

**MODELING TIME VARYING AND MULTIVARIATE
ENVIRONMENTAL CONDITIONS FOR EXTREME LOAD
PREDICTION ON OFFSHORE STRUCTURES IN A
RELIABILITY PERSPECTIVE**

ZHANG YI

NATIONAL UNIVERSITY OF SINGAPORE

2014

**MODELING TIME VARYING AND MULTIVARIATE
ENVIRONMENTAL CONDITIONS FOR EXTREME LOAD
PREDICTION ON OFFSHORE STRUCTURES IN A
RELIABILITY PERSPECTIVE**

ZHANG YI

(B. Eng. Nanyang Technological University)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2014

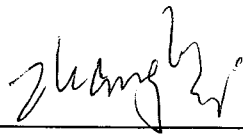
To my mother

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously

A handwritten signature in black ink, appearing to read 'Zhang Yi', is written above a horizontal line.

Zhang Yi

Nov 2013

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisors, Professor Quek Ser Tong and Professor Michael Beer, for their excellent guidance and warm encouragement throughout my Ph.D study. Their precious judgments and critical comments in the revision of my writing have helped me to refine my research tremendously. I am grateful to them for their contribution, not only the research works, but also to teaching me how to become a better researcher.

I would like to thank Dr Zhang Mingqiang for his generously sharing with me his knowledge on the uncertainty modeling and experience of Ph.D studies. I could never adequately express all the helps and supports that he has given to me.

I like to share my joy of completing the thesis with my friends, especially Dr Zhang Zhen, Miss Liu Mi, Miss Ge Yao, Mr Dai Jian, Dr Wang Yanbo, Dr Wang Li, Dr Ye Feijian, Mr Lu Yitan and Mr Luo Min. I also want to extend my sincere thanks to the other colleagues from the structural lab of National University of Singapore. Their helpful discussion and persistent friendship has made my Ph.D study quite enjoyable and fruitful.

My deepest gratitude goes to my family for their help, encouragement and support through all these years. Most importantly, I owe my loving thanks to my mother for her unlimited love and warm care. Without her continuous support and encouragement, I never would have been able to achieve my goals.

Finally, I wish to acknowledge the heartwarming support provided by my dear wife, who has always been the person most understand me and give me the unflagging love.

TABLE OF CONTENTS

TITLE PAGE	
DECLARATION PAGE	
ACKNOWLEDGEMENT	i
TABLE OF CONTENTS	iii
SUMMARY	vii
LIST OF TABLES	x
LIST OF FIGURES	xiii
LIST OF SYMBOLS	xvii
Chapter 1 Introduction	1
1.1 Background	1
1.1.1 Robust Extreme Models.....	4
1.1.2 Time Varying Environment	5
1.1.3 Multivariate Environment.....	6
1.1.4 Efficient Methods for Multivariate Analysis	7
1.2 Objectives and Scope of Thesis	8
1.3 Limitations	9
1.4 Organization of Thesis	10
Chapter 2 Literature Review	14
2.1 Environment Modeling in Analysis of Offshore Structures.....	14
2.2 Framework of Reliability Analysis	19
2.2.1 Measures of Reliability.....	21
2.2.2 Simulation Methods	22
2.2.3 Transformation Techniques	26
2.3 Long Term Assessment Criteria.....	28
2.4 Extreme Value Theory	32
2.4.1 Asymptotic Model	33

2.4.2	Inference for the Extreme Value Distribution	35
2.5	Concluding Remarks	39
Chapter 3	Establishing Robust Extreme Value Model	41
3.1	Introduction	42
3.2	Peak-Over-Threshold (POT) Method.....	44
3.2.1	Pareto Family	44
3.2.2	Poisson-GPD Model	45
3.2.3	Declustering	48
3.2.4	Parameter Estimate Method.....	53
3.3	Uncertainty Assessment of POT Method.....	57
3.3.1	Effects of Tail Behavior.....	59
3.3.2	Effects of Noise.....	64
3.3.3	Effects of Range of Dependency	70
3.4	Effects of Nonstationarity through Random Set Approach	76
3.4.1	Review of Random Set and Dempster-Shafer Structure.....	77
3.4.2	Selection of Threshold and Time Span.....	80
3.4.3	Uncertainty Quantification.....	85
3.5	Concluding Remarks	89
Chapter 4	Modeling the Time Varying Environmental Condition for Offshore Structural Analysis	92
4.1	Introduction	93
4.2	Field Data at Ocean Site.....	96
4.2.1	Seasonal Characteristics.....	96
4.2.2	Directional Characteristics.....	98
4.3	Test for Stationarity of Poisson-GPD model.....	103
4.3.1	Segmentation Algorithm for Seasonality.....	104
4.3.2	Segmentation Algorithm for Directionality	113
4.4	Time Varying Modeling.....	118
4.4.1	2D Fourier Series Characterization.....	118

4.4.2	Model Validation	123
4.5	Static Push-Over Analysis.....	129
4.5.1	Structural Model Description.....	129
4.5.2	Reliability Analysis with Importance Sampling.....	132
4.6	Concluding Remarks	139
Chapter 5	Modeling the Multivariate Environmental Condition for the Offshore Structural Analysis.....	142
5.1	Introduction	142
5.2	Bivariate Models for Sea State Parameters	145
5.2.1	Conditional Joint Distribution Model	145
5.2.2	Nataf Model	148
5.3	Copula Theory	150
5.3.1	Definition and Basic Properties	150
5.3.2	Examples of Copula.....	153
5.3.3	Dependence Concepts.....	158
5.4	Comparative Study in Multivariate Modeling	160
5.4.1	Data Pre-treatment	160
5.4.2	Application of Bivariate Models.....	167
5.4.3	Results and Discussions.....	175
5.5	Time Domain Structural Analysis.....	180
5.5.1	Proposed Discretized Copula Approach	182
5.5.2	Structural Analysis of a Fixed Offshore Platform	187
5.5.3	Results and Discussions.....	196
5.6	Concluding Remarks	200
Chapter 6	Conclusions.....	202
6.1	Summary of Thesis.....	202
6.2	Recommendation of Future Works	204

REFERENCES	209
Appendix A. Detailed Information of Four Discretization Steps	223
Appendix B. Uncertainty Assessment in POT Method	239
Appendix C. Selection of Threshold and Time Span in POT Method	254
Appendix D. Testing Values of Threshold and Time Span in POT Approach for Each Identified Time Sectors	259
Appendix E. USFOS Program Input for the Example Structure	264
Appendix F. Example of Constructing a Random Set Model	285
Appendix G. Information of Selected Wave Data and Basic Linear Wave Theory	288

SUMMARY

With the changing environmental conditions experienced over the last decade, design against failures of offshore structures has become even more challenging. Complexities exist in various steps, from translating and modeling of the environmental data, appropriate structural analysis, reliability assessment, installation, operations and maintenance. This is compounded by the presence of climatic influence which significantly affects the observed physical variables, leading to uncertainties in characterizing the system. It is essential to quantify and model the uncertainties associated with the real data collected and consider them in the long term safety assessment of new and existing offshore structures.

This thesis focuses on the characterization of the time varying characteristics of variables associated with the environmental loads, the dependencies between these variables, and investigating the impact of the uncertainties and dependencies on the long term assessment of a typical marine structure. The wave parameters which directly influence the loadings on offshore structures are studied in this work.

Considering the long time life-span and the limited data normally available, extreme value (EV) statistical models have been commonly adopted to describe environmental loads. Such class of models has assumptions that may not be fulfilled for some variables, such as independence between data points and the stationarity of the data. As such, compound extreme value statistical models have also been introduced, for example, the peak-over-threshold (POT) model. Using the

same set of simulated data, both the EV and POT models are compared in this study, where the latter is found to be a more inclusive and better approach. It is also found that the accuracy associated with each model is sensitive to the parameter estimate method used, especially if available data is limited. To characterize the uncertainties associated with the parameters, namely, the threshold u and time span Δt , a random set model is proposed in the context of the POT approach. Such elaborate measure in the form of imprecise probabilities could reflect the intensities of uncertainty in the selection of parameters more realistically.

By analyzing various segments of the collected data, stationarity of the parameters can be established. If the non-stationary is slowly time varying, then a simple means to account for this is through appropriate division of the data into stationary segments. A modified segmentation algorithm with specified fixed time interval is proposed to extract homogeneous data sets from non-stationary time series data. To select the time interval, the data fitting is recursively performed until the sample data can satisfactorily fit the Poisson-Generalized Pareto Distribution (GPD) model. Two-dimensional Fast Fourier transform is employed to characterize the variations of extreme values with time. A collected group of data is selected to demonstrate that such a discretized model can provide a more reasonable and accurate characterization of each parameter of interest. This approach of incorporating the time varying effect is examined through the reliability analysis of an existing offshore platform. The results show that incorporating the co-variate effects in the statistical model can better reflect the

underlying physical processes and help detect the most critical environmental conditions for the structure within a marine environment.

In a multivariate environment, the dependencies between the load-related environmental parameters can be a major concern. Extending beyond the linear correlation coefficient, the concept of copula is introduced to improve the quantification of the dependencies between variables as how the pairs of data are spread is considered. A further refinement is proposed through a discretized copula approach with a scheme to reduce the numerical efforts in deriving the long term response distribution of the offshore structure. A numerical example is presented to demonstrate that the proposed method produces more accurate results as the most critical sea states are identified. The newly developed method can be extended to a multivariate problem based on the copula model.

LIST OF TABLES

Table 2.1	Distribution models for selected ocean parameters.	16
Table 2.2	Target reliability levels in design codes.....	30
Table 3.1	Bias of estimated 99 th percentile for different tails $n=20, 100$	63
Table 3.2	Bias of estimated 99 th percentile for different noise condition $n=20, 100$	70
Table 4.1	Statistics of H_S over four defined seasons.....	97
Table 4.2	Statistics of wave directions over four defined seasons.....	99
Table 4.3	Summary of tested algorithms in different segmentations of H_S with respect to time.	111
Table 4.4	Summary of tested statistics for different segmentations of extreme H_S with respect to directionality.....	117
Table 4.5	Estimated shape parameter ζ in each discretized sector.	117
Table 4.6	Estimated scale parameter σ in each discretized sector.	117
Table 4.7	Random variables in reliability analysis.....	132
Table 4.8	Determined coefficients of Eq. (4.19) for different directions.	134
Table 4.9	Failure probability of jacket structure for different time and directional sector of wave model.	136
Table 4.10	Determined coefficients of Eq. (4.19) for rotated structure.....	138
Table 4.11	Failure probability of rotated jacket structure with respect to wave model.....	139
Table 5.1	Conditional joint distribution models for selected stochastic ocean parameters.....	147
Table 5.2	Examples of Archimedean Copulas.....	155
Table 5.3	Regression results for points representing H_S within interval values of 0.5.....	168
Table 5.4	Regression results for points representing H_S within interval values of 0.1.....	170
Table 5.5	Regression results for k vs H_S and λ vs H_S	172
Table 5.6	Results of marginal distribution model parameter estimates.....	174
Table 5.7	Comparison of parameter estimates and goodness-of-fit to the data for the 3 approaches.....	176
Table 5.8	Sea states and values of simulation parameters for the same accuracy.	196

Table 5.9 Results of proposed approach in determining long term base shear for $P_E=0.01$	199
Table A.1 Results of structural analysis in each subcopula.....	237
Table B.1 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=-0.5, \sigma=2$).....	240
Table B.2 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=-0.25, \sigma=2$).....	241
Table B.3 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=0, \sigma=2$).....	242
Table B.4 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=0.25, \sigma=2$).....	243
Table B.5 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=0.5, \sigma=2$).....	244
Table B.6 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1+N(0,0.1), \xi=-0.5, \sigma=2$).....	245
Table B.7 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1+N(0,0.3), \xi=-0.5, \sigma=2$).....	246
Table B.8 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1+N(0,0.5), \xi=-0.5, \sigma=2$).....	247
Table B.9 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=-0.5, \sigma=2+N(0,0.2)$).....	248
Table B.10 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=-0.5, \sigma=2+N(0,0.6)$).....	249
Table B.11 Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1, \xi=-0.5, \sigma=2+N(0,1.0)$).....	250

Table B.12	Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1$, $\xi = -0.5+N(0,0.05)$, $\sigma=2$).	251
Table B.13	Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1$, $\xi = -0.5+N(0,0.15)$, $\sigma=2$).	252
Table B.14	Bias of estimated results (shape parameter ξ , scale parameter σ and 99 th percentile) based on simulations from GPD model ($u=1$, $\xi = -0.5+N(0,0.25)$, $\sigma=2$).	253
Table C.1	P-value of K-S test for exceedances following GPD with different values of $(u, \Delta t)$.	254
Table C.2	P-value of K-S test for occurrences of exceedances following Poisson process with different values of $(u, \Delta t)$.	255
Table C.3	Estimated lower bound of 0.95 confidence interval for 100 year design value with different values of $(u, \Delta t)$.	256
Table C.4	Estimated upper bound of 0.95 confidence interval for 100 year design value with different values of $(u, \Delta t)$.	257
Table C.5	Number of exceedances above the threshold after applying the POT method with different values of $(u, \Delta t)$.	258

LIST OF FIGURES

Figure 1.1 Different types of offshore structures.	1
Figure 1.2 Controlled burning in Usumacinta jack-up after its first fire.....	3
Figure 1.3 Organization of the thesis.	13
Figure 2.1 Failure domain of performance function $G(.)$ in the variable space ...	21
Figure 2.2 Schematic showing of statistical response characterization in reliability analysis.....	25
Figure 2.3 Reliability analysis by using response surface method.....	28
Figure 2.4 Illustration of different tail behavior in Type I ($\xi = 0, \mu = 0, \sigma = 1$), Type II ($\xi = 0.5, \mu = 0, \sigma = 1$) and Type III ($\xi = -0.5, \mu = 0, \sigma = 1$).	35
Figure 3.1 (a) Scatter plot of time series measurement (b) Clusters identification (c) Identified extreme values after declustering.....	50
Figure 3.2 Bias of estimated scale parameter for different tails.....	62
Figure 3.3 Bias of estimated shape parameter for different tails.....	63
Figure 3.4 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in location parameter in GPD model.	67
Figure 3.5 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in scale parameter in GPD model.....	68
Figure 3.6 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in shape parameter in GPD model.....	69
Figure 3.7 Plots of autocorrelation functions for two tested time series (a) Case 1: $\varphi = 0.95$ (b) Case 2: $\varphi = 0$	71
Figure 3.8 Biases of estimated 99 th percentile in AMM and r largest order statistic method for two time series. (a) Case 1: $\varphi = 0.95$ (b) Case 2: $\varphi = 0$	73
Figure 3.9 Biases of estimated 99 th percentile in POT method for two time series (a) Case 1: $\varphi = 0.95$ (b) Case 2: $\varphi = 0$	74
Figure 3.10 Imprecise probability.	80
Figure 3.11 Mean residual plot with 95% confidence intervals	82
Figure 3.12 L-moment plot for exceedances over selected threshold with theoretical GPD curve.....	83
Figure 3.13 Autocorrelation plot for the wave time series records.....	84
Figure 3.14 Appropriate region for u and Δt	85

Figure 3.15 Constructed imprecise probability model for 100-year return value with mean value of all estimates.	87
Figure 3.16 Constructed imprecise probability model for 100-year return value based on different sample set: (a) Jan (b) Jan-Mar (c) Jan-Jun. Dotted line represents mean of all estimates.	89
Figure 4.1 Time varying effects on safety of structures.	94
Figure 4.2 Box plot of H_S for four different defined seasons.	97
Figure 4.3 Histograms of H_S in four different defined seasons.	98
Figure 4.4 Angular histograms of H_S over four different defined seasons (red line represents the resultant vector length).	100
Figure 4.5 2D kernel density plot of H_S with directionality and seasonality.	101
Figure 4.6 Isoline plot of H_S with directionality and seasonality.	102
Figure 4.7 Results of t -statistic for discretized time series.	108
Figure 4.8 2D kernel density plot of extreme H_S with directionality and time. .	112
Figure 4.9 Comparison of C_i values between different segmentations for time sector $[0.10, 0.35)$	116
Figure 4.10 Discrete spectrum for shape and scale parameters after DFT.	121
Figure 4.11 2D Fourier characterizations of shape and scale parameter changes along the time and direction axes.	121
Figure 4.12 CDF of original data and simulated data for $\theta \in (-10^\circ \sim 62^\circ)$	124
Figure 4.13 CDF of original data and simulated data for $\theta \in (62^\circ \sim 134^\circ)$	125
Figure 4.14 CDF of original data and simulated data for $\theta \in (134^\circ \sim 206^\circ)$	125
Figure 4.15 CDF of original data and simulated data for $\theta \in (206^\circ \sim 278^\circ)$	126
Figure 4.16 CDF of original data and simulated data for $\theta \in (278^\circ \sim 350^\circ)$	126
Figure 4.17 Comparison of simulated and original mean of H_S with time.	128
Figure 4.18 Comparison of simulated and original observed wave directions...	128
Figure 4.19 Jacket structure framing.	129
Figure 4.20 Plasticity utilization plot.	131
Figure 4.21 Comparison between calculated base shear and curve fit.	135
Figure 5.1 Sketch of several conditional probability density curves of T_P at various values of H_S	146
Figure 5.2 Comparison of different bivariate copulas for marginal distributions following standard normal distributions for correlation coefficient equals to 0.8.	156
Figure 5.3 Box plot of H_S and V_W over different months.	161
Figure 5.4 Comparison of dependencies between H_S and T_P for two different months.	164

Figure 5.5 Comparison of dependence between θ_S and θ_W over two different months.....	165
Figure 5.6 Rose plot of wave direction in Nov-Feb.....	166
Figure 5.7 Histogram of difference between wind direction and wave direction ($\theta_S - \theta_W$).....	166
Figure 5.8 Nonlinear fit of (a) $\mu[\ln(T_P)]$ vs H_S and (b) $\sigma^2[\ln(T_P)]$ vs H_S based on Eq. (5.1) and Eq. (5.3) for points representing H_S within interval values of 0.5.....	168
Figure 5.9 Nonlinear fit of (a) $\mu[\ln(T_P)]$ vs H_S and (b) $\sigma^2[\ln(T_P)]$ vs H_S based on Eq. (5.1) and Eq. (5.3) for points representing H_S within interval values of 0.1.....	170
Figure 5.10 Nonlinear fit of (a) k vs H_S and (b) λ vs H_S based on Eqs. (5.23) & (5.24).....	171
Figure 5.11 Marginal parametric model fit for (a) H_S , (b) T_P , and (c) V_W	173
Figure 5.12 Tail fittings of marginal parametric model for (a) H_S , (b) T_P , and (c) V_W	174
Figure 5.13 Comparison of contour plot between original data and (a) copula approach, (b) Nataf model, (c) conditional joint model for H_S and T_P	178
Figure 5.14 Comparison of contour plot between the original data and (a) copula approach (b) Nataf model (c) conditional joint model for H_S and V_W	179
Figure 5.15 Schematic showing of proposed discretization procedure.	187
Figure 5.16 Illustration of JONSWAP spectrum for different sea states.....	189
Figure 5.17 Illustration of JONSWAP spectrum for different peak shape parameters.....	190
Figure 5.18 Cut off frequency ω_{cut} for different peak shape parameters.	192
Figure 5.19 Determination of number of frequencies N for a given peak period.	193
Figure 5.20 Determination of time interval Δt for a given peak period.....	193
Figure 5.21 Typical simulated wave elevations.....	195
Figure 5.22 Schematic showing of discretization steps ($u=F_1(H_S)$, $v=F_2(T_P)$) and evaluation points.....	197
Figure 6.1 Schematic showing of discretization procedures for three-dimensional copula model.....	206
Figure A.1 Numbering of subcopulas in first discretization.....	233
Figure A.2 Numbering of subcopulas in second discretization.	234

Figure A.3	Numbering of subcopulas in third discretization.....	234
Figure A.4	Numbering of subcopulas in fourth discretization.	235
Figure A.5	Comparison of structural base shear for different sea states.	238
Figure C.1	Appropriate region for u and Δt for significance level equals to 0.01.	255
Figure C.2	Influence of u and Δt to (a) scale parameter, (b) shape parameter and (c) 100 year return value.	257
Figure D.1	Test of Poisson process in each identified time sector.	260
Figure D.2	Test of threshold by mean residual plot with 95% confident intervals for all four time sectors.	261
Figure D.3	L-moment plot for exceedances over the selected threshold with the theoretical GPD curve for all four time sectors.	262
Figure F.1	Random set (a) and its assigned probability weight (b).	286
Figure F.2	Probability box (a) and contour function (b).	287
Figure G.1	WIS data domain	288
Figure G.2	Geological location of the selected buoy	289
Figure G.3	Regular wave propagation properties	290

LIST OF SYMBOLS

A_n^2	Anderson-Darling statistic
AIC	Akaike Information Criteria
$Bel(x)$	belief function of x
$C(\cdot)$	copula function
C'	subcopula
$C_{Archimedean}(\cdot)$	Archimedean copula
CDF	cumulative distribution function
$C_{Gaussian}(\cdot)$	Gaussian copula
$C_\theta(\cdot)$	one-parameter copula
d	differentiation
DFT	discrete Fourier transformation
D_i	cumulative sum statistic for Poisson process
DI	dispersion index
D_n	Kolmogoriv-Smirnov statistic
$E(\cdot)$	expectation function
$Exp(\cdot)$	exponential distribution function
\overline{F}	upper bound of cumulative probability function
\underline{F}	lower bound of cumulative probability function
$F_{U_{10,max},1year}$	CDF of annual maximum of 10-minute mean wind speed
F_H	CDF of maximum wave height in a stationary sea state
$F_{V_w H_s}$	conditional CDF of 1-hour mean wind speed for a given value of H_s
$F_{V_c H_s}$	conditional CDF of current speed for a given value of H_s

$F_{H_{\max} H_S}$	conditional CDF of maximum wave height in a stationary sea state
$F_{H H_S}$	conditional CDF of wave height in a stationary sea state
$F_{U_{10}}$	CDF of 10-minute mean wind speed
F_C	CDF of current velocity
F_S	CDF of short term distribution of wave height
F_y	steel yield strength
F_ξ, F_σ	values of discrete spectrum for shape and scale parameters
F_{S_i}	cumulative distribution functions for sector S_i
$f_{T_p H_S}$	conditional PDF of T_p for a given value of H_S
GEV	generalized extreme value distribution
GPD	generalized Pareto distribution
$G_r(\cdot)$	r largest order statistics distribution function
$H_{100\max}$	maximum wave height in 100 years
H_{\max}	maximum wave height
H_S	significant wave height
$I_x(a, b)$	incomplete beta function
$L(\dots; \Theta)$	likelihood function
$\log L(\dots; \Theta)$	log-likelihood function
lim	limit
Log	natural logarithm
max	maximum
min	minimum
M_n	block maximum with size of n
M_n^r	r th largest value in a sequence of random variables

N	number of simulated wave frequency
PDF	probability density function
P_E	exceedance probability
P_f	failure probability
$Pls(x)$	plausibility function of x
Poi(.)	Poisson distribution
POT	peak over threshold method
$Q(. \theta)$	short term distribution function for $X(.)$
R	structure system's resistance
$R_{ultimate}$	ultimate strength of jacket structure
S	applied load
S_J	<i>JONSWAP</i> spectrum
S_{PM}	<i>Pierson-Moskowitz</i> (P-M) spectrum
T	length of wave simulation
T	steel tube thickness
t_i	seasonality value for observed wave height
T_p	wave peak period
T_z	zero crossing wave period
u	threshold in POT
U	vector of standard normal variables
$u(.)$	horizontal water particle velocity
$\dot{u}(.)$	horizontal water particle acceleration
$Var(x)$	variance of variable x
V_w	wind speed
$X(.)$	short term response conditioned on the ocean state parameters θ
z_p	long term value with return level $1/p$

β	reliability index
γ	non-dimensional peak shape parameter
ε	random noise
$\eta(\cdot)$	ocean surface elevation
θ_i	directionality value for observed wave height
θ_S	direction of waves
θ_w	direction of wind
μ	location parameter
μ_x	mean of variable x
ξ	shape parameter
ρ	linear correlation coefficient
ρ_s	Spearman's coefficient for measuring the dependency
ρ_θ	angular linear correlation coefficient
σ	scale parameter
σ_x	standard deviation of variable x
τ_k	Kendall's coefficient for measuring the dependency
$\phi(\cdot)$	probability density function for standard normal variables
ω_{cut}	cut-off wave spectrum frequency
ω_k, ω_l	discrete frequencies in DFT
\in	element of
\emptyset	empty set
$\lambda(\cdot)$	intensity function
\cap	intersection
τ_4 / τ_3	L-moment ratio
$\tilde{\sigma}$	modified scale parameter

\mathbb{R}	real numbers
\subset	subset of
\cup	union
$\Phi(\cdot)$	cumulative distribution function for standard normal variables
(\dots, \dots)	open interval
(\mathfrak{F}, m)	focal set
$[\dots, \dots]$	closed interval
∂	partial differentiation
ΔC	discretized subcopula
ΔP_E	subcopula exceedance probability
Δt	time span
Δt_{time_step}	time step in time domain analysis
ΔT	time interval in the time segmentation algorithm
\prod	product
\sum	sum
\int	integration
\rightarrow	mapping
∞	infinity

Chapter 1 Introduction

1.1 Background

Offshore structures facilitate the exploitation of the vast ocean resource which contributes significantly to technological and economic development. Since the first oil platform started operating at the ocean of Couissana in 1947, thousands of marine structures have been built over the last six decades. Up to now, there are more than 4000 platforms operating at the Gulf of Mexico alone. The evolution of marine technology is rapid and numerous kinds of structures are now available in the open sea, encompassing fixed structures (for example, jackets and jack-ups) in shallow waters and floating structures (for example, tension leg platforms and semi-submersibles) in deep waters, see Fig. 1.1.

