*, ASCE, 129(2), pp. 104-113.

Harper, B.A. 2002. "Tropical Cyclone Intensity Estimation for the 10^{-4} Waves Study", Systems Engineering Australia report for WEL, 24 p.

Holland G.J., 1997, The maximum potential intensity of tropical cyclones, J. Atmos. Sci. 54, 2519-2541

Holland, G.J. (1981). "On the quality of the Australian tropical cyclone data base." *Aust. Met. Mag*. 29,169-181.

Holland, G.J. (Ed.) (2004). "Global Guide to Tropical Cyclone Forecasting." http://www.bom.gov.au/bmrc/pubs/tcguide/globa_guide_intro.htm. Australian Bureau of Meteorology Research Centre, Melbourne, Australia. Oct. 27, 2004.

McConochie, J.D, Hardy, T.A., and Mason, L.B. (2004) Modeling tropical cyclone over water wind and pressure fields, *Ocean Engineering*, Vol. 31, pp. 1757-1782.

James, M.K. and Mason, L.B. (in Press). "A synthetic Tropical Cyclone Database", *J. of Waterway, Port, Coastal and Ocean Engineering*, ASCE

McGuffie, K. 2002. "Regional MPI Analysis". Report pprepared for Systems Engineering Australia, commissioned by WEL, 24 p.

Scheffner, N.W., Borgman, L.E. and Mark, D.J. (1996). "Empirical simulation technique based storm surge frequency analyses." *J. of Waterway, Port, Coastal and Ocean Engineering*, ASCE, 122(2), 93-101

Stephens, M.A. (1974). "EDF statistics for goodness-of-fit and some comparisons." *J. Amer. Stat. Assoc*., 69(347), 730-737.

Vickery, P.J., Skerlj, P.F. and Twisdale, L.A. (2000). "Simulation of hurricane risk in the U.S. using empirical track model." *J. of Structural Engineering*, ASCE, 126 (10), 1222-1237.

WNI, 2003. "Review of North West Shelf and Timor Sea Tropical Cyclone Database for the 10^{-4} Waves Study", unpublished report prepared for WEL.