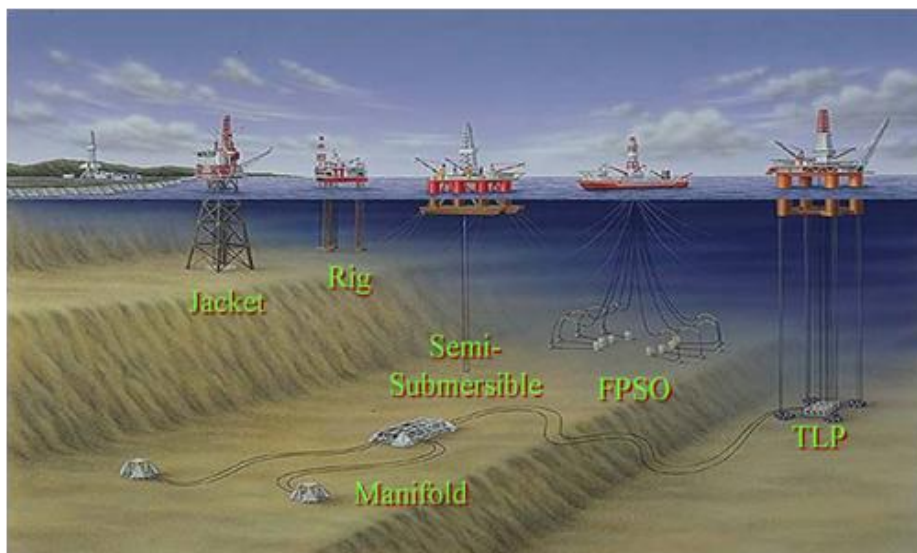


Figure 1.1 Different types of offshore structures (Roy 2013).

Compared with normal structures on land, offshore structures are bulky, expensive and in some cases have complex geometry. The marine environment for offshore structures can be severe, adverse, varied and uncertain. It covers a broad area of climatic factors which generally includes wind, waves, current, ice, tide and other catastrophe events such as earthquake, storm and tsunami. Under such environment, many unfavorable phenomena like marine corrosion, marine growth, foundation scouring, material deterioration and fatigue damage will cause a weakening of the overall strength of the structure and thus lead to an unexpected accident. Historical records depicted our inadequate understanding of the ocean environment leading to accidents with drastic consequences and huge economic loss. Examples include the first UK-built semi-submersible rig Ocean Prince and the first rig to find oil in UK waters, which broke up off England's east coast during a storm in 1967 (Oo 1974). Bohai No. 2 jack-up structure located in the Gulf of Bohai between China and Korea encountered a storm and sank on 25 November 1979, resulting in the deaths of 72 out of the 74 personnel on board (Santos and Feijo 2010). In 1980, the semi-submersible accommodation rig Alexander L. Kielland capsized during a storm after the brace supporting a leg failed in the Norwegian Continental Shelf (BMT 2006). An example of recent failures is the Usumacinta jack-up in the Gulf of Mexico which was struck by a strong storm on 21 October 2007, resulting in a fatal blowout causing 21 reported deaths with one worker missing during the evacuation (OGP 2010), see Fig. 1.2. These painful lessons provide strong motivation to improve our knowledge of the

climate-dependent ocean and its effects so as to enhance the safe operation of offshore structures.



Figure 1.2 Controlled burning in Usumacinta jack-up after its first fire (OGP 2010).

To enforce the safety requirements imposed by design codes (such as Ultimate Limit State (ULS) and Fatigue Limit State (FLS) criteria to achieve a target level of safety not exceeding more than once in 100 years (DNV 2007)) are followed, an assessment of the structure under consideration must be carried out. The objective is to ensure safe performance and limit fatalities as well as damages caused by the environment and operational loads during its service life.

Various methods for safety assessment of existing offshore structures have been developed in relation to severe environmental loading such as hurricanes (Moan 2005). The results from various methods are very much dependent on the validity of the assumptions imposed (Fitzwater and Winterstein 2001; Moriarty et

al. 2006; Saranyasontorn and Manuel 2006; Naess et al. 2007). The design code recommends several statistical models to represent physical variables such as wind and waves. Due to the complexity of the environment, these models may not be able to capture the complete characteristics of the physical variables. For example, the wave load which constitutes roughly 80% of the overall external load, is the major environmental loads. The load is affected by factors such as its period, time and direction of loading, that a single statistical model could hardly be accurate in describing the randomness of the waves. Obtaining an appropriate wave load model remains a challenge. The model should be sufficiently robust to be able to represent the most critical situation and produce a good estimate of the reliability of the structure under investigation.

1.1.1 Robust Extreme Models

The traditional modeling based on probabilistic theory has been well established and used in the characterization of environmental factors in the offshore industry (Muir & El-Shaarawi 1986). The parameters governing the models are estimated statistically using data collected from monitoring stations. Often only the extreme values govern the design loads and the design life span becomes a primary determining variable.

A commonly adopted approach is to use extreme value statistical models to arrive at a design value corresponding to the structure's lifespan. The direct annual maximum method, which models the largest value from each year, is most

widely adopted (Gumbel 1958; Leadbetter et al. 1983). However, significant inaccuracies or misrepresentation may occur if only limited amount of data are available for constructing a statistical model. Usually, the length of field data is limited to several decades which are considered short in relation to the design life of the structure of interest. Consequently, other approaches have been proposed, such as the Peak-Over-Threshold (POT) method, which characterizes the exceedance over a high threshold, to utilize more data in forming the extreme value model.

Even with such methods as POT, several important associated questions have not been addressed. These include issues on the choice of the parameter estimation method, the applicability of the method for data sample size that is small, the effect of serial dependencies, the effect of noise in the observed time series data, and the quantification of uncertainties associated with the selected threshold and limited time span of the data. Not all these issues have been adequately addressed in the literature (Næss 1998; Mackay et al. 2010; Jocković 2013).

1.1.2 Time Varying Environment

Based on different channels of observations, it is increasingly evident that the effects of climate change, especially with regards to the ocean environment, cannot be ignored. The contention is whether the assumption that the averages and extremes of sea states are stationary is valid (Vanem 2011). If non-

stationarity is assumed, a simple statistical model for the ocean parameter may not be appropriate in assessing the safety of a structure under such environment. Using the classical extreme statistics approach for climate-induced time varying phenomena may lead to bias estimations for the parameters of the extremal models and in turn, may have drastic consequences. It is important to derive a statistical model that could reflect the changes in the value with time to ensure that the reliability of the structure is estimated with minimal error. Recent researches showed that a time-dependent version of extreme value model could help to capture this non-stationary characteristic in the ocean data (Méndez et al. 2006; Marcos et al 2011). This has sparked interest in incorporating such model in the safety assessment of marine structure over its design life (Menendez et al. 2009), in particular accounting for all the uncertainties associated with such an approach (Vanem and Bitner-Gregersen 2012).

1.1.3 Multivariate Environment

It is natural that the environment parameters, such as wave period, significant wave height and wind speed, are correlated. To assume that they are independent may lead to unconservative results and hence is a potential cause for under-design leading to possible catastrophic consequences. Multivariate statistical model would be an appropriate tool to handle such complications. There exist practical guidelines on the choice of joint distribution models to characterize a multivariate environment (DNV 2007). The issue is whether these models can be applied to

all the ocean sites. The suitability of various computational methods to handle dependencies has yet to be completely answered. It is therefore of interest to perform a comparative study highlighting the characteristics in each approach in modeling multivariate data.

1.1.4 Efficient Methods for Multivariate Analysis

To assess the long term performance of a structure within a multivariate environment associated with many environment conditions requires many numerical simulations and is computationally very expensive. The current practice is to divide the scatter diagram describing the environmental parameters into blocks and to calculate the response of the structure associated with for each block. The results for all blocks are then combined to obtain the overall response distribution. Such an approach involves numerous combinations even with a small number of environment conditions, not to say that the computation of the response for each combination can be computationally cumbersome. A complete evaluation may demand hundreds of hours of CPU time on a personal computer (Agarwal and Manuel 2009). Besides, it may not be easy to deduce from such computational results the critical environment conditions that contribute significantly to the overall long term performance of the structure (for example, floating structures could be quite sensitive to wave loading pertaining to certain critical wave periods and/or directions). It is thus of interest to develop efficient

methods, including one that allows for quick sensitivity studies to deduce the significant parameters and their range of critical values.

1.2 Objectives and Scope of Thesis

As a further contribution towards a more realistic assessment of the reliability of offshore structures, the main objective of the proposed research is to develop an accurate statistical model and reliability computation procedure which will consider the time varying uncertain non-stationary characteristics of the wave loads, as well as the dependencies amongst the key parameters. To achieve this objective, the main foci of the research are:

1. To study the uncertainties related to an existing extreme value model. The aim is to have an improved understanding of the importance of the methods selected for the construction of an extreme value model. The performances of the methods are compared for different types of data groups. In addition, the concept of random set to tackle the difficulties associated with parameter selections in the Peak-over-Threshold extreme method is presented.
2. To manage the non-stationarity inherent in long term data by developing a discrete statistical model to represent the time varying effects in the ocean parameters and show the importance of this in a reliability analysis. A segmentation approach is employed and 2-D Fourier transforms is used in conjunction with a probabilistic model to reflect the smooth transition in the parameters amongst the segments. The estimated reliability associated with

each segment can be studied to ascertain whether effects such as seasonality and directionality are significant.

3. To provide a better model for the multivariate environment beyond the linear correlation coefficient representation. Specifically, the discretized copula approach is introduced and evaluated for its suitable application towards the structural reliability of offshore structures by comparing the results with other existing multivariate models. The computational efficiency of the proposed technique is discussed.

1.3 Limitations

While this work endeavors to develop statistical procedures for accurate modeling of environmental effects for offshore design, many simplifications were made in the reliability analysis. The load model uncertainty, including the wave kinematics and hydrodynamic load, for a jacket structure are not considered. The current analysis only focuses on the ultimate limit state of jacket structure and the complex behavior of individual elements in the structure will not be analyzed in detail. The soil foundation is simplified to act like a spring in the structural analysis. With regards to the environment loads, only wave loads are considered using an idealized random linear wave model. Besides, the effects of breaking wave, the coupling between the wave and structure, wave and current are not addressed in this study. Rather, the key focus is on structural reliability analysis by considering environmental statistical model. The effects of corrosion, fatigue

damage and deterioration which are not likely to occur in a short time are neglected in the structural model. The developed procedures in handling the field data serves as a pre-cursor step to a full numerical analysis of an offshore engineering problem. The investigated statistical model in this thesis is only based on the collected data which is independent of the structural types. Therefore, other types of offshore structures can also be applied in this case. Lastly, the results and findings of this study is based on limited data from a given geological area. The models adopted in this study may not be the most appropriate ones in every case, especially if the field data is taking from one place that is different from the chosen site.

The conclusions drawn from the thesis should be seen in the light of these limitations. The influence of these limitations to the reliability results may need further investigations in the future.

1.4 Organization of Thesis

This thesis is organized into six chapters as follows.

Chapter 1 outlines the motivation for the research, leading to the proposed research objectives and methodology. This chapter addresses several issues in performing the reliability analysis of offshore structures under its operational environment. Three key topics associated with the long term safety assessment of offshore structures are highlighted and forms the research topics presented in the subsequent chapters.

Chapter 2 presents a concise literature review of the current state of knowledge for the reliability analysis of offshore structures. The basic mathematical model of characterizing extreme values is presented. Several probabilistic analysis techniques and their applicability in offshore engineering are reviewed, with emphasize on impact of the environment.

Chapter 3 discusses the performance of different extreme value approaches in establishing a robust statistical model. The importance of parameter estimation method associated with the effects of distribution tail behavior, noise in the data and nature of dependency are investigated. To address the difficulties in selecting the threshold and time span in POT approach, a random set based imprecise probability model is proposed to quantify this subjective uncertainty.

Chapter 4 presents an improved method for the long term safety assessment of an existing structure where the time-varying characteristics of the environment are incorporated. For this purpose, a segmentation algorithm is proposed to improve the quality of the extreme value model for the ocean parameters. Comparison of the reliability analysis results from the proposed approach against those from the traditional approach is illustrated using a real structure.

Chapter 5 discusses the probabilistic models for multivariate environment variables and the associated reliability estimation method proposed in this thesis.

The numerical procedure in evaluating the safety of the structure with environment factors having complex dependency is illustrated.

This thesis concludes with Chapter 6 which summarizes the main findings. Recommendations of possible future works are provided.

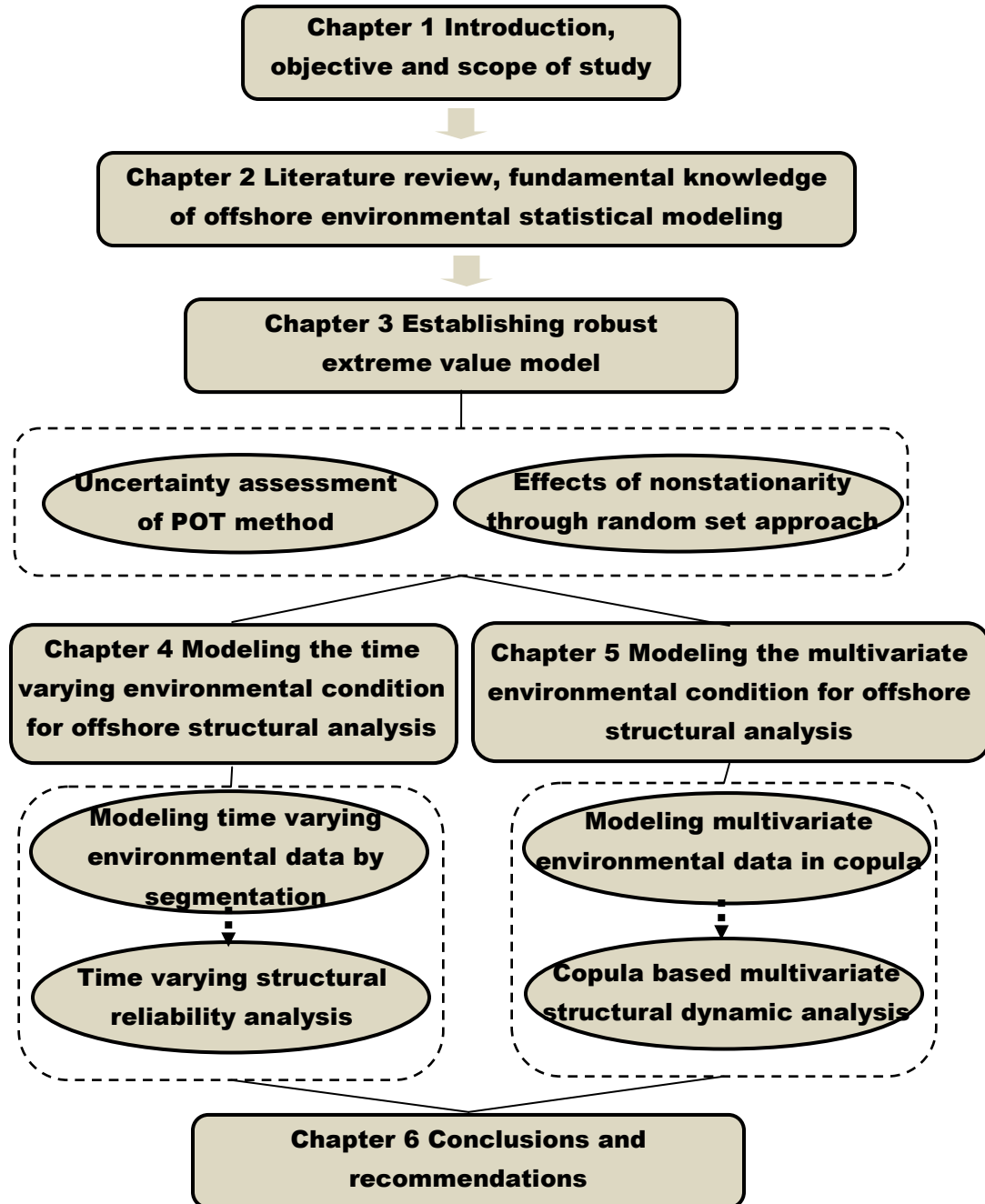


Figure 1.3 Organization of the thesis.

Chapter 2 Literature Review

This chapter presents a review of various techniques related to the reliability assessment of an offshore structure and also the available mathematical models for the ocean environment variables. For processing the available information through engineering computations, various aspects through the reliability analysis are discussed such as the required level of accuracy, the ease of handling as well as the computational efficiency. Particularly, the difficulties in assessing an offshore engineering reliability problem are emphasized in this chapter. The analysis of the influence of environment variability to the overall structural design is achieved by structural analysis. Moreover, if the long term performance of an offshore structure is a primary concern, the analysis could be realized by the probabilistic approach with the basic concept of extreme value modeling. The most key features relevant to the extreme value modeling are discussed in this chapter.

2.1 Environment Modeling in Analysis of Offshore Structures

The presence of uncertainties in engineering systems and models has been widely acknowledged (Ang and Tang, 1975; Hokstad et al. 1998; Straub and Faber 2005; Gao and Moan 2009; Lee and Song 2011; Agarwal and Manuel 2011). The uncertainties associated with offshore structural analysis lies in the structural

characteristics and associated environmental loadings. The former arises from the uncertainty in the response transfer function, the variability of the structure's strength, geometry as a result of production and manufacturing, and deteriorations of the materials. The latter results from the variability in the environmental factors, such as wind velocity, wave characteristics (magnitude, direction, height and period), and the uncertainty in characterizing the loads.

An important task in the design and reliability assessment of an offshore structure is the modeling of the environmental parameters. A proper choice of the distribution type of the environmental dependent load parameter is very critical as the results of a reliability analysis may be very sensitive to the tail of the probability distribution. However, since there is no theoretical basis to support the choice of any particular model, there are many different methods and models applied to the ocean parameters by various authors. The most widely-used distribution for 10-min average wind speeds is the 2-parameter Weibull distribution (Manwell et al. 2002; Ramirez and Carta 2005; Morgan et al. 2011). Celik (2004) showed that the simple 1-parameter Rayleigh distribution sometimes offers a better fit to the sampled data. The Gumbel distribution is usually selected to model the current velocity (Pugh 1982; Robinson and Tawn 1997; Sauvaget et al. 2000). However, Mazumder and Mazumber (2006) argued that the model should take care of the directional effects. Jaspers (1956) was the first to propose a lognormal distribution model to represent the wave height. This was later discussed and compared with many other models, such as the Weibull distribution (Battjes 1972), the mixed Weibull-Lognormal distribution (Haver 1985), the

generalized Gamma distribution (Ochi 1992) and the Beta distribution (Ferreira and Guedes Soares 1999). In addition, Nolte (1973) and Muir and El Shaarawi (1986) were among the earliest to apply the extreme value theory to the wave

Table 2.1 Distribution models for selected ocean parameters (DNV 2007).

Parameter	Description	Distribution
$F_{U_{10}}$	10-minute mean wind speed in a given height z	Weibull Distribution $F_{U_{10}}(u) = 1 - \exp\left\{-\left(\frac{u}{A}\right)^k\right\}$ scale parameter A and shape parameter k are site and height dependent
$F_{U_{10,\max,1\text{year}}}$	Annual maximum of 10-minute mean wind speed in a given height z	Gumbel Distribution $F_{U_{10,\max,1\text{year}}}(u) = \exp\left(-\exp(-a(u-b))\right)$ a and shape parameter b are site and height dependent
F_H	Maximum wave height in a stationary sea state	Weibull Distribution $F_H(h) = 1 - \exp\left\{-\left(\frac{h-\gamma}{\alpha}\right)^\beta\right\}$ γ is location parameter, α is scale parameter and β is shape parameter
F_S	Short term distribution of wave height	Rayleigh Distribution $F_S(h) = 1 - \exp\left\{-\left(\frac{h}{\alpha_S H_S}\right)^2\right\}$ $\alpha_S = \frac{1}{2}\sqrt{1-\rho}$ and H_S is the significant wave height
F_C	Current velocity	Gumbel Distribution $F_C(x) = \exp\left\{-\exp\left(\frac{x-e}{d}\right)\right\}$ e and d are site and time dependent parameters
F_{H_S}	Distribution of long-term significant wave height	Weibull Distribution $F_{H_S}(h) = 1 - \exp\left\{-\left(\frac{h-\gamma}{\alpha}\right)^\beta\right\}$ γ is location parameter, α is scale parameter and β is shape parameter

height modeling. These were further developed by Ferreira and Guedes Soares (1998; 2004) through the peak over threshold method and r th largest order statistic method. Recent works include using the regressive support vector machine (Mahjoobi et al. 2009), quantile functions (Muraleedharan et al. 2012) and the maximum entropy (Petrov et al. 2013) to derive the probabilistic model. Examples of models utilized in various design code and guidelines are summarized in Table 2.1.

More importantly, among these works, the characterizing of environmental time varying effects in the statistical model has been realized to be quite critical (Vanem and Bitner-Gregersen 2012). Recent works have demonstrated several ways of employing time-dependent versions of statistical model to capture this non-stationary characteristic in the ocean data. Méndez et al. (2006) had used linear functions to model the increasing pattern over decades for the extreme value of wave height. Jonathan and Ewans (2011) adopted trigonometric function to characterize the wave value changes due to the seasonality. Vanem (2011) had applied the wavelet transformation technique to model the long term variations for the wave parameter. However, most of these works are utilizing a time dependent regression model which is depending on a defined regression function. The quality of these approaches would be highly affected by the nonstationarity along the time series. For example, the nonstationarity among the time series may reach a certain extent that none of the defined regression function can be applied. Thus, from this point of view, the discrete statistical model is more preferred (Menéndez et al. 2009). Nevertheless,

most of the discrete model approach partitioned the data into different groups in a coarse manner. It lacks a clear and accurate way in guiding people how to perform the segmentation to the environmental data. This is going to be investigated more deeply in Chapter 4.

Besides the modeling of an individual environmental parameter, the establishing of a reliable statistical model for multivariate data is more challenging. Quite a number of multivariate models for the ocean parameters have been discussed by the former researchers (Guedes Soares et al. 2001; Baarholm et al. 2010). Detailed discussion of these investigated multivariate models are provided in Section 5.2. However, the applicability of the traditional models may be quite weak especially when there are strong nonlinear relationships among the parameters. It lacks a clear criterion for selecting the appropriate model for different groups of data which may have their own characteristics. There is a need to conduct such a comparative study to highlight the characteristics in each approach in handling the multivariate data. Despite this, it should be recognized that the developed model should be applied in engineering practice. The processing of some newly developed model, for example the copula model, in the safety assessment of an offshore structure requires further development. More detailed discussion on this topic will be provided in Chapter 5.

The use of statistical models is implicit in reliability analysis. The selection of appropriate probability distributions and estimating the values of their parameters are key issues. In offshore applications, the environmental parameters are quite site specific. For example, in the North Sea, the wave loading during

storms is much more important for the design of structures in deep water environment where wind loads only represent a contribution of less than 5% of the total environmental loads. However, wind loads are of major concern in the Gulf of Mexico where the wind speed reaches 50 m/s during hurricanes. The design parameters used to describe the environmental condition is thus determined based on observations from the actual location as well as knowledge of the environmental conditions in the area. The assessment of model uncertainty requires the comparison from field measurements against model predictions.

2.2 Framework of Reliability Analysis

The identification, modeling and quantification of uncertainties in engineering are initial steps towards incorporating them into practice in design. The next step is to assess the risk associated with a particular design or system under the presence of such uncertainties. Various approaches have been used, such as the use of probability concepts and fuzzy logic. Arising from the use of probability concepts, various reliability analysis and related computational techniques have been developed over the years (Baarholm and Moan 2002; Naess and Gaidai 2008; Naess et al. 2007, 2009). The analysis yields the probability of non-performance for various critical states of interest providing a quantitative measure of structural safety. It facilitates the sensitivity study of significant parameters in relation to non-performance. In offshore engineering, the usefulness of reliability analysis has been demonstrated. The safety assessment of fixed structure has been studied

(DNV 1992; Onoufriou and Forbes 2001) including evaluation of fatigue reliability (Karadeniz 2001). These have been extended to floating structures such as tension leg platforms (DNV 1995) and ship/FPSO hull structures (Moan et al. 2006). Moan and Song (2000) and Onoufriou (1999) developed reliability-based techniques to optimize the plans with regards to the inspection, monitoring, maintenance and repair of various structures. Zhou (2013) studied the effects of ice loading on ice-breaking tanker both numerically and experimentally. Melchers (2006) incorporated the effects of corrosion in the safety assessment of the offshore structure. Wirsching (1984) discussed the reliability assessment of fatigue problems by means of a probabilistic approach. Low et al. (2011, 2012) has presented a novel approach from a frequency domain analysis to assess the long term cumulative fatigue damage for the riser and mooring system within the wave environment. The application of the upcrossing theory in the reliability assessment of the offshore engineering has also been studied by Naess (1998) and Zayed et al. (2013). Other techniques like characterizing the uncertainty in a form of interval (Zhang et al. 2010), fuzzy set (Tao et al. 2012; Zhang et al. 2012) and imprecise probability (Beer 2013) could also be found in the literature.

Before continuing the discussions, a remark on the basic reliability theory is necessary. The following section gives a more detailed description of the fundamental law of assessing the performance of offshore structures.

2.2.1 Measures of Reliability

A common way of assessing the safety level of an existing structure is to compare the estimated probability of failure P_f against a target level specified by a design code. For a basic structural reliability problem (Melchers 1999), P_f corresponds to the probability of limit state violation for a structural system, that is, its resistance R is less than the applied load S . The failure probability is thus related to the probabilistic descriptions of load effects S and resistance R as follows:

$$P_f = \Pr(G(R,S) = R - S < 0) \quad (2.1)$$

where $G(\cdot)$ is called the limit state function or performance function dividing the load-resistance domain into a safe state ($G(\cdot) > 0$) and a failure state ($G(\cdot) < 0$) as shown in Fig. 2.1.

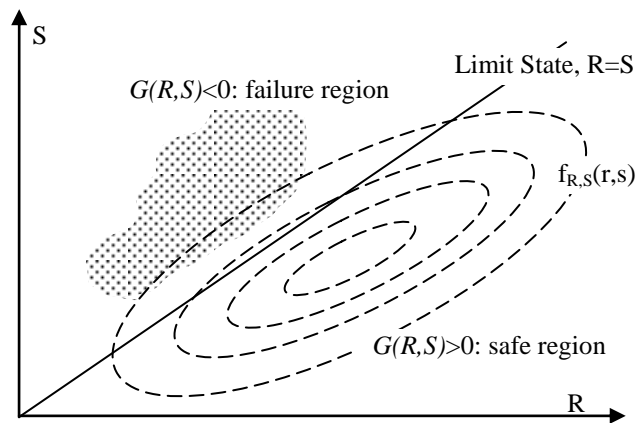


Figure 2.1 Failure domain of performance function $G(\cdot)$ in the variable space.

In most engineering problem, the resistance R and load S are vectors that contain several random variables, such as structural dimensions, material strength,

wave period, and wave height in relation to offshore applications, which influence the behavior of a structure. In that case, the performance function $G(x_1, \dots, x_n)$ would be written in an equation containing all the random variables $\{x_1, \dots, x_n\}$ in the input parameters. Thus, the failure probability can be calculated as:

$$P_f = \Pr(G(R, S) < 0) = \int \dots \int_{G(R, S) < 0} f_{x_1, \dots, x_n}(x_1, \dots, x_n) dx_1 \dots dx_n \quad (2.2)$$

where $f_{x_1, \dots, x_n}(\cdot)$ is the joint probability distribution function for all the variables. The above equation, which involves multi-dimensional integrals, is computationally feasible only for problems with few significant variables. However, not only are engineering problems governed by a sizeable number of variables, but their probability distributions may not be of the same family and the variables are not all independent of each other (Chakrabarti 1987). Such diverse possibilities give rise to numerous techniques being developed to obtain good estimates of Eq. (2.2). Broadly, the techniques can be categorized as simulation methods and transformation methods.

2.2.2 Simulation Methods

In simulation-based methods, samples of random input parameters $\underline{\mathbf{X}}$ are generated from their underlying probability models and then each set are then used to assess whether the system survives or fails. With sufficient number of sets in relation to the failure probability, the latter is estimated and should theoretically converge to the exact solution if the number of simulated data is very

large. This corresponds to the so-called direct Monte Carlo Simulation (MCS) method, which is the simplest and convenient method. Mathematically, the basic idea is to estimate an indicator function defined as

$$I_{[G(\underline{\mathbf{X}})<0]} = \begin{cases} 1 & \text{if } G(\underline{\mathbf{X}}) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

where it is directly related to the failure probability by:

$$P_f = \Pr(G(R, S) < 0) = \int \dots \int I_{[G(\underline{\mathbf{X}})<0]} f_{x_1, \dots, x_n}(x_1, \dots, x_n) dx_1 \dots dx_n \quad (2.4)$$

It is easy to see that P_f is actually an estimated value of the indicator function such as $E[I_{[G(\underline{\mathbf{X}})<0]}]$. By applying the Central Limit Theorem, the failure probability can be approached by using the sample mean of a simulated data set that has a sample size $N \rightarrow \infty$,

$$P_f \approx \bar{P}_f = E\left[I_{[G(\underline{\mathbf{X}})<0]}\right] = \frac{1}{N} \sum_{i=1}^N I_{[G(\underline{\mathbf{X}})<0]} \quad (2.5)$$

\bar{P}_f is a general unbiased estimate of P_f , and the efficiency of the estimated results can be measured by the variance of the estimator as:

$$\text{Var}[\bar{P}_f] = \frac{1}{N^2} \sum_{i=1}^N \text{Var}[I_{[G(\underline{\mathbf{X}})<0]}] = \frac{\text{Var}[I_{[G(\underline{\mathbf{X}})<0]}]}{N} \quad (2.6)$$

The advantage of applying the MCS method is its generality. The applicability and efficiency of MCS method will not be affected by the type of performance function, number of variables or dependencies among the random variables.

In estimating P_f , a response parameter is computed and then compared with the limit state value. So in reality, we can plot a histogram of the computed response parameter (see Fig. 2.2). As most well-designed structures have low P_f , only a small fraction of the computation shows exceedance of the limit state or failure. Thus, to reduce the variability, very high number of simulations is needed, making the direct MCS method infeasible especially if the problem involves a large number of stochastic variables. It thus becomes obvious that some strategy to reduce computational efforts especially for small P_f without loss of accuracy is necessary and can be realized by focusing on the relevant tail part of the histogram (that is, those sets which gives rise to failure). Indeed this leads to the class of importance sampling techniques, which limits the simulation to some “importance region” of the random parameter space, namely, those that mainly contributes to the failure probability P_f , instead of spreading the samples over the entire domain of the variables. The failure probability is evaluated as:

$$P_f = \int \dots \int I_{[G(\underline{x}) < 0]} \frac{f_{x_1, \dots, x_n}(x_1, \dots, x_n)}{g_{x_1, \dots, x_n}(x_1, \dots, x_n)} g_{x_1, \dots, x_n}(x_1, \dots, x_n) dx_1 \dots dx_n \quad (2.7)$$

where x_1, \dots, x_n are simulated based on the probability density function $g(\cdot)$. The unbiased estimator would be simply approximated as:

$$\bar{P}_f = \frac{1}{N} \sum_{i=1}^N I_{[G(\underline{x}) < 0]} \frac{f_{x_1, \dots, x_n}(x_1, \dots, x_n)}{g_{x_1, \dots, x_n}(x_1, \dots, x_n)} \quad (2.8)$$

It can be easily seen from the above equation that the selection of an appropriate sampling function $g(\cdot)$ is quite important in the importance sampling

approach. A commonly used strategy is to establish $g(\cdot)$ based on the points that have the highest probability density among all other points and are located within the failure region (Papadimitriou et al. 1997, Der Kiureghian and Dakessian 1998). Developments around this technique include for example, a kernel-based method introduced by Ang et al. (1992) for constructing the sampling function, directional sampling method (Melchers 1999) to efficiently find “important” points for nonlinear performance functions, Markov Chain Monte Carlo method to identify regions of interest (Au and Beck 1999), subset simulation (Au 2001) and line sampling (Pradlwarter et al. 2010) to efficiently estimate the results for a specific shape of performance function or group of variables. It should be noted that importance sampling method is very sensitive to the sampling function but once the appropriate one is employed, it is extremely efficient for problems with low failure probability P_f which even with today’s computing power may not be feasible with the direct MCS method.

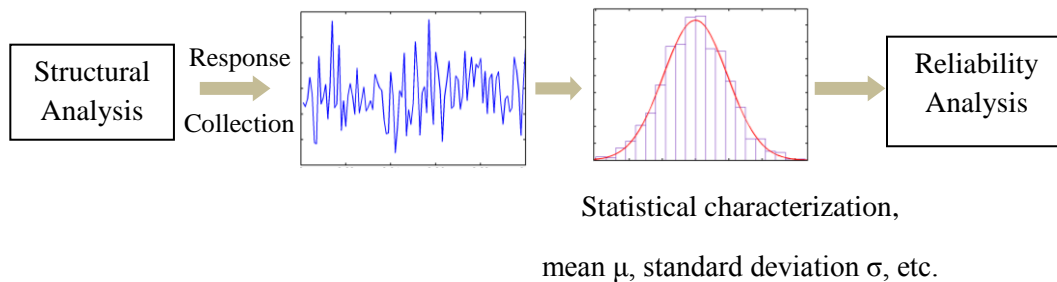


Figure 2.2 Schematic showing of statistical response characterization in reliability analysis.

2.2.3 Transformation Techniques

In classical structural reliability computation, the random variables are represented in the standard normal space such that the short distance from the origin to the limit state surface provides a measure of the reliability. If the limit state surface is planar or linear, the distance can be used to compute the actual P_f . In view of this special property, the transformation approach has been developed to compute P_f by first transforming the performance function in the original space $\underline{X} \in R^n$ to that in the standard normal space $\underline{U} \in R^n$; that is, the transformation which maps the original joint probability density function to a multi-normal probability density function.

The first of this class is the Hasofer-Lind transformation involving all components of \underline{X} that are uncorrelated normal variables (Hasofer and Lind, 1974). In the case of correlated normal variables of \underline{X} , orthogonal transformations are employed (Ang and Tang, 1984). Rosenblatt transformation has also been advocated for cases where the joint probability distribution $F_{\underline{X}}(\underline{x})$ are known whereas Nataf transformation is appropriate for the case that only partial information is known (Nataf 1962).

A practical extension of the transformation methods is to obtain good approximate estimates of P_f for nonlinear limit states or performance functions. Well-developed methods include the First Order Reliability Method (FORM) and Second Order Reliability Method (SORM) based on the Taylor-series expansion of the limit state function (Ang and Tang 1984). FORM approximates the

performance function by a linear equation at the point giving the shortest distance from the limit state ($G(.)=0$) to the origin of the standard space. The location of this point in a multi-dimensional normal space is estimated by solving the optimization equation given by

$$\beta = \min \left(\underline{U}^T \cdot \underline{U} \right)^{\frac{1}{2}} \quad (2.9)$$

subject to $G(\underline{U}) = 0$

where $G(.)$ is the performance function for the transformed random variables in the standard normal space. The optimal result of β is called the Hasofer-Lind reliability index and the point of interest is often known as “most probable failure point”. Although FORM is able to give a quick assessment of the structure’s performance, the assumed idealized linear function gradually becomes inaccurate as the dimensionality increases (Pradlwarter and Schueller 2010). To overcome this limitation, the response surface method was developed to provide an approximate relationship between the input variables and the response, especially near the most probable failure point. To obtain the response surface, regression using a polynomial function to fit the response data is performed,

$$G(\underline{\mathbf{X}}) \approx Y(\underline{\mathbf{X}}) = C_0 + \sum_{i=1}^n C_i x_i + \sum_{i=1}^n \sum_{j=1}^n C_{ij} x_i x_j \quad (2.10)$$

where $Y(\underline{\mathbf{X}})$ is the fitted function, C_i and C_{ij} are the coefficients of the polynomial expansion. This response surface method is often employed together with the simulation method to predict the response characteristics based on the input variables. The general procedure is shown in Fig. 2.3.

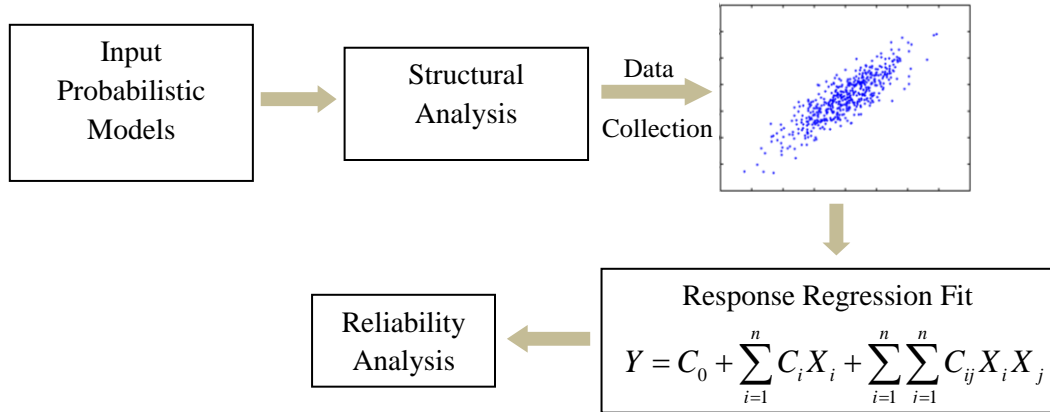


Figure 2.3 Reliability analysis by using response surface method.

Although the numerical effort is greatly reduced in the response surface method to approximate the system response function, its accuracy still remains an issue of contention. Generally, the response surface method is a good approach in continuous slightly nonlinear systems since the polynomial function can easily and accurately model the relationship. But if the system is highly nonlinear or discontinuous, the response may not be simply represented by a polynomial function. The number and accuracy demanded in the response surface method have been investigated by researchers (Myers 1999; Franchin 2002).

2.3 Long Term Assessment Criteria

The objective for offshore structural design is to design and construct the structure that could fulfill all requirements with respect to functionality, safety and economy. The current practice is to analyze the limit state or the maximum

capacity of the structure for a set of specified loads, such as that caused by the design wave. The structure, or structural element, is considered to satisfy the design requirement only if all the limit states are not exceeded. These are four limit states defined by the Norwegian Petroleum Directorate (NPD 1996) regulations:

- *Ultimate Limit State* (ULS) is defined on the basis of danger of failure, large displacement or movement, free drifting, capsizing and sinking.
- *Fatigue Limit State* (FLS) is defined on the basis of danger of fatigue due to cyclic loading.
- *Progressive Collapse Limit State* (PLS) is defined on the basis of danger of failure, free drifting, capsizing or sinking of the structure when subjected to abnormal effects.
- *Serviceability Limit State* (SLS) is defined on the basis of criteria applicable to functional capability, or durability properties under normal conditions.

The offshore structure is designed to achieve the safety levels throughout the whole operation period by meeting the limit state criteria. Associated with these criteria are implied target probability levels (Farnes and Moan 1994), examples of which are given in Table 2.2.

Normally, the evaluation of the long term behavior of an offshore structure is determined by the probability of the load level which exceeds the capacity of the structure. This requires the long term probability distribution of the load which can be obtained numerically using direct integration. That is the

exceedance probability P_E corresponding to a load level l accounting for the variability of the load related parameters θ is:

$$P_E = \Pr[L > l] = \int_{\theta} \Pr[L > l | \theta] f(\theta) d\theta \quad (2.11)$$

where $f(\theta)$ represents the joint distribution of the random variables associated with the load L represents the load. In offshore engineering, the set of load related parameters are mostly related to environmental factors, such as the wave height and wind velocity.

Table 2.2 Target reliability levels in design codes.

Eurocode 1		
Limit State	Annual	Lifetime
ULS	0.72×10^{-4}	0.13×10^{-5}
FLS	$0.67 \times 10^{-1} \sim 0.72 \times 10^{-4}$	-
SLS	0.67×10^{-4}	0.13×10^{-2}
Canadian Standards Association		
Safety Class	Consequences of Failure	P_{target}
1	Great risk of life or high potential for environmental pollution or damage	10^{-5}
2	Small risk to life and low potential for environmental pollution or damage	10^{-3}
3	Impairment function/in-serviceability	10^{-1}
Det Norske Veritas		
	Consequences	
Class of Failure	Less Serious	Serious
I – Redundant	10^{-3}	10^{-4}
II – Non-redundant structure, significant warning before failure	10^{-4}	10^{-5}
III – Non-redundant	-	-

However, the change in the ocean characteristics is generally slow such that the concept of sea state, which represents a short term sea condition, can be applied. The environmental condition is assumed to be stationary within a short duration in the ocean (Goda 2000). For example, the sea surface is normally considered to be stationary for a period of time from 20 minutes to 3 hours, represented by the significant wave height H_s and the peak period T_p . The response caused by these sea loads within this short term could be considered to be described by a random process $X(t)$. Then the evaluation of Eq. (2.11) corresponds to how likely the short term processes will exceed a load level. If the input stochastic processes are Gaussian and the system is linear, the statistical information contained in $X(t)$ can be easily computed. For most offshore dynamic problems, the response process may be non-Gaussian especially when one deals with extremes. The nonlinearity of the structure in the transfer of environmental process to the load process may cause a change in the statistical property from the input to the output. Analytical solution may thus be formidable and hence approximate or numerical analysis is required for each of the given stochastic environmental conditions. The overall exceedance probability is given by

$$P_E = \int_{\theta} Q(X(t) > l | \theta) f(\theta) d\theta \quad (2.12)$$

where $X(t)$ is the short term response conditioned on the ocean state parameters θ , $Q(\cdot)$ is the short term exceedance probability for a response over a load level l . To obtain the long term solution, assuming that the assumptions of stationary and independent and identical distribution (i.i.d.) are valid, the extreme value model

would be used for $Q(\cdot)$. A prediction of long term design value would then be estimated by setting a low exceedance probability, for example, 0.01 is normally used for a return level of 100 years.

2.4 Extreme Value Theory

Structures are designed for a service life span, and depending on the application, it can be from less than 10 to more than 100 years. It may not be possible to have 100 years of data for assessing the structure's performance over its design life. If the data are assumed to be stationary and iid, then the short term data may be used in conjunction with extreme value statistical models for the long term assessment of offshore design and operations. As introduced in Section 2.3, the design of offshore structures hinges on employing a critical value which has a low probability of exceedance, usually associated with a return period (Galambos 1994; Coles 2001; Castillo et al. 2004). Extreme statistics has emerged an important tool for many real problems over the past few decades, such as in public economics (Webb and Zank 2011), hydraulic engineering (Leviandier 2010; Dourte et al. 2013), ocean engineering (Morgan et al. 2011; Hodapp et al. 2013), non-linear beam analysis (Markovic et al. 2013), climate change assessment (Kollat et al. 2012) and medical engineering (Mano 2012). The basic properties of the extreme value model are presented in the following.

2.4.1 Asymptotic Model

The classical extreme value theory is based on the statistical behavior of block maximas (Gumbel 1958):

$$M_n = \max \{X_1, \dots, X_n\} \quad (2.13)$$

where X_1, \dots, X_n is a collection of independent random variables having the same common distribution F . The value of n could be taken as the number of a certain group of data to characterize the statistical properties of its extremes. Thus, the distribution of M_n is derived in relation to the original distribution F :

$$\Pr \{M_n \leq z\} = \Pr \{X_1 \leq z, \dots, X_n \leq z\} = \Pr \{X_1 \leq z\} \times \dots \times \Pr \{X_n \leq z\} = \{F(z)\}^n \quad (2.14)$$

The estimation of the maxima hinges on the distribution F . Unfortunately, the distribution of the observed data is generally unknown in reality. One possible way is to use standard statistical techniques to estimate F from observed data, and then substitute into Eq. (2.14). The error in F will propagate according to the power of n . Moreover in reliability analysis, the tail probability is of interest and hence the error in this probability may be large even with small discrepancies in the estimated F .

When the size of the block maxima approaches infinity, $n \rightarrow \infty$, Eq. (2.14) tends towards a stable function asymptotically if there exists sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that:

$$\Pr \{(M_n - b_n) / a_n \leq z\} \rightarrow G(z) \text{ as } n \rightarrow \infty \quad (2.15)$$

where G is a non-degenerate distribution function. Then G can be classified into one of the following types:

$$\text{I: } G(z) = \exp \left\{ -\exp \left[-\left(\frac{z-b}{a} \right) \right] \right\}, \quad -\infty < z < \infty \quad (2.16)$$

$$\text{II: } G(z) = \begin{cases} 0, & z \leq b, \\ \exp \left\{ -\left(\frac{z-b}{a} \right)^{-\alpha} \right\} & z > b, \end{cases} \quad (2.17)$$

$$\text{III: } G(z) = \begin{cases} \exp \left\{ -\left[-\left(\frac{z-b}{a} \right)^\alpha \right] \right\}, & z \leq b, \\ 1, & z > b, \end{cases} \quad (2.18)$$

where each type has parameters a and b as scale and location parameter, and α is the shape parameter in extreme types II and III. The three classes of distribution are termed the extreme value distribution with the widely known names of Gumbel, Fréchet and Weibull families, respectively. The three types of limit functions have distinct form of behavior, corresponding to the different forms of tail behavior of F rather than on the entire function F . The density of G is Weibull distributed if F has a bounded tail whereas it is Gumbel distributed if the tail of F decays exponentially and follows the Fréchet distribution if F has a polynomial decaying tail, see Figure 2.4. The three distributions above can be collapsed into the *Generalized Extreme Value* (GEV) Distribution:

$$G(z) = \begin{cases} \exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & \xi \neq 0 \\ \exp\left\{-\exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right]\right\} & \xi = 0 \end{cases} \quad (2.19)$$

defined on the set $\{1 + \xi((z-\mu)/\sigma)\} > 0$, where the parameters satisfy $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. The types I, II and III classes correspond respectively to $\xi = 0$, $\xi > 0$ or $\xi < 0$.

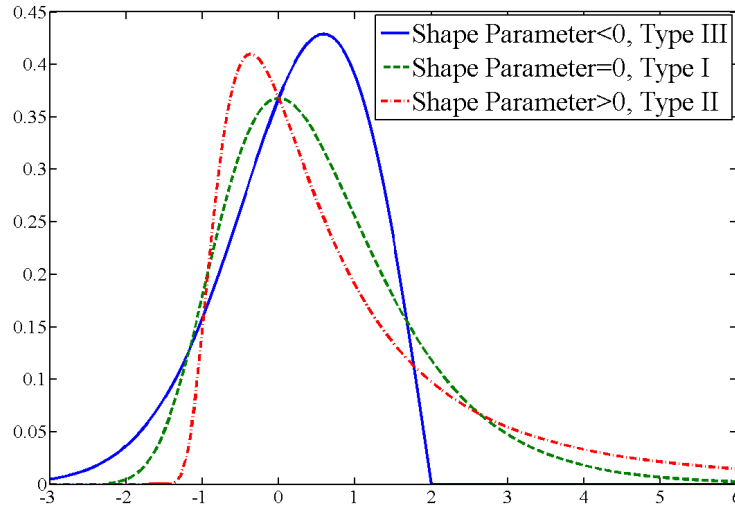


Figure 2.4 Illustration of different tail behavior in Type I ($\xi = 0$, $\mu = 0$, $\sigma = 1$), Type II ($\xi = 0.5$, $\mu = 0$, $\sigma = 1$) and Type III ($\xi = -0.5$, $\mu = 0$, $\sigma = 1$).

2.4.2 Inference for the Extreme Value Distribution

The estimates for long-term return value z_p , corresponding to an average return of once in every $1/p$ times the reference period, can be obtained by evaluating the appropriate quantile value in the GEV by setting $G(z_p) = 1 - p$. By substituting the parameter values, the return value can be calculated as:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} \left[1 - (-\log(1-p))^{-\xi} \right], & \text{for } \xi \neq 0, \\ \mu - \sigma \log(-\log(1-p)), & \text{for } \xi = 0, \end{cases} \quad (2.20)$$

By applying the Central Limit Theorem or the delta method (Oehlert 1992), the variance of z_p can be obtained as:

$$\text{Var}(z_p) \approx \nabla_{z_p}^T V \nabla_{z_p} \quad (2.21)$$

where V is the covariance matrix of (μ, σ, ξ) and

$$\nabla_{z_p}^T = \left[\frac{\partial z_p}{\partial \mu}, \frac{\partial z_p}{\partial \sigma}, \frac{\partial z_p}{\partial \xi} \right]$$

$$= \begin{bmatrix} 1 \\ -\xi^{-1} \left(1 - (-\log(1-p))^{-\xi} \right) \\ \sigma \xi^{-2} \left(1 - (-\log(1-p))^{-\xi} \right) - \sigma \xi^{-1} (-\log(1-p))^{-\xi} \log(-\log(1-p)) \end{bmatrix} \quad (2.22)$$

The return level is an approximation if the distribution is not binomial or normal. In cases where the distribution is significantly different from binomial or normal, then return level inferred may not be correct.

2.4.2.1 Annual Maximum Method

The annual maximum method, which has a block size of one year leading to the GEV distribution, is well-advocated by researchers (García-Ruiz et al. 2000; Winterstein et al. 2001, Morgan et al. 2011, Li et al. 2012). The design value corresponding to a return value can be obtained through an extreme distribution

plot for the observed data. For example, the analytical distribution function for the Type I Gumbel distribution, as shown in Eq. (2.16), can be generally expressed as:

$$G(z) = \exp\left\{-\exp\left[-k_o(z - z_o)\right]\right\} \quad (2.23)$$

Hence, theoretically, the value of z must have a linear relationship with $-\ln[-\ln(G(z))]$ such that:

$$k_o(z - z_o) = -\ln[-\ln G(z)] \quad (2.24)$$

The coefficients k_o and z_o could then be determined from linear regression based on the realized extreme value data z and its corresponding empirical cumulative distribution function value $G(z)$. The return level value would thus be simply estimated by the $1-p$ quantile from the regressed linear plot.

Although the annual maximum method is convenient and easy to apply, the time series data collected at the field may not always be sufficient for the proper statistical analysis. Moreover, useful information about the extremes are not used and are inherently discarded. The DNV code (2007) has emphasized that at least 20 years of data must be available for this approach.

2.4.2.2 r th Largest Order Statistics Model

The r th largest order statistics model provides an alternative way to utilize more extremes within a block or reference period (e.g. one year) for constructing the GEV model (Guedes Soares and Scott 2004, An and Pandey 2007, Zhao and Gu 2010). The limiting distribution function is defined for the r th largest value in a

sequence of independent and identically distributed random variables. This can be presented as:

$$M_n^r = \text{rth largest of } \{X_1, \dots, X_n\}, \quad (2.25)$$

Under the same asymptotic rule of extremes with the limit function as described in Eq. (2.15), the asymptotic function $G_r(z)$ corresponding to the order statistics can be expressed as:

$$G_r(z) = \exp\{-\tau(z)\} \sum_{s=0}^{r-1} \frac{\tau(z)^s}{s!} \quad (2.26)$$

where

$$\tau(z) = \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \quad (2.27)$$

When applying the r th largest order statistic model to a long time series, the data are grouped into several blocks. Let n denotes the number of blocks in the time series data with the largest r observations in each block, a total series of maximas can be obtained as $M_i^r = (z_i^1, \dots, z_i^r)$, for $i = 1, \dots, n$. The total likelihood of this model for all the series of data can then be expressed as

$$L(\mu, \sigma, \xi) = \prod_{i=1}^n \left(\exp \left\{ - \left[1 + \xi \left(\frac{z_i^{r_i} - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \prod_{k=1}^{r_i} \sigma^{-1} \left[1 + \xi \left(\frac{z_i^k - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi} - 1} \right) \quad (2.28)$$

The parameters determined for r th largest order statistics are strictly linked to the extreme value models. The return period level and other inference can be

estimated from the parameters evaluated in Eq. (2.28). In the special case where $r=1$ for each year, the model corresponds to the annual maximum method as described earlier.

The only difficulty in the application of the r th largest order statistic model is the selection of an appropriate value r . The number of selected extremes in each block often comprises a bias-variance tradeoff; small values of r generate few data leading to high variance whereas large values of r will violate the asymptotic theory in the extreme value distributions. Some of the previous works (Smith 2001) suggested a value of 5.

2.5 Concluding Remarks

Various available probabilistic models for the ocean environment parameters are discussed. Available numerical approaches in the structural probabilistic analysis have been summarized. Both simulation and transformation techniques used in the analysis of reliability problems are introduced. The simulation method is good at estimating the failure probability but however demands much efforts in the calculation. The transformation method could somehow require fewer calculations but may only give rough estimate of the failure probability. In the case of a complicated structural system, the approximation method like FORM and response surface method could be more suitable for engineering computations as it avoids much computations.

The long term performance of offshore structures and the assessment criteria is formulated based on the exceedance probability. Based on the results, a long term return value is usually employed in the structural design. Characterization of the long term structural response is realized by the extreme value theory when the maximum is of interest. A mathematical basis of the extreme values, regarding the method and inference, is shortly reviewed. Further applications and development will be investigated in Chapter 3.

Chapter 3 Establishing Robust Extreme Value Model

In this chapter, the peak over threshold (POT) method is examined in terms of its practical usage in relation the analysis of offshore structure with specific considerations to extreme events. The selection of the threshold and the issue of ensuring dependency between extreme events data points are discussed. In particular, de-clustering, mean residual life plot, L-moment plot and dispersion index plot are presented. A key element in the application of probability and statistical theory is the estimation of parameters. This chapter will compare the POT method with the other two approaches namely, annual maximum method and r th largest order statistic method for establishing an extreme statistical model. The performance of these methods are tested and compared through a numerical simulation study. The main focus of this chapter is to understand the performance and critical issues regarding applying the POT method as this method will be further utilized in the next chapters for assessing offshore structure's long term performance. This is also the reason to choose review POT method in this chapter instead of the previous chapter. In view of the subjective nature of some of the uncertainties, a random set based model that can handle imprecise probability is proposed to model the uncertainties associated with the selection of the threshold and the time to separate extremes from the original time series data in the POT method. It is shown that this utilization of Dempster-Shafer structure, or random set, in achieving a design value leads to improvements including a better understanding of imprecision in the observed data. The extended model is also

applied to different reference periods of the sampled data to evaluate the significance of the climatic conditions on the uncertainties of the parameters.

3.1 Introduction

Occurrences of extreme events have been known to be catalyst for the failure of constructed facilities. Hence, the prediction of extreme values is an important component in the design of coastal and offshore structures. As described in Chapter 2, design codes or standards have specifications with regards to the design values that is consistent with the design life (long term), often extrapolated from short term values based on statistical concepts. The theory of statistical extremes can be employed for this purpose but its underlying assumptions must be recognized and enforced. Observation data collected from a particular locality and environment do have its peculiarity (Kyselý et al. 2010) and some environmental data may present high serial correlation and large variations in magnitude for different reference periods. As such, the data are not independent and likely to be identically distributed, thus do not satisfy the assumption implied in statistical extreme theory (Menendez et al. 2008). Some form of data grouping may need to be done to ensure that each subgroup does fit the assumptions. If there is no subgroup that can satisfy the assumptions, then the extreme value model cannot be applied.

In practice, real data are used in establishing the model in relation to a specific application and without doubt uncertainties exist in various forms

depending on how sophisticated the model or analysis is. How these uncertainties are propagated to the end results, namely, establishing the return significant wave height level corresponding to a specified level of reliability over the design life of the offshore structure will be of primary interest. This will also help establish the need to gather more accurate information so that a consistent, acceptable and optimal design value leading to a safe and economical structure can be realized.

In this chapter, a numerical simulation based study is used to address the aforementioned concerns about establishing an extreme value model with regard to real practices in offshore engineering. In particular, the quality and performance of different parameter estimators in establishing an extreme value model is discussed when limited data is used. The issues of different kinds of data uncertainties including tail behavior, noise and range of dependencies in the time series will be investigated. The effect of extrapolation in the construction of a statistical model will also be discussed. This study also proposes using a random set model to account for the subjectivity in the selection of the threshold and time span, leading to an imprecise probability based quantification. The results obtained from this study of uncertainty assessment of performing POT and extreme value approaches serve to provide the basic knowledge for the rest of this dissertation.

3.2 Peak-Over-Threshold (POT) Method

Threshold exceedance model has been widely applied in the extreme value framework for environment parameters (Quek and Cheong 1992, Dukes and Palutikof 1995; Palutikof 1999 et al.; Smith 2001). The threshold approach is quite useful in treating and effectively utilizing time series data and has been widely employed in various applications (Ferreira and Guedes Soares 1998; Mackay et al. 2010; Teena et al. 2012; Petrov et al. 2013). It is suited for dealing with realizations of a stochastic process which is approximately stationary or can be split into stationary parts (Kyselý et al. 2010). It has this clear advantage over the other method, since it does not require the time series to be strictly stationary. Even time varying properties can be modeled in a POT approach, while the shape, scale parameters within the model can be treated as time varying functions (Kharin & Zwiers 2005; Parey et al. 2007). This will be addressed in Chapter 4 of this thesis.

3.2.1 Pareto Family

Consider a set of data extracted from an original set of data with probability distribution F such that their values are above a certain threshold value of u . Then the probability of a variable X exceeding a value z is given by the conditional probability (Pickhands 1975):

$$\Pr\{X > z | X > u\} = \frac{1-F(z)}{1-F(u)} = \frac{\left[1 + \xi \frac{z-\mu}{\tilde{\sigma}}\right]^{-1/\xi}}{\left[1 + \xi \frac{u-\mu}{\tilde{\sigma}}\right]^{-1/\xi}}, \quad z > u \quad (3.1)$$

If F also obeys the asymptotic rules mentioned in Section 2.4.1 for large n , the cumulative probability function for the exceedance can be expressed as:

$$G(z) = \begin{cases} 1 - \left[1 + \xi \left(\frac{z-u}{\tilde{\sigma}}\right)\right]^{-1/\xi} & \xi \neq 0 \\ 1 - \exp\left[-\left(\frac{z-u}{\tilde{\sigma}}\right)\right] & \xi = 0 \end{cases} \quad (3.2)$$

where ξ is the shape parameter, u is the threshold and $\tilde{\sigma}$ is the scale parameter which has a relationship with other parameters in the GEV model (e.g. Eq. (2.19)) as $\tilde{\sigma} = \sigma + \xi(u - \mu)$. Equation (3.2) belongs to the family of the *Generalized Pareto Distributions* (GPD). The concept parallels the GEV in the modeling of maxima, including the classification of the distribution into types I, II or III.

3.2.2 Poisson-GPD Model

In practice, the peaks over a sufficiently high threshold of time series data of engineering interest are usually rare and memoryless events. As such, their occurrences can be appropriately modeled as a Poisson process.

A point process defines a stochastic rule for the occurrence and position of point events on a set T , where T could represent time, geographical space, or even more general spaces. This can be used to describe the occurrence of extremes

within a reference period. The statistical properties of a point process can be defined by a non-negative integer-valued random variable, $N(A)$, for each $A \subset T$, such that $N(A)$ is the number of points in the set T . The probability distribution of each of the $N(A)$ characterises the point process. The expected number points in subset $A \subset T$ is given by

$$\Lambda(A) = E[N(A)] \quad (3.3)$$

and defines the *intensity measure* of the process. Considering time series data, the extreme value may be written as set $A = [x_1, t_1] \times \dots \times [x_n, t_n] \subset \mathbb{R}^n$, and the intensity function of the process is given by

$$\lambda(x) = \frac{\partial \Lambda(A)}{\partial x_1 \dots \partial x_n} \quad (3.4)$$

provided the derivative function exists.

3.2.2.1 Poisson Point Process

The most classical point process is the Poisson process which has two underlying conditions:

1. The number of occurrences $N(A)$ within a time period $A = [t_1, t_2] \subset T$ follows a Poisson's distribution:

$$N(A) \sim \text{Poi}(\lambda(t_2 - t_1)) \quad (3.5)$$

2. For all non-overlapping periods A and B , $N(A)$ and $N(B)$ are independent identical Poisson distributed random variables with mean value equals $\lambda(t_2 - t_1)$.

In other words, the number of occurrences for a given reference period of time is proportional to the increase of the time length. For Poisson process, the intensity function in Eq. (3.4) can be written as a time-dependent function $\lambda(t)$. Considering a family of parametric models $\lambda(\cdot, \theta)$, the expected intensity measure within a small time interval around the observation $I_i=[x_i, x_i+\delta_i]$ can be expressed as:

$$\Lambda(I_i; \theta) = \int_{x_i}^{x_i+\delta_i} \lambda(z) dz \approx \lambda(x_i) \delta_i \quad (3.6)$$

The probability function derived for a set of data exactly occurred at this set of observations $I_i=[x_i, x_i+\delta_i]$ for $i=1, \dots, n$, and no occurrence elsewhere, could be obtained in a likelihood function as:

$$\begin{aligned} L(\theta; x_1, \dots, x_n) &= \Pr \{ N(\mathfrak{T}) = 0, N(I_1) = 1, N(I_2) = 1, \dots, N(I_n) = 1 \} \\ &= \Pr \{ N(\mathfrak{T}) = 0 \} \prod_{i=1}^n \Pr \{ N(I_i) = 1 \} \approx \exp \{ -\Lambda(T; \theta) \} \prod_{i=1}^n \lambda(x_i; \theta) \delta_i \end{aligned} \quad (3.7)$$

where T is the set of total observations and $\mathfrak{T} = T \setminus \bigcup_{i=1}^n I_i$. Maximization of the likelihood function can give the parameter estimations θ in the intensity function $\lambda(\cdot, \theta)$.

3.2.2.2 Combined Poisson-GPD Model

An extension of the threshold model in Section 3.2.1 to characterize the extremes both in frequency and intensity can be obtained by a combination of the Poisson process and GPD model. The Poisson property of exceedances (Leadbetter et al. 1983) suggests the following model which is called the Poisson-GPD model:

- The number, N , of exceedances of the level u in any one block (e.g. one year) has a Poisson distribution with mean λ .
- Conditional on $N \geq 1$, the excess values Y_1, \dots, Y_i for $i=1, \dots, N$ are identical independent distributed following the GPD.

For $x > u$, the probability that the block maximum of the process just described is less than x is:

$$\begin{aligned} \Pr \left\{ \max_{1 \leq i \leq N} Y_i \leq x \right\} &= \Pr \{N = 0\} + \sum_{n=1}^{\infty} \Pr \{N = n, Y_1 \leq x, \dots, Y_n \leq x\} \\ &= e^{-\lambda} + \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \left\{ 1 - \left(1 + \xi \frac{x-u}{\sigma} \right)^{-1/\xi} \right\}^n = \exp \left\{ -\lambda \left(1 + \xi \frac{x-u}{\sigma} \right)^{-1/\xi} \right\} \end{aligned} \quad (3.8)$$

where

$$\tilde{\sigma} = \sigma + \xi(u - \mu), \quad \lambda = \left(1 + \xi(u - \mu) / \sigma \right)^{-1/\xi} \quad (3.9)$$

Thus the concepts of GEV and GPD models could be combined together for modeling the values above the GPD threshold, and moreover, Eq. (3.9) shows exactly how the Poisson-GPD parameters vary with the threshold u . From a mathematical point of view, this is the basic properties of extremes in stationary processes, which shows that under very general conditions, the magnitude of the exceedances can be modeled in a Pareto distribution while the occurrence rate approximately follows the Poisson process.

3.2.3 Declustering

In the POT approach, the exceedances must be regarded as independent and identically distributed variables. For some real events, the extremes may have

some degree of clustering, leading to the issue of dependency between exceedances above the threshold. To resolve this, *declustering* has been suggested, which is a process to filter the dependent values to obtain a set of threshold excesses that are approximately independent (Coles 2001). This is performed by the following steps:

1. Using an empirical rule to define clusters of exceedances.
2. Identifying the maximum exceedance within each cluster.
3. Assuming cluster maxima to be independent, with conditional exceedance distribution given by the GPD.
4. Fitting the GPD to the cluster maxima.

The method is simple but has limitations. In particular, results can be sensitive to the arbitrary choices made in cluster determination and there is arguably a wastage of information in discarding all data except the cluster maxima.

In the identification of maximum exceedance within each cluster, a possible way is to choose a time span Δt (e.g. 1 day, 3 days or 1 week), such that the extreme events separated by less than this period of time are considered as one “event”, and the highest value is identified as this extreme (Corti et al. 1997; Morton et al. 1997; Chen & Yoon 2002). This is chosen in an optimal way which is a minimum value to guarantee a persistent Poisson process for the extremes in the time series. The appropriateness of this time span can be checked from the basic Poisson process model where the interarrival time between the exceedances

should follow an exponential distribution (Luceño 2006). The whole procedure of applying POT to a time series data can be seen from Figure 3.1.

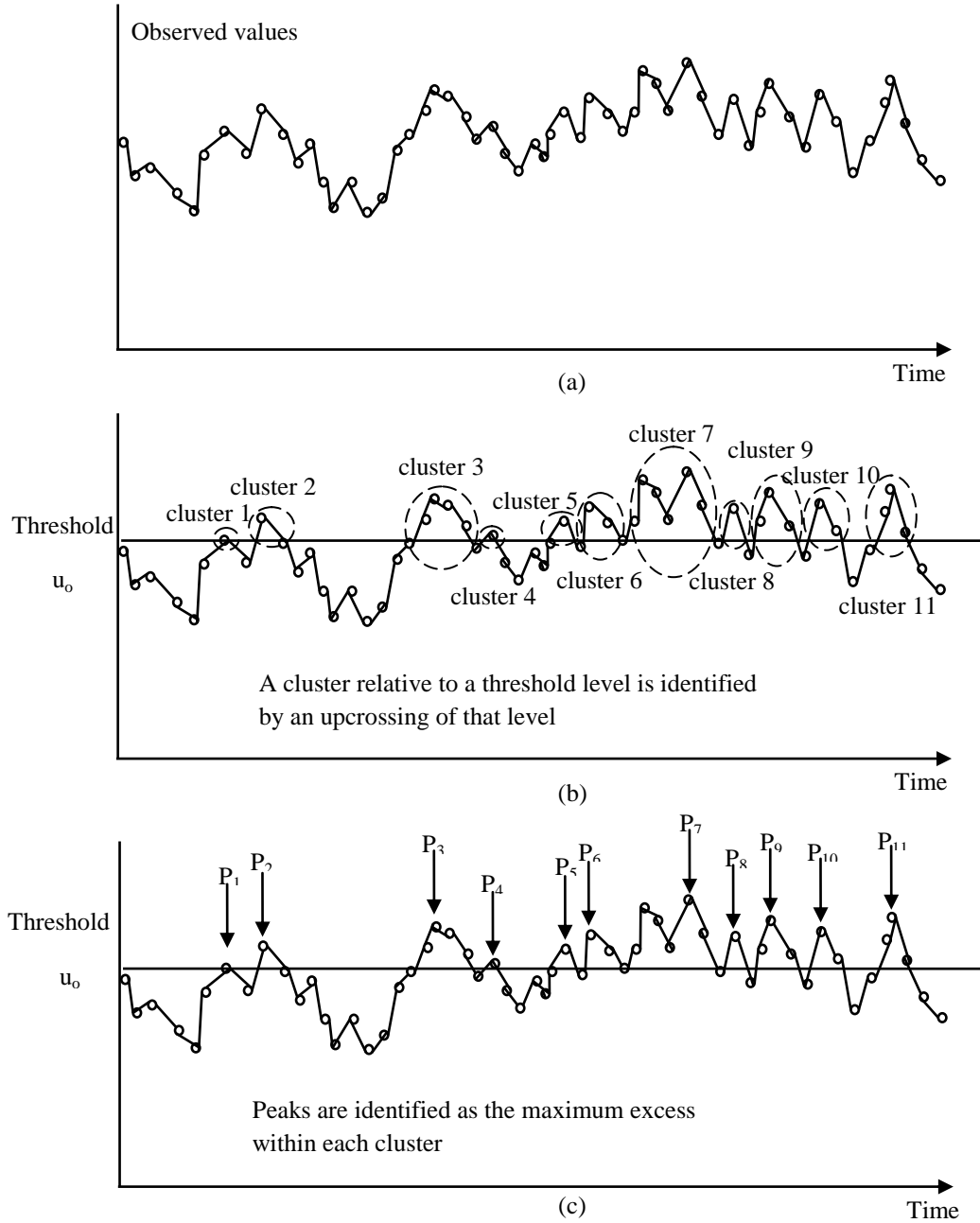


Figure 3.1 (a) Scatter plot of time series measurement (b) Clusters identification (c) Identified extreme values after declustering.

3.2.3.1 Model Check

The accuracy of the results obtained from the POT model depends on the fit of the Pareto distribution. The suitability of the chosen threshold needs to be checked using some criteria, such as through statistical checks of the fit to the GPD model. For example, the stability in the fitted distribution with slight changes in the threshold may be one factor to consider. The most common means to check the POT models are mean residual life plot, L-moment plot and dispersion index plot.

- *Mean Residual Life Plot*

The appropriateness of a threshold can be verified by the stability of the parameters of the Pareto distribution. One particular example is to use the mean residual life plot. The mean residual life plot, or sometimes called mean excess plot, is used to test the stability of the mean value of the exceedances with the change of threshold. Theoretically, the mean of a GPD can be estimated as:

$$E[x] = \mu + \frac{\tilde{\sigma}}{1-\xi}, \quad \xi < 1 \quad (3.10)$$

Since the estimated scale parameter is a linear function with the threshold, by substituting $\tilde{\sigma} = \sigma + \xi(u - \mu)$ in this function, the estimated mean function also becomes a linear one with a gradient of:

$$k = \frac{\xi}{1-\xi} \quad (3.11)$$

This consistent linear relationship between the mean excess and the threshold must be maintained for the possible thresholds.

- *L-moment Plot*

The L-moment theory is established based on order statistics that could give a measure to the properties of a distribution such as skewness and kurtosis. It parallels the theory of conventional moment, but more robust in the inference when there are outliers. The adequacy of the GPD fit to a sample data may be assessed through comparing the L-moment ratio (τ_4 / τ_3) against the fitted model, where the theoretical L-moment ratio of the GPD is approximately (Hosking 1990):

$$\tau_4 = \frac{\tau_3(1+5\tau_3)}{5+\tau_3} \quad (3.12)$$

By increasing the threshold, the plot will show a trajectory around the theoretical curve. An appropriate threshold is selected when the value of the L-moment ratio is close to the value given by Eq. (3.12).

- *Dispersion Index Plot*

The suitability of the threshold can also be tested by means of the dispersion index (DI) (Cunnane 1979):

$$DI = \frac{s^2}{\lambda} \quad (3.13)$$

where s^2 is the estimated variance of occurrence rate λ for the extremes in each block. The Poisson hypothesis is not be rejected when the dispersion index lies in

the confidence interval $\left[\chi_{\alpha/2, n-1}^2 / (n-1), \chi_{1-\alpha/2, n-1}^2 / (n-1) \right]$ for n blocks of sample data with a significance level of α .

3.2.4 Parameter Estimate Method

The importance of parameter estimation cannot be under-estimated as any error will be propagated in estimating the quantiles and long term design value. There are numerous parameter estimation methods such as the likelihood-moment estimations (Zhang 2007), least-squares error method (Moharram et al. 1993), empirical percentile method (Castillo and Hadi 1997), robust method (Peng and Welsh 2001) as well as Bayesian method (Zhang and Stephen 2009). However, most of these methods may not be easily implemented and some require intensive computations. The difficulties and disadvantages of using these methods have been discussed in De Zea Bermudez and Kotz (2010). Some of the better known GPD model parameter estimation methods are briefly summarized herein.

3.2.4.1 Method of Moments Method

The simplest method in estimating the parameters in POT could be the method of moments (MOM). The basic idea is to equate the sample mean and variance to the theoretical population mean and variance. For example, the theoretical expressions of the mean and variance for a random variable x following GPD(ζ, σ) model, assuming the threshold is given, are:

$$\begin{aligned}
E[x] &= \frac{\sigma}{1-\xi} && \text{for } \xi < 1 \\
\text{Var}[x] &= \frac{\sigma^2}{(1-\xi)^2(1-2\xi)} && \text{for } \xi < \frac{1}{2}
\end{aligned} \tag{3.14}$$

Based on these relationships with the GPD parameters, the MOM estimates are given by:

$$\xi = \frac{1}{2} \left(1 - \frac{\bar{x}^2}{\bar{s}^2} \right), \quad \sigma = \frac{1}{2} \bar{x} \left(\frac{\bar{x}^2}{\bar{s}^2} + 1 \right) \tag{3.15}$$

where \bar{x} and \bar{s}^2 stand for the sample mean and variance. Other alternative approaches, for example, by using the skewness or kurtosis in the sample estimate, can also be applied. However, the application of MOM requires a limiting value in the shape parameter, since a heavy tail GPD model may not have an estimate in the moment (the estimate of mean, variance, skewness and kurtosis will be infinity for shape parameter ξ larger than 1, 0.5, 1/3 and 1/4 respectively). Although the moments are easy to compute, the heavy tailed situation may increase the errors in the estimation. Such outliers existed in the data may cause certain distortions in the results.

3.2.4.2 Probability Weighted Moments Method

Based on a similar idea of the MOM, the probability weighted moments (PWM) method utilizes the sample PWM in estimating the parameters in the GPD. The PWM is defined as:

$$M_{p,q,r} = E \left[x^p (F(x))^q (1-F(x))^r \right] \tag{3.16}$$

where the p , q and r are real numbers. With this basic formula, the parameters of many distributions can be expressed as functions of PWM, rather than the traditional moments. It is especially handy if the simple inverse CDF is available. For example, the most useful and special moments that is used in showing the characteristics of a distribution are the following two cases:

$$\alpha_r = M_{1,0,r} = E \left[x(1-F(x))^r \right] \quad (3.17)$$

$$\beta_q = M_{1,q,0} = E \left[x(F(x))^q \right] \quad (3.18)$$

where q and r are non-negative integers. There is a close connection between these two PWM for any underlying distributions (Greenwood 1979). The number of PWM needed depends on the number of parameters of the distribution that needs to be estimated. By substituting the Pareto distribution into Eqs. (3.17) and (3.18), the PWM can be expressed as:

$$\alpha_s = E \left[x(1-F(x))^s \right] = \frac{\sigma}{(s+1)(s+1-\xi)}, \text{ for } \xi < 1, s = 0, 1, 2, \dots \quad (3.19)$$

By using the first two PWMs, the GPD parameters can be easily estimated as:

$$\xi = 2 - \frac{\alpha_0}{\alpha_0 - 2\alpha_1} \quad \text{and} \quad \sigma = \frac{2\alpha_0\alpha_1}{\alpha_0 - 2\alpha_1} \quad (3.20)$$

where α_0 and α_1 are the associated PWMs in Eq. (3.19). The values of these two quantities can be approximated by:

$$\alpha_s = \frac{1}{n} \sum_{i=1}^n x_{i:n} (1 - p_{i:n})^s \quad (3.21)$$

where $p_{i:n}$ is the plotting position which is general approximation to the value of $1-F$. An unbiased estimate is $p_{i:n}=(i-0.5)/n$, while for other cases, various expressions are available. For example, Hosking et al. (1985) suggested a biased estimate of $p_{i:n}=(i-0.35)/n$.

An advantage of using PWM and MOM is that they always exist and can be easily computed. Compared to maximum likelihood method, PWM and MOM will not need to encounter difficulties that may arise in the optimization step. However, the limiting values imposed on the shape parameter required in PWM method reduce the attractiveness of PWM (for example, PWM cannot be performed for $\xi > 1$).

3.2.4.3 Goodness-of-fit Method

Other than utilizing the specific statistical properties in estimating the parameter values, the goodness-of-fit method estimates the statistical parameters in the most obvious way, from a plot of the data. The result of a fitted parametric model should give the least sum of squares and must be visually compared against the empirical data plot, for example, quantile-quantile (QQ) plot. In the test statistics, the null hypothesis is $H_0: F(x) = F_0(x)$ where F is the empirical CDF and F_0 is the distribution being tested. Two of these well known statistics are:

Kolmogoriv-Smirnov (K-S) statistic:

$$D_n = \max \left\{ \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F_o(x_i) \right\}, \max_{1 \leq i \leq n} \left\{ F_o(x_i) - \frac{i-1}{n} \right\} \right\} \quad (3.22)$$

Anderson-Darling (A-D) statistic:

$$A_n^2 = - \sum_{i=1}^n \frac{2i-1}{n} \left\{ \log F_o(x_i) + \log(1 - F_o(x_{n+1-i})) \right\} - n \quad (3.23)$$

The K-S test measures the maximum discrepancy between the theoretical model and the empirical data whereas the A-D test places more weight or discriminating power on the tails of the distribution. The critical values of these statistics (e.g. D_n^α) for a significance level α to accept a parametric model are tabulated in standard statistics textbooks, for example, Ang and Tang (1984). Theoretically, the smaller the statistic is, the better is the fit. Thus, the estimators for the Pareto model could be obtained by minimizing these statistics (Luceno 2006). Other similar approaches like minimizing the sum of square of the differences between empirical and theoretical CDF have also been proposed (Moharram et al. 1993), but optimization is involved which requires much more computational effort.

3.3 Uncertainty Assessment of POT Method

In selecting the POT method to model ocean data, numerous factors which affect the accuracy of the results of interest need to be considered, such as the length and number of data available, the criteria used to identify extremes, the choice of

threshold and dependency effects (Mackay et al. 2011). These issues in extreme value statistics applications have been addressed previously (Ashkar and Tatsambon 2007; Deidda and Puliga 2009; Ribereau et al. 2011). However, these studies only examined comprehensively the uncertainty conditions; for example, many consider the effects of limited sample size but ignore other characteristics inherent in the collected data sample, such as serial correlations. The performance of POT method by using different parameters, like the estimator and threshold, is highly dependent on the collected data.

The uncertainties associated with the POT method is quantified herein. The performance of the POT method is first examined using Monte Carlo simulations considering different parameter estimation methods, sample size, tail effects and noise. The former include the method of moments (MOM), maximum likelihood method (MLE), unbiased probability weighted moments method (PWMU: $p_{i:n}=(i-0.5)/n$ in Eq. (3.21)), biased probability weighted moments method (PWMB: $p_{i:n}=(i-0.35)/n$ in Eq. (3.21)), A-D test based goodness-of-fit method (AD), and the K-S test based goodness-of-fit method (KS). The effect of sample size on the GPD model parameters are investigated using simulated data with sample sizes of $n = 10, 20, 30, 50, 80, 100, 150$ and 200 . For data sample size n , the simulation will be repeated 10,000 times and their average used as a means for comparison. To investigate the tail and noise effects, the threshold is set to the location parameter in fitting the GPD model. The importance of threshold selection will be discussed in Section 3.4.2.

The results of interest are the shape parameter, scale parameter and a high percentile value (for this purpose, a non-exceedance probability of 0.99 is used). The accuracy of the estimators are compared using a normalized measure of deviation from the theoretical value, herein denoted as the relative bias, as:

$$\mathbf{Relative\ Bias}[\hat{\theta}] = \frac{E[\hat{\theta}] - \theta}{\theta} \quad (3.24)$$

where θ is the parameter of interest and $\hat{\theta}$ is the parameter estimator. Here, we set the value of the parameter for different case studies. That is, the exact “true” value of θ is known. And each estimate $\hat{\theta}$ from one set of simulations is considered as one realization. The average of all the estimates from the simulations is used to compare with the “true” value to determine how much bias is about this estimation according to different methods.

The simulation is performed in R-programming with the help of open source code packages “fields”, “evd”, “evdbayes”, “ismev”, “SpatialExtremes” and “POT”. The detailed programming code and the verification of these packages have been documented by Ribatet (2009; 2010) and Stephenson (2002; 2010a; 2010 b).

3.3.1 Effects of Tail Behavior

As shown in Chapter 2, the tail characteristics, or the value of ζ , of a GPD model can critically influence the parameter estimations, which in turn will affect the

expected return values. Theoretically, the GPD is valid for any value of ξ . However, not all the estimation methods will yield estimates that can cover the entire range of possible values of ξ in a GPD model.

To illustrate this, the simulated data based on a GPD model for $\xi = -0.5, -0.25, 0, 0.25$ and 0.5 with $\sigma = 2$ and $\mu = 1$ are used to compute the scale and shape parameters. This range of values in $\xi \in [-0.5, 0.5]$ is commonly observed for environmental variables, such as the significant wave height (Beguer á 2005). The relative bias in the shape and scale parameters with respect to sample sizes from $n = 10$ to 200 using various estimated methods are presented in Figs. 3.2 and 3.3 respectively. The relative bias in estimating the values corresponding to the 99th percentile for $n = 20$ and 100 are tabulated in Table 3.1. The detailed results are recorded in Appendix B.

The findings based on the simulated results can be summarized as follows:

- Generally, the relative bias of each estimator decreases with increasing sample size. However, for the heavy tails, the bias of KS and MOM in estimating the shape and scale parameters is still large even the sample size is increased to 200 .
- The change of tail behavior from light tail ($\xi < 0$) to heavy tail ($\xi > 0$) will increase the relative bias in the shape parameter, scale parameter and 99th percentile estimations for all the estimators.

- Amongst the estimators, the MLE is most sensitive to sample size. It produces the largest bias estimate for data that has a light tail, for all the sample sizes considered.
- MOM estimator is the most sensitive to the tail behavior. For a heavy tail that has a value of ζ around 0.5, the MOM estimator is not suitable as the bias is about 30% and does not significantly improve by increasing the sample size.
- PWMU and PWMB show consistent good parameter estimates for different tail behavior compared to the other estimators. However, compared to PWMB, PWMU is a slightly less sensitive to the effects of tail behavior and sample size.
- AD gives a low bias in the estimated shape and scale parameters and is not sensitive to the change of tail behavior. However, AD estimators are quite sensitive to the sample size effect. The bias can go up to 20% for a sample size around 10.
- KS estimators give large bias in estimating the shape and scale parameters for $\zeta > 0$. But for $\zeta < 0$, the bias is relatively not significant. The performance of KS estimator is very poor in estimating the high percentile value for a data set that has a heavy tail or small sample size.

Obviously, it is not easy to tell which estimator is the best one for all conditions as shown by the results. However, for a data set with $\zeta < 0$, MOM, AD, PWMU and PWMB are reasonably good estimators for GPD model parameters and 99th percentile for the problem of interest as the relative bias are fairly small

(<10%) even for a sample size of 20. If the sample size is greater than 100, the MLE is a suitable alternative.

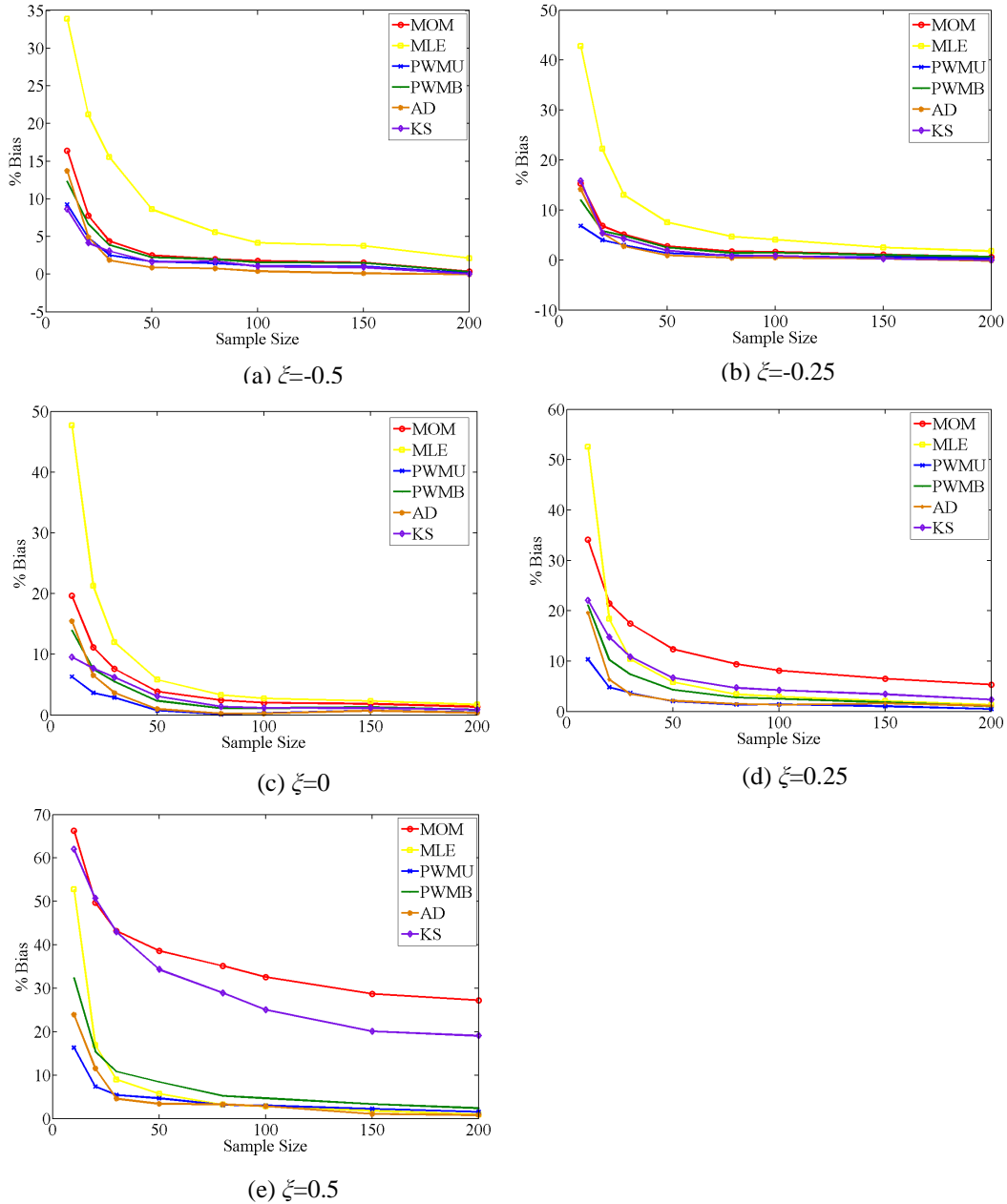


Figure 3.2 Bias of estimated scale parameter for different tails.

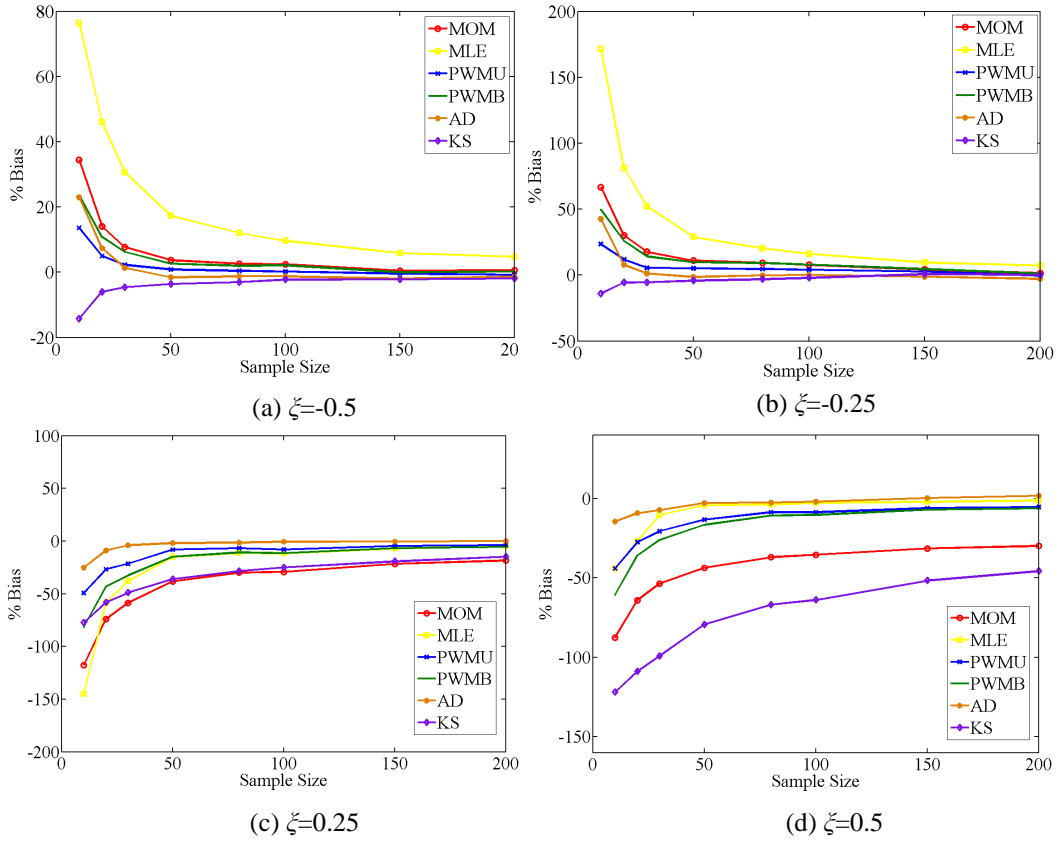


Figure 3.3 Bias of estimated shape parameter for different tails.

Table 3.1 Bias of estimated 99th percentile for different tails $n=20, 100$.

	$\xi=-0.5$	$\xi=-0.25$	$\xi=0$	$\xi=0.25$	$\xi=0.5$
$n=20$					
MOM	0.02%	-1.95%	-7.02%	-15.51%	-27.54%
MLE	-7.28%	-8.07%	-5.75%	70.54%	24.76%
PWMU	4.41%	4.53%	2.90%	0.18%	-12.29%
PWMB	1.98%	0.90%	-1.72%	-4.42%	-14.12%
AD	2.87%	9.00%	20.86%	56.52%	79.08%
KS	19.14%	25.76%	30.62%	55.23%	14.3%
$n=100$					
MOM	0.01%	-0.74%	-1.49%	-8.64%	-16.79%
MLE	-2.05%	-2.64%	-1.89%	-2.51%	6.69%
PWMU	0.88%	0.62%	0.83%	-2.28%	-2.90%
PWMB	0.46%	-0.05%	-0.11%	-3.32%	-3.58%
AD	0.86%	1.77%	3.42%	4.25%	14.64%
KS	4.11%	7.58%	5.18%	-0.71%	-11.07%

However if $\zeta > 0$, both PWMB and PWMU stand out as the best estimation methods. If $n > 100$, the AD and MLE estimators can be adopted in view of their small bias for a large size of data sample. The results obtained here can be generalized to other data analysis problems since this investigation is performed based on purely Monte Carlo simulations.

3.3.2 Effects of Noise

Another common uncertainty arises from noise in the collected data. For example, as most of the environmental data collected at a site is not enough, the data collected at a nearby site may also be utilized together for the statistical analysis. This is also called site averaging effect (Elsignhorst et al. 1998). However, this procedure introduces some non-stationary data into the data group for the analysis since these two sites may not be perfectly the same and discrepancies are expected. Then, it may cause some variations in the statistical parameters of the GPD when the POT method is employed (Jonathan and Ewans 2007).

The effect of noise on the parameter estimates in the GPD is investigated in this study by polluting the simulated data with Gaussian noise. Noise are added to the parameters of the GPD having $\zeta = -0.5$, $\sigma = 2$ and $u = 1$ before the data are simulated. The following cases of noise are simulated:

- Noise in location parameter: $u=1+N(0,\varepsilon^2)$, $\sigma=2$, $\zeta=-0.5$ for $\varepsilon=0.1, 0.3, 0.5$
- Noise in scale parameter: $u=1$, $\sigma=2+N(0,\varepsilon^2)$, $\zeta=-0.5$ for $\varepsilon=0.2, 0.6, 1.0$
- Noise in shape parameter: $u=1$, $\sigma=2$, $\zeta=-0.5+N(0,\varepsilon^2)$ for $\varepsilon=0.05, 0.15, 0.25$

where $N(0, \varepsilon^2)$ is a value drawn from a standard Gaussian distributed random number generator having a mean of 0 and variance equals to ε^2 . Three noise intensity ε^2 are chosen, corresponding to coefficients of variation of 0.1, 0.3 and 0.5. This considered range of the noise intensity in this study has been reported by Jonathan and Ewans (2006) for the ocean parameter such as wave height.

The calculated bias estimates to the shape and scale parameters are presented in Figs. 3.4-3.6. The bias in the estimated 99th percentile for $n = 20$ and 100 are shown in Table 3.2. The detailed results are given in Appendix B. Comparison of the results yields the following conclusions:

- The noises in the location parameter yield the largest bias compared to the noises in the scale and shape parameters. While the effect of the noises in scale and shape parameters can be reduced by increasing the sample size, the bias for large noises in the location parameter cannot be reduced, at least between $n = 20$ and 100.
- All the parameters in the GPD experience increase in relative bias with increase in noise intensity irrespective of the parameter estimation methods, with the location parameter being the most affected. The 99th percentile estimate is very sensitive to the noise especially from the shape parameter.
- MLE gives the largest relative bias when noises are present in the scale and shape parameters. However, MLE is relatively the best estimator when noise occurs in the location parameter. It also produces good results for the 99th percentile as the bias for all the cases are less than 10% even for a small

sample size $n=20$. The quality of MLE method is highly sensitive to the sample size.

- MOM, PWMU and PWMB produce similar results, giving large relative bias with noise in the location parameter but low bias with noises in the scale and shape parameters. None of the parameter estimation methods are able to give reliable results when the noise in the location parameter is very high,
- Among all the estimators, AD shows the best performance if there is noise in the location parameter and very good estimates if there are noises in the shape and scale parameters. However, for noises in the shape parameter, AD gives a large bias in the 99th percentile estimate.
- KS method gives the largest relative bias in estimating 99th percentile for all the noise effects. It gives a negative bias in the shape parameter estimate for the noises in the scale and shape parameters.

The effects of noise are clearly not insignificant and the parameter estimation method needs to be carefully selected. For noise in the location parameter, AD would be the most suitable method in estimating the parameters in the GPD model. If the intensity of noises in location parameter is high ($\varepsilon > 0.5$ in this study) and the sample size is not small ($n > 100$), the MLE method is another good choice. However, one should note that the use of MLE and AD when the noise is large still give large relative bias ($> 30\%$). For noises in the shape and scale parameters, MOM, PWMU, PWMB and AD are all applicable as long as sample size is not too small ($n > 20$). However, if the noise in the shape

parameter is high ($\varepsilon > 0.25$ in this study), the estimated 99th percentile may still have a large bias ($\approx 10\%$) even though the sample size is 200.

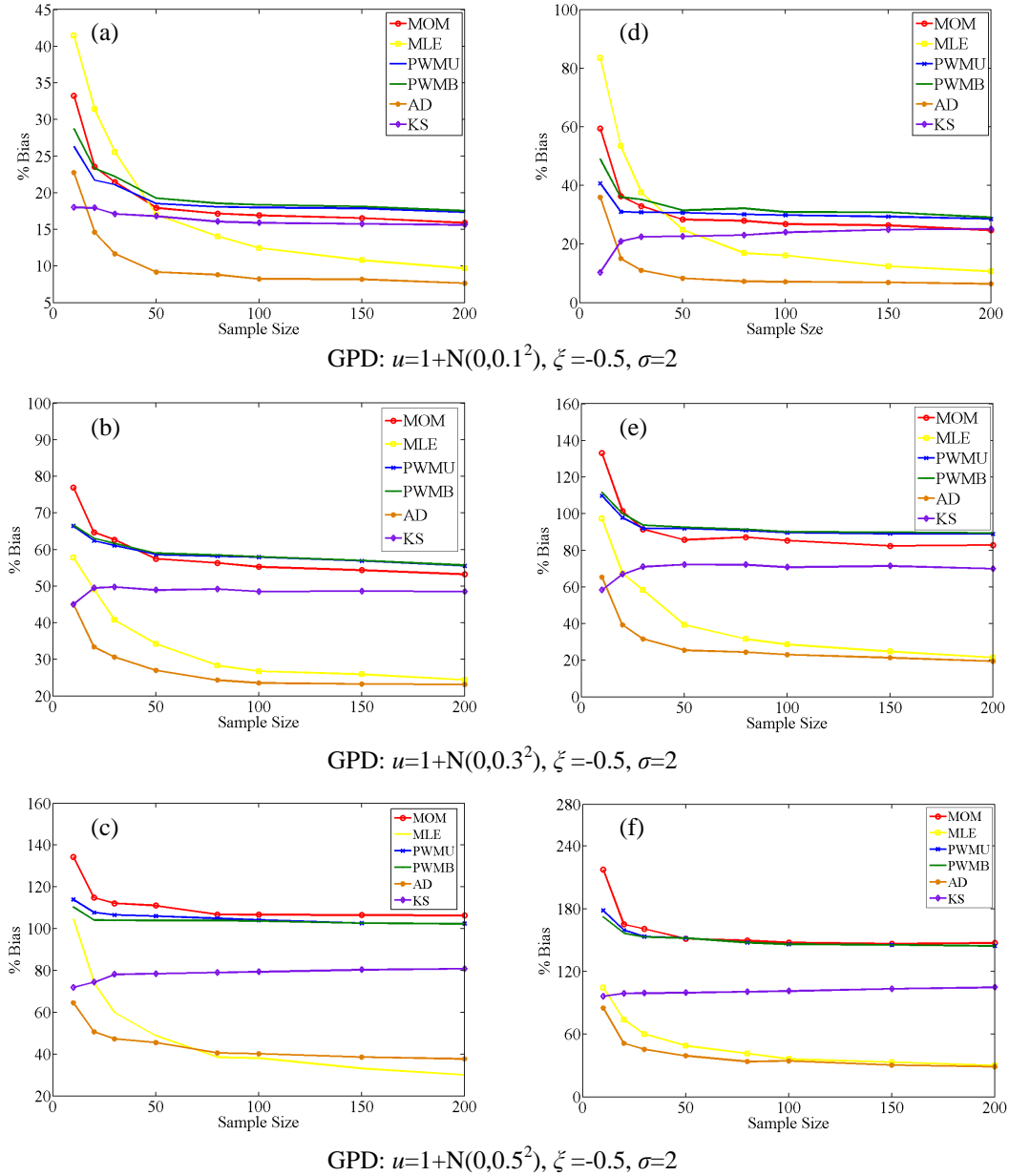


Figure 3.4 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in location parameter in GPD model.

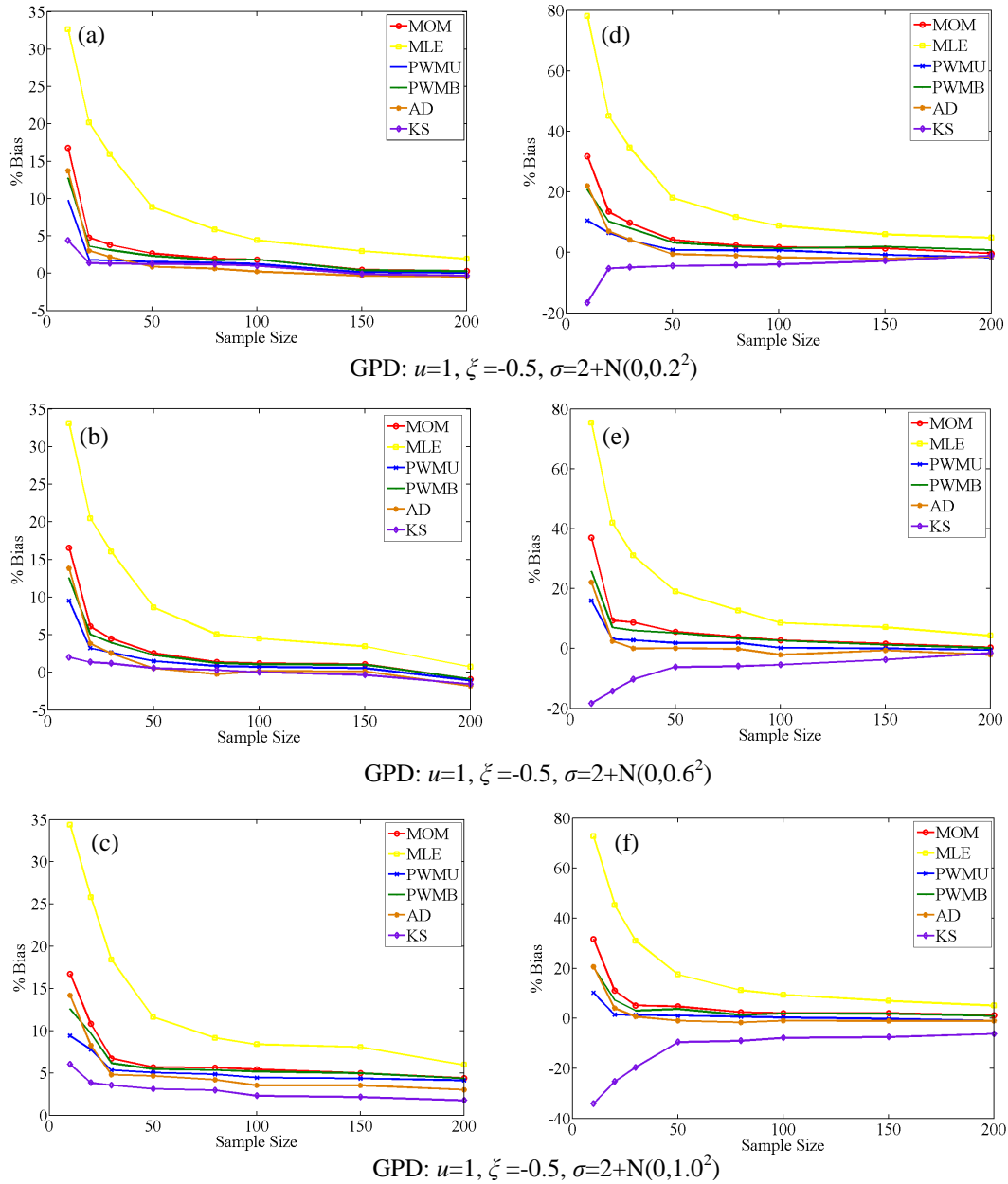


Figure 3.5 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in scale parameter in GPD model.

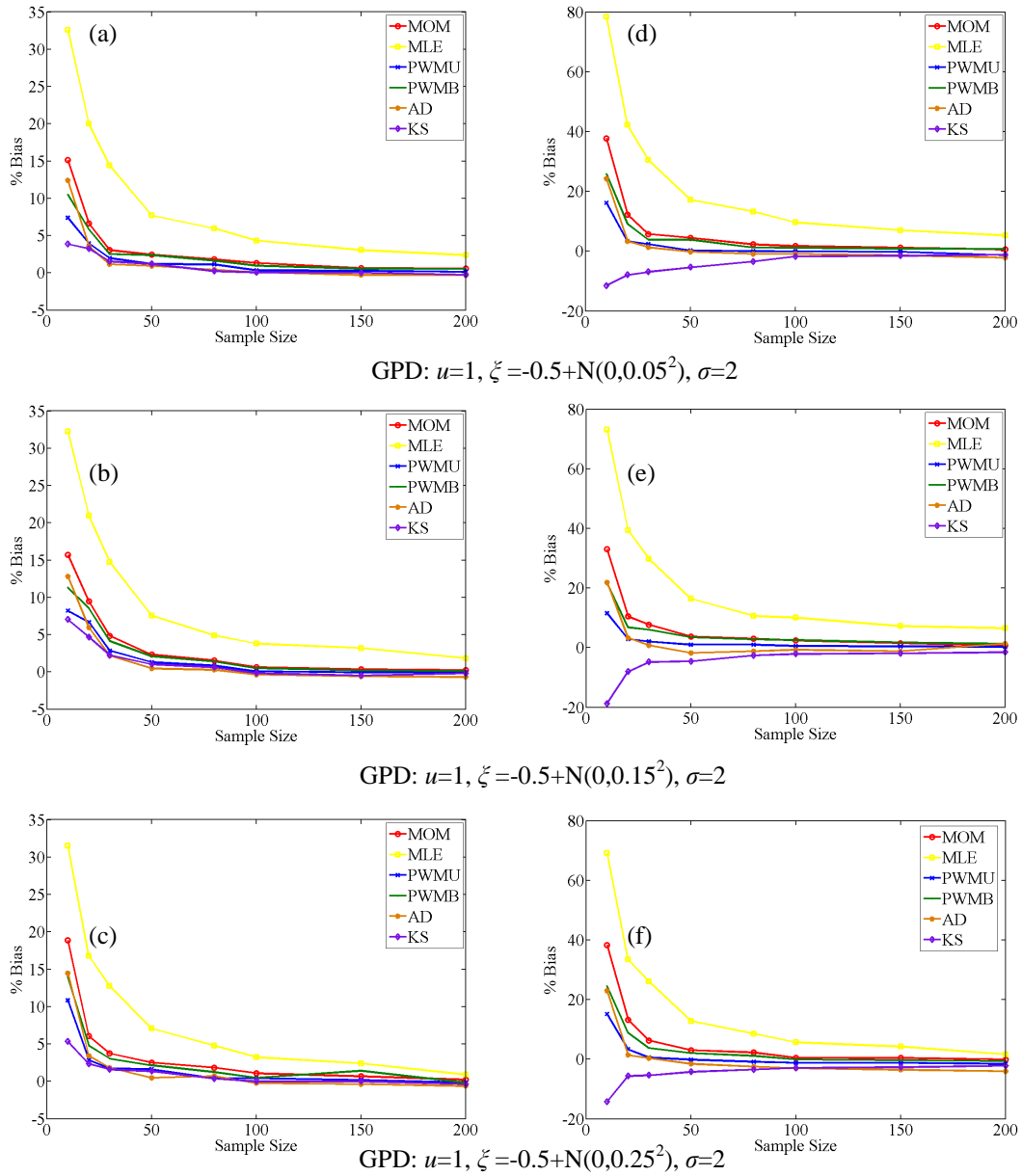


Figure 3.6 Bias of estimated scale parameter ((a), (b) and (c)) and shape parameter ((d), (e) and (f)) with the noise effect in shape parameter in GPD model.

Table 3.2 Bias of estimated 99th percentile for different noise condition $n=20, 100$.

	Noise in u		Noise in σ		Noise in ξ	
	$\varepsilon=0.1$	$\varepsilon=0.5$	$\varepsilon=0.2$	$\varepsilon=1.0$	$\varepsilon=0.05$	$\varepsilon=0.25$
<i>n=20</i>						
MOM	-2.09%	-6.22%	-0.22%	5.70%	0.48%	8.22%
MLE	-5.97%	0.19%	-7.81%	-2.28%	-7.06%	2.21%
PWMU	0.35%	-5.55%	4.43%	10.47%	5.15%	14.13%
PWMB	-1.49%	-6.13%	1.99%	7.85%	2.64%	11.03%
AD	2.58%	5.14%	3.74%	9.89%	4.34%	16.85%
KS	14.15%	18.56%	19.07%	36.80%	23.13%	35.43%
<i>n=100</i>						
MOM	-2.56%	-6.92%	-0.38%	0.50%	0.53%	10.14%
MLE	-0.53%	7.03%	-2.21%	-1.57%	-1.46%	8.56%
PWMU	-3.11%	-7.51%	0.40%	1.35%	1.36%	11.17%
PWMB	-3.41%	-7.59%	-0.01%	0.93%	0.93%	10.60%
AD	1.41%	7.68%	0.62%	1.38%	1.44%	11.92%
KS	0.25%	11.40%	3.09%	7.68%	4.16%	14.50%

3.3.3 Effects of Range of Dependency

The issue of dependency between points in time series data needs to be investigated. A potential cause is when the time interval between points are close, in which case, either a correction is needed or the data used must be separated by specified interval to reduce the serial correlation. Another cause is when an extreme event takes place and induces subsequent significant events.

To understand the effects of dependency, an autoregressive model (AR) of order one is utilized to simulate a weakly stationary and dependent time series. It is basically a linear relationship between the current and the previous value in time t , given by:

$$X_t = c + \varphi X_{t-1} + \varepsilon_t \quad (3.25)$$

where φ is the parameter of the model, c is a constant and ε_t is the noise term. For comparison purpose, the values of φ are set to 0.95 and 0 which correspond to a highly-correlated and an uncorrelated time series, respectively. Their difference is best depicted through their autocorrelation function (ACF) as plotted in Fig. 3.7.

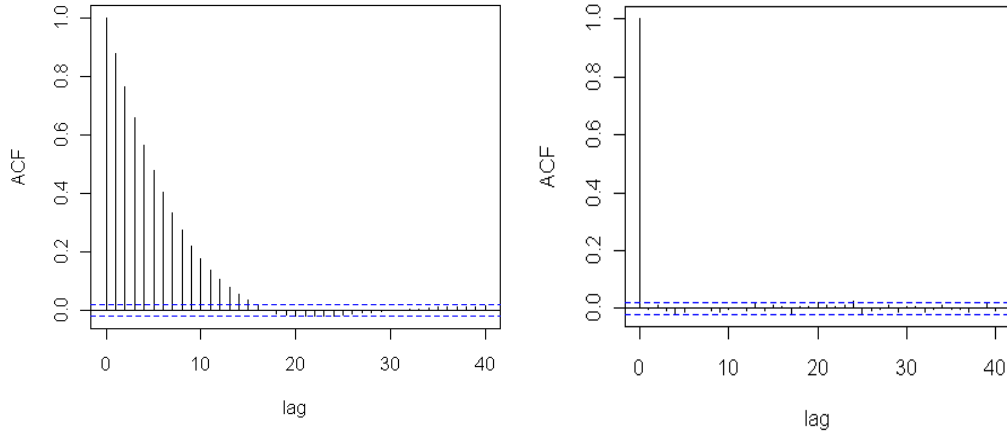


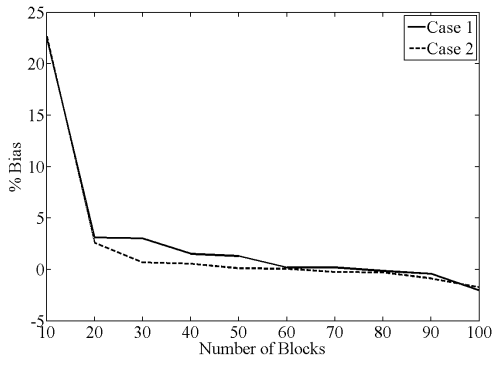
Figure 3.7 Plots of autocorrelation functions for two tested time series (a) Case 1: $\varphi=0.95$ (b) Case 2: $\varphi=0$.

In this study, a value of 0 is given to c and ε_t is assumed to follow an exponential distribution which has a rate parameter equals to 1 (that is, $\varepsilon_t \sim \text{Exp}(1)$). Theoretically, the parameters of the extreme value model for these simulated time series have the corresponding values $u=4.605$, $\sigma=1$ and $\zeta=0$. For each simulated realization, a group consisting of 100 continuous time series data is defined as a block. The block is used here to represent a reference time unit (for example, the AMM will only utilize the maximum value within each block). In order to test the estimates for different lengths of time series, the data simulated will have lengths of 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 blocks for comparison purpose. The results of interest here is to estimate the 99th percentile

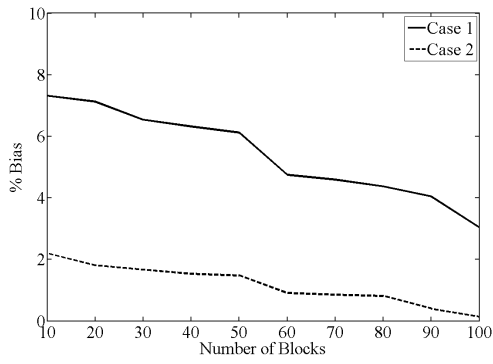
from each simulated time series. Each estimate is calculated based on an average of the results from 100 simulations.

Data from the two φ cases of time series are employed to estimate the 99th percentile using the (a) Annual maximum method (AMM), (b) r largest order statistic method and (c) POT method are applied. Within the r largest order statistic method, four values of r are considered, namely, $r = 5, 10, 15$ and 20 . Within the POT method, four different values of threshold and time span are used, denoted as U3T0, U3T10, U5T0 and U5T10, where the notation U_iT_j refers to a threshold value of i in identifying the excess values, and j represents the value of time span (number of continuous time series data) used in de-clustering the extremes as discussed in Section 3.2.3. The results in terms of bias in estimating the 99th percentile for each time series are plotted in Figs. 3.8-3.9. The findings based on the simulated results are summarized as follows:

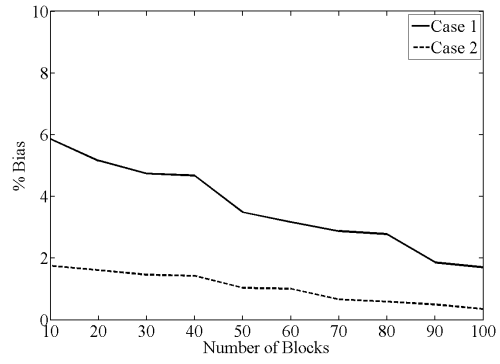
- Compared to POT and r largest order statistic methods, AMM is least affected by dependencies within the time series with regards to the 99th percentile value, provided the sample size is larger than 20 blocks. For example, for time series that only have 10 blocks, the bias of estimated 99th percentile is quite large (>20%). This is because the AMM filters out only a small amount of data (only the maximum value within each block is filtered) in the time series. When the number of blocks (for example, it could be units of years or months) is limited, the statistical uncertainty resulting from small sample size is high.



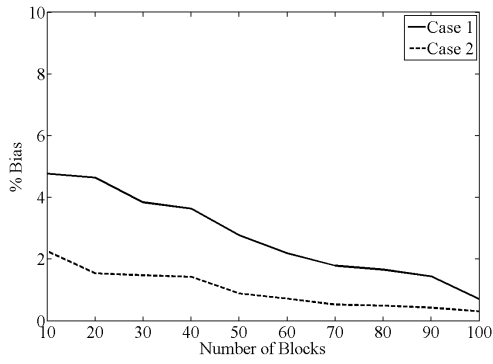
(a) AMM



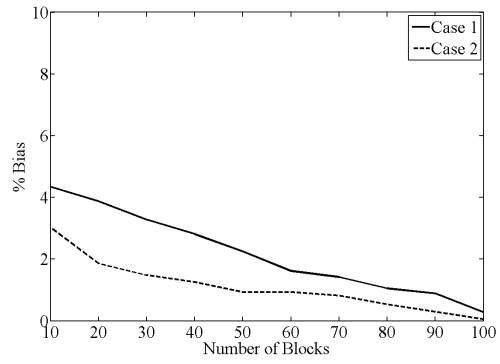
(b) 5 largest



(c) 10 largest



(d) 15 largest



(e) 20 largest

Figure 3.8 Biases of estimated 99th percentile in AMM and r largest order statistic method for two time series. (a) Case 1: $\varphi=0.95$ (b) Case 2: $\varphi=0$.

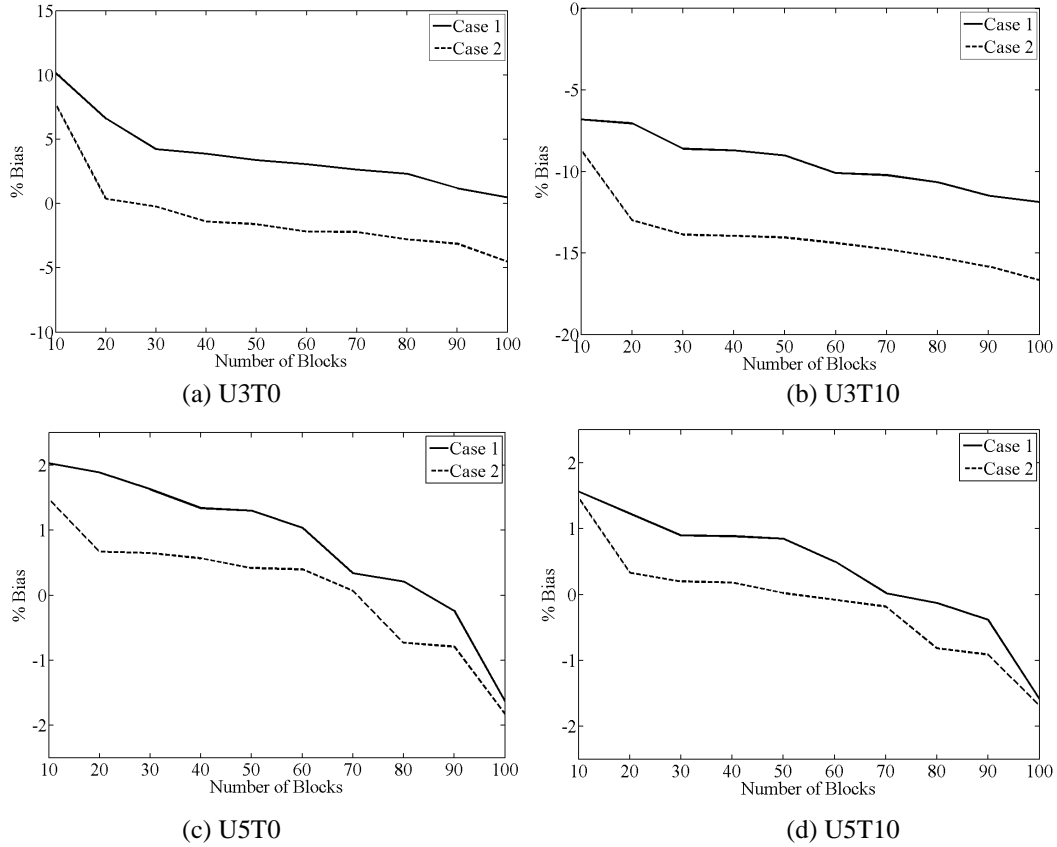


Figure 3.9 Biases of estimated 99th percentile in POT method for two time series (a) Case 1: $\varphi=0.95$ (b) Case 2: $\varphi=0$.

- The r largest order statistic method filters out more data per block and hence the statistical uncertainty is smaller. However, it is more sensitive to the dependency that may exist within the time series. This is particularly obvious when r is small where fewer data are filtered. For example, when $r = 5$ and only 10 blocks of data are available, the estimated biases in case 1 is much higher (7.3%) compared to case 2 (2.2%). These biases can be reduced by utilizing more data within the block. For instance, for $r = 20$, the bias estimate for case 1 is much less than for $r = 5$. But it does not imply utilizing more data within the block is always helpful because the assumption of

asymptotic property in Eq. (2.26) is violated for large r . For this reason, the results using the 20 largest order statistics is less accurate compared to the 10 largest order statistics in the case of uncorrelated time series.

- The performance of POT is quite dependent on the given values of threshold and time span. As shown in Fig. 3.9, U3T0 gives a large positive bias, while U3T10 gives a large negative bias. However, if the threshold changes to 5, the error associated with the estimations in U5T0 and U5T10 are very small (-2%~2%). This implies that the threshold value of 3 is too small and is not suitable value for use with the POT method. It is noted that sample size has lesser influence on the results from POT method, as it filters more data compared to AMM and r largest order statistic method. It can be seen from the comparison between (c) and (d) in Fig. 3.9, the use of time span in U5T10 leads to a smaller bias compared to U5T0. However, the dependency in the time series has very little influence to the accuracy in POT (the difference between case 1 and case 2 in Fig. 3.9 (c) and (d) are quite small) and only leads to a small positive bias in the estimate.

The main conclusion from this exercise is that each method has its own advantages and disadvantages in establishing an extreme value model for the time series data. AMM has very good performance when there is a large amount of data and it is not affected by the dependency effect in the time series. The r largest order statistic method does not need a large amount of data compared to AMM, but it is not suitable for highly correlated time series. POT gives the most desirable results even for time series that has high serial correlations. However,

the accuracy of performing POT method is quite sensitive to the selected values of threshold and time span.

Unfortunately, the current available approaches can hardly find out the “most” appropriate values of threshold and time span from the sample data. The use of model check plot to find suitable threshold and time span values as described Section 3.2.3.1 may only give rough estimation. This is demonstrated in the following sections.

3.4 Effects of Nonstationarity through Random Set Approach

As illustrated in the previous section, the POT method is quite efficient in utilizing data compared to other methods. However, the choice of both the threshold and the length of data (or time span of data) have a large impact on the quality of the exceedances as illustrated by the simulation study in Section 3.3.3. The uncertainties associated with the selected threshold and time span are important quantities that affected the robustness of the POT method. As discussed in Section 3.2, most of the methods to select the threshold and time span are based on expert judgment, which can be highly subjective. In fact, there may be no known true answers as any non-stationarity within the time series implies that the parameters are time-varying instead of constant. Therefore, this type of uncertainty cannot be eliminated and may perhaps not be appropriately treated in a traditional probabilistic way. A better alternative to model these uncertainties is proposed in this study, namely to use a random set based model. Random set model, which is developed from evidence theory (Shafer 1976),

provides a suitable basis for the consideration of imprecise observations with a random background.

In this section, the basic theory of random set and evidence theory will be briefly introduced. A wave data set taken from a U.S. project is then used to demonstrate a random set based extreme value model to obtain the long term design value. To understand the influence of different climate conditions on the uncertainties, the results from various defined reference periods are compared.

3.4.1 Review of Random Set and Dempster-Shafer Structure

3.4.1.1 Random Set

The basic probability theory states that, for discrete random variable x_i in the space X , each observation is associated with a non-zero probability mass $p_x(x_i)$. Similarly under the Dempster-Shafer structure (Dempster 1967; Shafer 1976) in evidence theory, random sets are a collection of many imprecise observations A_i , $i=1, 2, \dots, n$ of a given fundamental set X , called the focal sets, with basic assigned probability weight $m_i = m(A_i)$, $\sum m(A_i) = 1$ as a measure of the degree of confidence. This can be formally expressed as follows (Nguyen 2006):

Let X denotes a non-empty set containing all the possible values of a variable. A finite random set on X can be defined as pairs (\mathfrak{A}, m) , where $\mathfrak{A} = \{A_i : i=1, \dots, n\}$ represents finite subsets, and m represents the mapping such that for all the sets belonging to the power set of X , there is a mass assignment:

$$m: \mathfrak{S} \rightarrow [0,1] \quad (3.26)$$

Because the subsets may overlap each other, the evidence theory does not comply with the traditional probability theory. However, the random set can be understood as a generalization of traditional probability theory when measures for the occurrence of an event E in the space of X are needed. This is derived from the Dempster-Shafer structure which allows the definition of a degree of belief and a degree of plausibility, respectively:

$$\text{Bel}(E) = \sum_{A_i \subseteq E} m(A_i) \quad (3.27)$$

$$\text{Pls}(E) = \sum_{A_i \cap E \neq \emptyset} m(A_i) \quad (3.28)$$

An example demonstrates the use of random set is given in Appendix F. The plausibility and the belief are actually upper and lower probabilities of a certain set of probability distributions. The constructed random sets model which is bounded by the upper and lower distribution functions is then called an imprecise probability model.

3.4.1.2 Imprecise Probability

Imprecise probability is characterized by a mixed case which specifies the bounds of probability for an uncertain quantity with underlying randomness that is not known in detail. Suppose \bar{F} and \underline{F} are non-decreasing functions mapping the real line \mathbb{R} onto $[0,1]$ and $\underline{F}(x) \leq \bar{F}(x)$ for all $x \in \mathbb{R}$. Let $[\underline{F}, \bar{F}]$ denote the set of all non-decreasing functions F from the reals into $[0, 1]$ such that

$\underline{F}(x) \leq F(x) \leq \overline{F}(x)$. When the functions \overline{F} and \underline{F} circumscribe an imprecisely known probability distribution, the model of $[\underline{F}, \overline{F}]$, specified by the pair of functions, is called a “probability box” or imprecise probability (Ferson 2002) for that distribution. This means that, if $[\underline{F}, \overline{F}]$ is a “probability box” for a random variable X whose distribution F is unknown except that it is within the “probability box”, then $\underline{F}(x)$ is a lower bound on $F(x)$ which is the (imprecisely known) probability that the random variable X is smaller than x . Likewise, $\overline{F}(x)$ is an upper bound on the same probability. From a lower probability measure \underline{P} for a random variable X , one can compute upper and lower bounds on distribution functions using (Walley 1991):

$$\overline{F}_X(x) = 1 - \underline{P}(X > x) \quad (3.29)$$

$$\underline{F}_X(x) = \underline{P}(X \leq x) \quad (3.30)$$

As shown in Fig. 3.10, the left bound \underline{F} is an upper bound on probabilities and a lower bound on quantiles (that is, the x -values). The right bound \overline{F} is a lower bound on probabilities and an upper bound on quantiles.

In the framework of imprecise probabilities, both the probabilistic uncertainty and non-probabilistic uncertainty can be considered simultaneously and transferred separately to the final results. This situation can occur when only certain statistical information is available but a pure probabilistic approach cannot be carried out due to a lack of pertinent information or irreducible uncertainty in the data (Fellin et al. 2005).

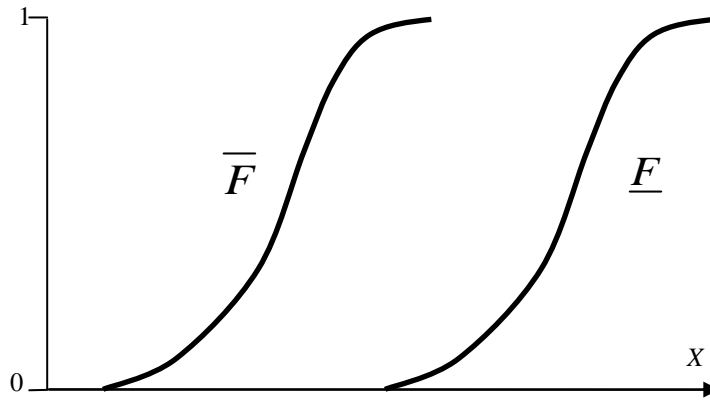


Figure 3.10 Imprecise probability.

The random set based imprecise probability is an attractive model for the bounding of probabilities and there are attempts to incorporate this in engineering analysis. Oberguggenberger and Fellin (2008) applied it to a sheet pile wall analysis. Rubio et al. (2004) applied the Dempster-Shafer concept to quantify the uncertainties in a slope stability problem. Both studies are able to model variables with inconsistent probabilistic behaviors and produce bracketing probability estimates arising from different sources of uncertainties as well as for combining information of different type without a need of treatment before the combinations.

3.4.2 Selection of Threshold and Time Span

A set of wave data provided from Wave Information Studies (WIS), a sponsored project supported by US Army Corps of Engineers (USACE), is used in this study to demonstrate the application of random set based model to quantify the uncertainties associated with the threshold and duration used in POT method. The

buoy is located in the south coast of Alaska (56.5°N 203.25°E) at 124m depth and a data set containing 25 years of hourly wave records (1985/1/1 01:00 to 2010/1/1 01:00) is filtered. The long term record of significant wave height (H_s), with an overall mean value of 1.897m, shows obvious variations in the mean with time. The highest monthly mean is 2.544m in January and lowest monthly mean is 1.238m in July. The pattern in extreme waves shows clearly that more severe conditions occur in winter, which is more likely to have storms, compared to summer. The annual maximum H_s is observed to vary significantly between years, with a largest value of 12.32m and lowest of 6.65m. Hence, such climatic variations may not support the assumption of independent and identically distributed data in statistical analysis. However, as long as the data are relatively stationary and their dependency structure is weak during its reference period, the POT is appropriate for modeling the extremes (Coles 2001). For illustration, the data which covers the January and February (two months) period are used as the sample data in the following POT approach.

As discussed in Section 3.2.3.1, the choice of the threshold (u) and time span (Δt) is a trade-off between the quantity of filtered data and the dependency between data points. Small values of u and Δt lead to higher number of exceedances but the later will have dependency characteristics, and may violate the basis of the Poisson and GPD models. Higher u and Δt may result in few extreme data points which may lead to inaccuracy (or higher uncertainty) in the estimated GPD parameters. Thus, the usual procedure is to derive a range of possible values for $(u, \Delta t)$ from the statistical fit of the exceedances, where large

deviations in the fit imply inappropriateness of the chosen $(u, \Delta t)$.

Based on a range of assumed threshold values, the asymptotic stability of the GPD model in POT method are investigated using the mean residual plot, and the L-moment plot, as introduced in Section 3.2.3. Figure 3.11 shows the mean exceedance plot for the 25 years of H_s data after applying POT with threshold ranging from 2.0 to 10.0. The time span used in the declustering step is $\Delta t = 24$ hours. For the GPD model to be valid, a linear trend is expected, see Eq. (3.10). Hence, the results show that the threshold should not be higher than 6.0 beyond which the sample size drops significantly to less than 100 (see Table C.4) causing large fluctuations in the predicted mean exceedances. This gives an initial sense of the range of appropriate threshold values. Note that a different choice of Δt may result in a different set of filtered data for the same threshold and thus lead to a different range of appropriate threshold values.

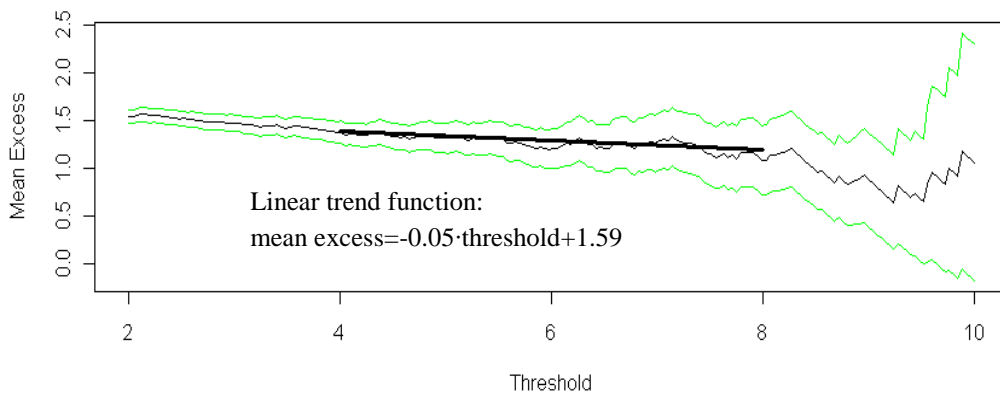


Figure 3.11 Mean residual plot with 95% confidence intervals (green line).

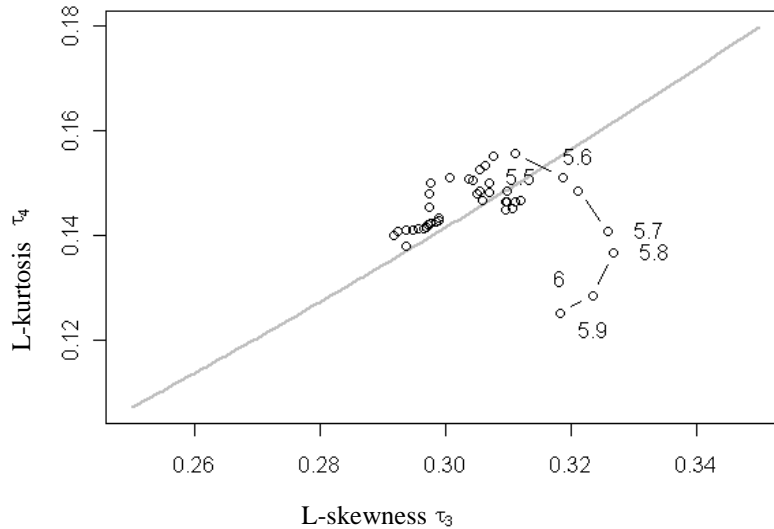


Figure 3.12 L-moment plot for exceedances over selected threshold (from 2.0 to 6.0 m) with theoretical GPD curve (grey line).

Figure 3.12 shows the L-moment plot for the threshold ranging from 2.0 to 6.0 with points corresponding to $u = 5.5$ to 6.0 connected by a curve. Compared to the theoretical GPD L-moment line in the plot, the range of $2.0 < u < 5.6$ appears yield compatible L-moments, which is not too different from the inference by the mean exceedance plot.

With regards to selecting an appropriate time interval Δt , various values are tested here. First the autocorrelation values in the H_S time series can be computed and the plot of Fig. 3.13 shows that records separated by larger than about 140 hours gives virtually zero correlation. The correlation is less than 0.4 for lag greater than 15. Thus, Δt ranging from 12 hours to 144 hours (nearest multiples of 12 hours) are further studied.

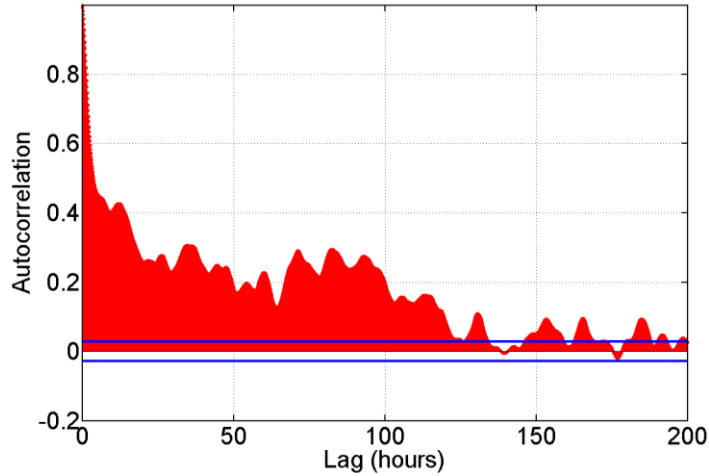


Figure 3.13 Autocorrelation plot for the wave time series records (blue lines represent upper and lower confidence bounds for assuming a moving average process).

Using these smaller ranges of u and Δt , the K-S test is applied to check the quality of the Pareto distribution fit and the Poisson property of the exceedances for the data series at a significance level α of 0.05. Figure 3.14 shows the “appropriate” region of $(u, \Delta t)$ marked by the combined left and right diagonal lines. The detailed results for various values of $(u, \Delta t)$ are recorded in Appendix C. Effectively, these two statistical tests have reduced the feasible domain from the earlier mean exceedance and L-moment plots.

One could see the feasible region of $(u, \Delta t)$ for passing the K-S test in Poisson process model is smaller compared to Pareto model. It implies that the adequacy of Poisson process model is more sensitive to the selected value of $(u, \Delta t)$ compared to the Pareto model. This can also be observed from Table C.1-C.2 where the tested p-value for Poisson process model is lower than the Pareto model for each use of $(u, \Delta t)$. Other than this, it is expected that if the significance level is change to 0.01, the feasible region will be increased to include more

combinations of u and Δt . The reason is because the criteria for rejecting the null hypothesis of following Poisson process or Pareto model becomes less stringent. Therefore, the exceedances filtered by the $(u, \Delta t)$ having p-values higher than 0.01 in both tests are considered appropriate to be applied in a Poisson-GPD model. The allowable range of $(u, \Delta t)$ for passing the level of significance of 0.01 is illustrated in Fig. C.1. From the calculated results, it shows that a total number of 48 $(u, \Delta t)$ are appropriate for the criteria of 0.01 in the significance level.

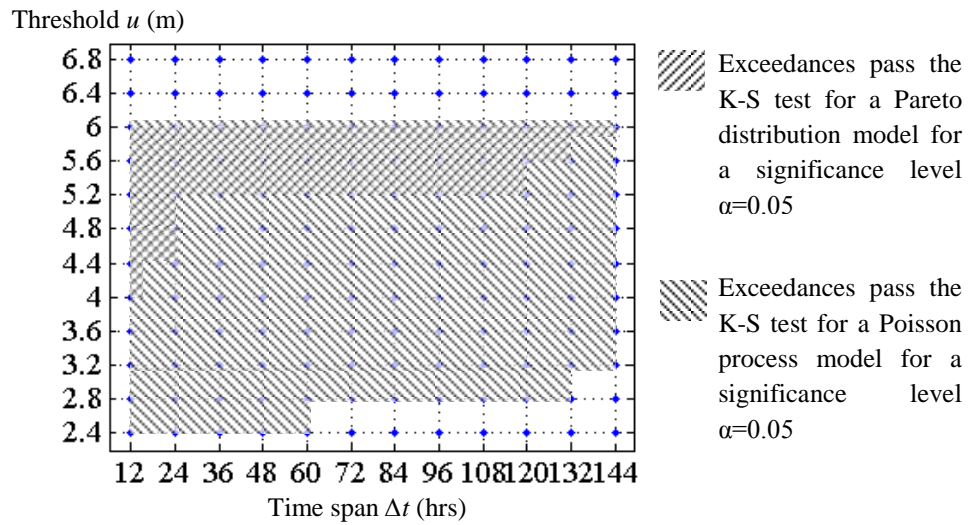


Figure 3.14 Appropriate region for u and Δt .

3.4.3 Uncertainty Quantification

As indicated above, the use of only one threshold and time span does not seem to be sufficient to ensure the statistical robustness in the prediction of long term value. In fact, the estimates for the 100-year return value deviate quite a lot from 16.38m ($u=6.0, \Delta t=36$ hours) to 14.04m ($u=4.0, \Delta t=12$ hours) in the upper bound of the 0.95 confidence interval. The details are provided in Appendix C. It is

very difficult to predict the errors contained in the results since the estimated confidence interval may not be unique. The reason is because the statistical difference of the exceedances for using different values of threshold and time span are quite small which may not be easily differentiated. This high level of uncertainties seems to be inherent and is hard to remove by conventional statistical method since the traditional POT approach only adopts one threshold and time span.

Here, the random set theory is employed to describe the uncertainties associated with the selection of u and Δt . In this study, each combination of u and Δt within the feasible range shown in Fig. 3.14 is considered as an observation for the true value of $(u, \Delta t)$. For example, by taking a spacing of 0.4m for u and 12 hours for Δt within the allowable region (the number of the intersected shaded cells between Table C.1 and Table C.2 in Appendix C), combinations of u and Δt are used in the POT to estimate the confidence intervals $[x_{100\text{-year},l}^i, x_{100\text{year},u}^i]_{\alpha=0.95}$ ($i=1, \dots, 38$) for the 100-year return value. Based on evidence theory, the interval $[x_{100\text{-year},l}^i, x_{100\text{year},u}^i]_{\alpha=0.95}$ is regarded as a focal subset A_i with an associated probability p_i . The available intervals are simply assumed to be independent and an unbiased averaging procedure is provided to assign each of these intervals an equal probability mass $p_i = 1/38$. Based on the pairs $(A_i = [x_{100\text{-year},l}^i, x_{100\text{year},u}^i]_{\alpha=0.95}, p_i)$, empirical distribution functions can be constructed for the bounds of the intervals. By using the Dempster-Shafer approach as described in Section 3.4.1, the estimations for the bounding functions $F_{100\text{-year}}(x)_u$ and $F_{100\text{-year}}(x)_l$ can be calculated as:

$$F_{100\text{-year}}(x)_l = \sum_{x \geq x_{100\text{-year},u}^i} p(A_i) \quad (3.31)$$

$$F_{100\text{-year}}(x)_u = \sum_{x \geq x_{100\text{-year},l}^i} p(A_i) \quad (3.32)$$

The constructed model is illustrated in Fig. 3.15.

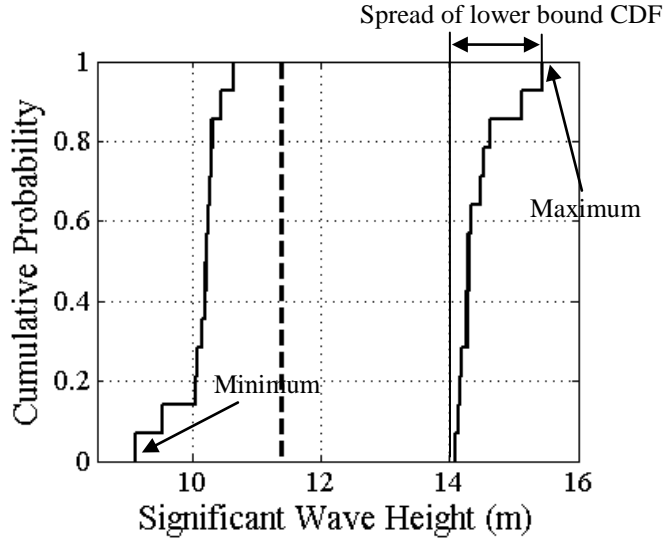


Figure 3.15 Constructed imprecise probability model for 100-year return value with mean value of all estimates (dotted line).

The result shows obvious spread in both the upper and lower bound cumulative distribution functions. This shows the significance of uncertainties associated with the selection of u and Δt to the estimated long term design value. In other words, the width of the bounded cumulative distribution function gives an indication of the sensitivity of the estimated 100-year return value with respect to the choice of u and Δt . The minimum of all the estimated lower bound and maximum of all the estimated upper bound give an approximate interval as an envelope for the 100-year return value. This interval is believed to be more robust than the traditional confidence interval since it accounts for the uncertainties

involved in the subjective selection of u and Δt .

To demonstrate how the imprecise probability model can be used to assess the uncertainties induced by different environmental conditions, the investigation is now extended to different data set. Here, three specific reference periods are selected. These are 1 month (January), 3 months (January-March) and 6 months (January-June), which are picked out from the 25 years of time series data. The same procedure is applied to the data and Fig. 3.16 shows the corresponding random set approximations for the 100-year return value in selected cases. It can be seen that the spread of lower bound cumulative distribution function is quite large for the case of 6 months (12.99~16.27m) compared to the others, such as the 1 month data set (14.09~15.45m). The reduction in the width of the imprecise probability could be viewed as a decrease in the uncertainties associated with the selection of parameters u and Δt . But this reduction does not imply a good model of the extremes, since the filtered data may lead to biased estimate. For example, the case for considering only the data in January may tend to give large weight to the statistical characterization in winter. Though the 1 month case is relatively stable and the wave record seems to be quite stationary, which results in low uncertainties, the results must be used with caution.

This model provides an indication of the intensities of the imprecision in u and Δt as indicated by the width of the imprecise probabilities (the spread of bound CDF as seen in Fig. 3.15). In this manner, global sensitivities of the results such as failure probabilities with respect to the imprecision in u and Δt can be revealed directly in one analysis. In such case, acceptable input imprecision

(acceptable intervals) can be determined based on specified constraints for the results such as an acceptable failure probability. This helps to decide whether further information should be gathered to reduce the uncertainty of the output.

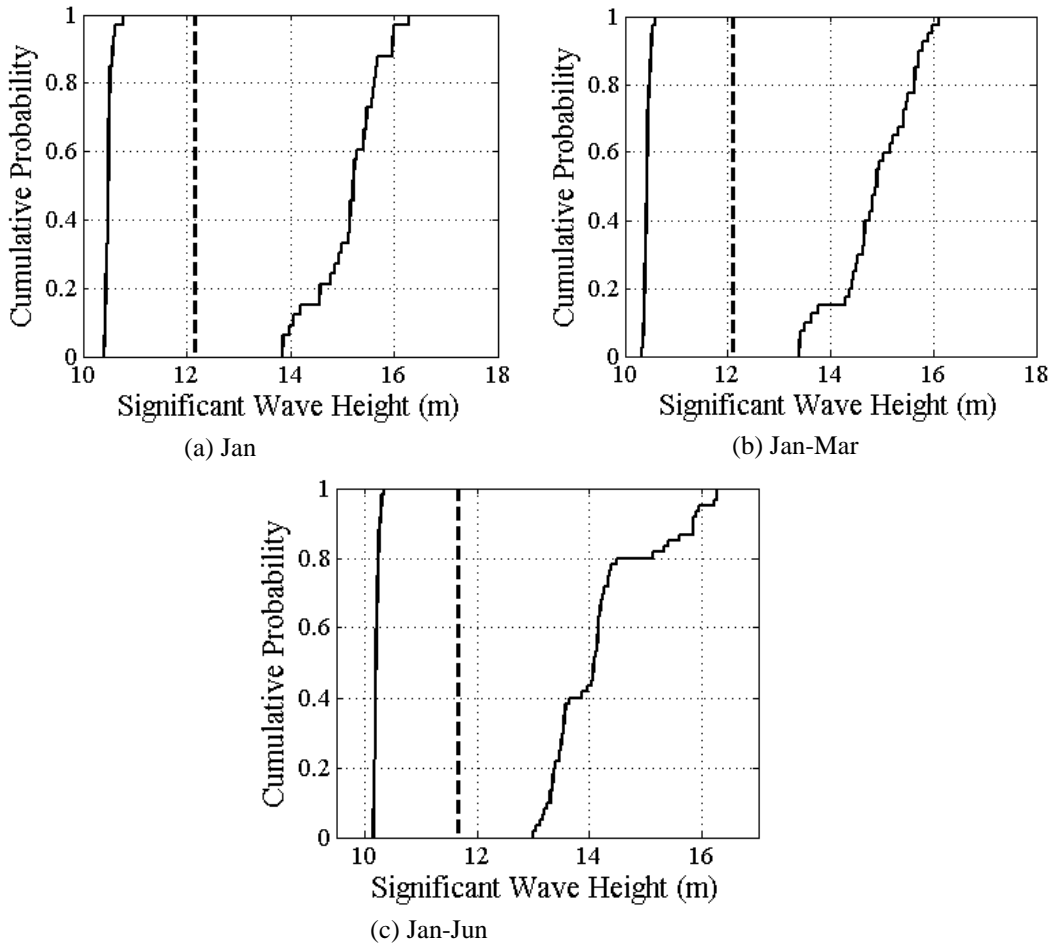


Figure 3.16 Constructed imprecise probability model for 100-year return value based on different sample set: (a) Jan (b) Jan-Mar (c) Jan-Jun. Dotted line represents mean of all estimates.

3.5 Concluding Remarks

In this chapter, several issues regarding the establishing of a robust extreme value model have been investigated, focusing primarily on the peak over threshold

method. Simulation studies are conducted to test the robustness of the established extreme value model from various methods by comparing the difference between the theoretical value and the estimated value. The methods are tested with simulated sample data that have different sample size, extreme tail behavior, random noise and serial correlation to see the overall performance in predicting the parameters in the extreme value model. Investigation of the POT method is also extended to an imprecise probability which is used to account for the subjectivity in the modeling.

The key conclusions are summarized as follows:

- Modeling the Pareto distribution using a limited number of data produces significant uncertainties in the shape, scale and high percentile estimates. Generally, the simulation studies show that MOM, PWMB and PWMU are the better parameter estimate methods. Besides the sample size effect, the tail behavior can influence the accuracy of the estimated parameters significantly, especially for light tail in the extreme data.
- The presence of random noise in the collected data increases the uncertainty in the parameter estimations. Noise in the location parameter has the most influence and the bias of the estimate arising from this may not be reduced much with more data.
- When limited time series data are available, the POT method may be the most appropriate compared to the annual maximum method and the r largest order statistics method. In addition, serial correlation in the times series data have little impact on the results from the POT method. From a practical point of

view, it implies that pretreatment of the data collected is minimal. However, the performance of the POT method is largely dependent on the appropriate choice of time span and threshold.

- A random set based imprecise probability model is proposed to describe the uncertainty associated with the selection of threshold and time span in the POT method. As the feasible values of threshold and time span may cover a finite domain, the imprecise probability model provides a consistent means to present the uncertainties in a quantitative form. The imprecise probabilities and bound for the results provide information as to whether further information to reduce the uncertainty is essential.

The conclusions drawn from this chapter will be used as reference for choosing the most appropriate parameter estimation method for the research presented in the subsequent chapters.

Chapter 4 Modeling the Time Varying Environmental Condition for Offshore Structural Analysis

The end of earlier chapter shows that the amount of non-stationarities contained in the environmental data is quite depending on the observed period. It generally implies that the environment is time varying and the associated established environmental statistical model should also be time dependent. This is particularly obvious for the ocean since the ocean climate may always exhibit phenomena of non-stationarity. Such change in the environmental characteristics may have significant effects on the load related factors and the irregular variations, if not accounted for in the statistical model, may lead to unreliable safety analysis of offshore structures. In this chapter, structural reliability analysis of offshore structures subjected to a time varying environment is investigated. An extreme value statistical model for the wave height is adopted as a basis for the performance assessment of a jacket structure. Two segmentation algorithms are proposed and applied to observation data to derive piecewise stationary processes for modeling the time varying effect. The results are compared with the traditional extreme values approach with regards to the accuracy and information content. The investigation also compares a case where the design of the structure ignores the time varying property.

4.1 Introduction

The consideration of time varying environmental hazards in the design of offshore structures is essential to ensure safe operation over the planned life-span. Coupled with the presence of inherent, statistical and/or model uncertainties, appropriate stochastic models of the environmental parameters are needed for a realistic reliability assessment to produce a reliable design. The focus should be on the identification and quantification of exceptional environmental conditions and associated uncertainties in view of the severe consequences of non-performance. The pre-requisite for this is adequacy and accuracy of recorded relevant environmental data, which can pose a challenge.

The assessment of design or system in the context of exceptional loads has been reported, often through a probabilistic framework. Examples include the stability analysis of slopes (Dijkstra & Dixon 2010), performance assessment of concrete structures (Biondini et al. 2006), as well as risk analysis considering climate change (Hughes et al. 2009). Generally, a changing environment can have two classes of effects on the safety of an existing structure: (i) deterioration of the structure, such as corrosion, erosion and aging, and (ii) variations in the environmental loads, including seasonal changes and inter-decade changes, see Fig. 4.1. Their combined effects make the performance of an existing structure significantly environment-dependent and hence must be factored in the design.

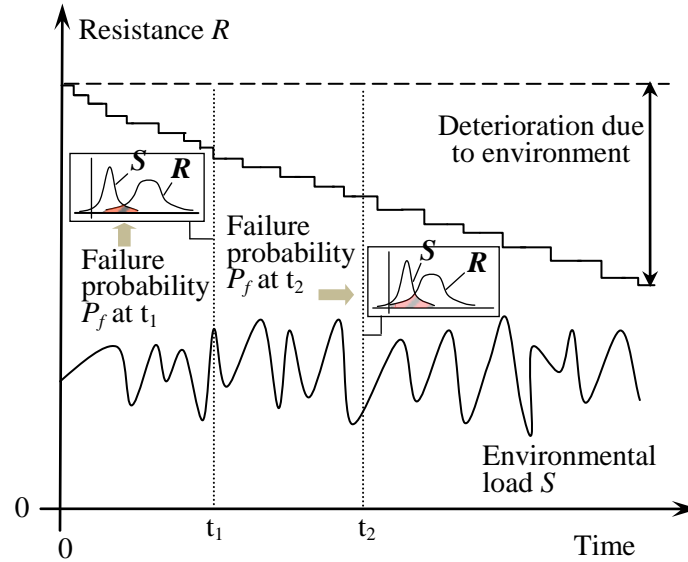


Figure 4.1 Time varying effects on safety of structures.

Generally, the design of structures requires long-term prediction of the environmental loads which is obtained through statistical extrapolation from field data. For example, the long-term variation of wave climate in offshore engineering is described in terms of the generic distributions for the governing sea-state parameters such as significant wave height H_S and zero-up crossing period T_Z . A long-term “safe” sea-state is usually estimated based on the observed data and used in the structural analysis to see whether the required safety level can be achieved. Such stochastic analysis poses several challenges. Firstly, the recorded environmental data are usually limited and may not be sufficient to predict the environmental conditions over the service life of the structure. A robust statistical method is needed to describe the resulting uncertainties and model the extremes where safety is an issue. Secondly, the collected field data are evidently time-varying. The collected data for each

parameter may not be stationary depending on the time scale of interest, and the data may be auto-correlated. In addition, the correlation between different parameters, for example, wind speed and wave height, needs to be considered in the stochastic model (Ditlevsen, 2002). Several researchers have reported the importance of considering covariate effects in establishing a statistical model for offshore parameters (Coles and Walshaw 1994; Robinson and Tawn 1997; Davison et al. 2012), and neglecting such dependencies result in under-estimating the structural reliability (Anderson et al. 2001; Jonathan & Ewans 2007).

This chapter introduces a procedure for the reliability assessment of offshore structure with the time varying effect. The objective of this current study is to derive an improved model to characterize the environment effects experienced by offshore structures. The scope is limited to the wave loading, where the primary concern is to construct a reliable statistical model of the wave height for use in the structural analysis. A non-stationary statistical model, which accounts for correlation between time and directionality, is introduced for the extreme significant wave height, H_S . The Poisson-GPD model is used to fit a set of wave height records taken from the WIS project. The variation in the value of statistical model parameters with time and directionality is approximated through a Fourier series expansion. In addition, a segmentation algorithm is proposed to discretize the data set. The model is used to characterize the loads for the reliability analysis of a jacket structure. The dependency between the environmental loads and structural response, and the advantage of using this model is highlighted.

4.2 Field Data at Ocean Site

For demonstration purpose, the same set of wave record data studied in Section 3.4 is analyzed. In addition to the magnitude of H_S , the observed direction and time for the waves are analyzed. For convenience, a parameter t representing the observed time for a particular H_S is defined in an interval $[0, 1 \text{ year})$, whereas the direction θ of H_S is defined in an interval $[0, 360 \text{ degrees})$, measured clockwise from North. The construction of a reliable statistical model for H_S is first considered, followed by its time varying effect on the overall safety assessment of a real offshore structure.

4.2.1 Seasonal Characteristics

When establishing the probability distribution of the H_S , one should understand the influence of the seasonal effects. As the data collected from the field measurement covers the varying seasons within a year, a summary of the statistical information for H_S over quarterly periods is tabulated in Table 4.1 and depicted in the box plot as shown in Fig. 4.2. The mean and quartile values are larger in the 1st and 4th (winter) season compared to the 2nd (summer) and 3rd season. The winter period produces more severe conditions as there are more storms occurring compared to summer.

It is observed that the winter period has a relatively larger dispersion in the H_S data (variance is 1.4886 for the 1st season and 1.5921 for the 4th season)

compared to the summer period (variance is 0.8008 for the 2nd season and 0.8000 for the 3rd season). To better understand this, the histogram which

Table 4.1 Statistics of H_S over four defined seasons.

	Min	1st Quartile	Median	Mean	3rd Quartile	Max
1st Season	0.19	1.56	2.19	2.437	3.09	10.49
2nd Season	0.12	0.85	1.26	1.478	1.89	7.68
3rd Season	0.09	0.71	1.1	1.319	1.68	12.32
4th Season	0.17	1.46	2.13	2.362	2.99	9.75

Notes: 1st Season: 1st Jan 00:00 ~ 31st Mar 23:00; 2nd Season: 1st Apr 00:00 ~ 30th Jun 23:00; 3rd Season: 1st Jul 00:00 ~ 30th Sep 23:00; 4th Season: 1st Oct 00:00 ~ 31st Dec 23:00.

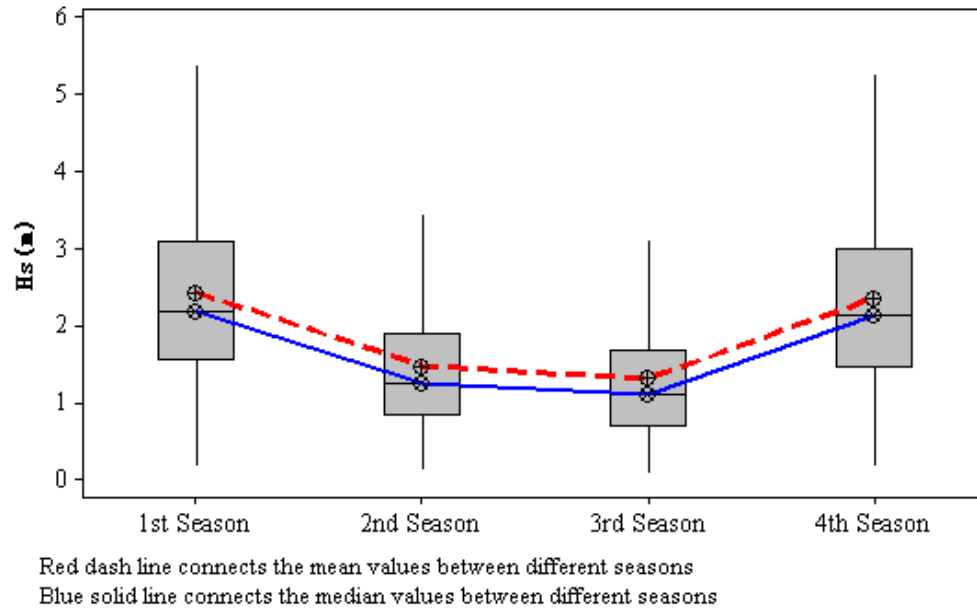


Figure 4.2 Box plot of H_S for four different defined seasons.

compares the frequencies of H_S values over these seasons is plotted in Fig. 4.3. It is easy to see that the histogram of H_S shows a larger spreading in the values during the winter period compared to summer. This difference implies a more

severe season may increase the dispersion of the environment parameter. In fact, if the statistical differences of the data are too large, the data should not be grouped together for a single statistical analysis. The establishing of the statistical model for the environmental parameter should account for all these influences.

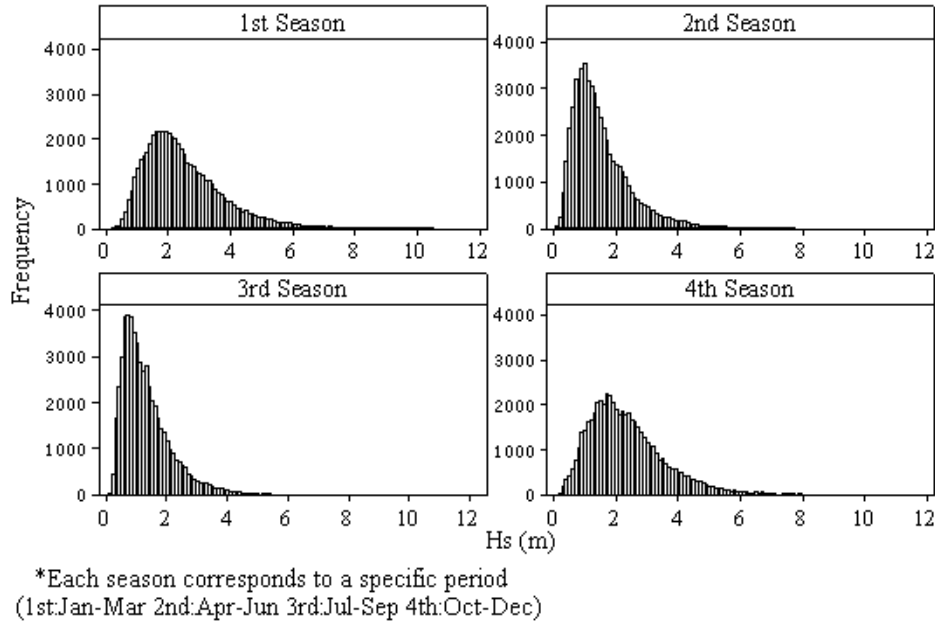


Figure 4.3 Histograms of H_S in four different defined seasons.

4.2.2 Directional Characteristics

The effect of seasonality in the wave data may not only exist in terms of the magnitude of wave height, but also in the direction from which the observed wave comes from. In most oceanic regions, particularly for the storm dominated regions, the observed wave direction is quite dependent on the storm direction. The requirement to consider the directional effect in developing the ocean parameter statistical model has been well demonstrated. It has been discussed for

quite some time (e.g. Graham 1981) about dealing with covariate effects, but only realized and adopted in the met ocean design criteria recently (Jonathan et al. 2011). It has also been proven that ignoring the directional effect is not accurate because the design value of the wave height from the most severe directional sector can be higher than the normally used omni-directional value (Forristall 2004; Jonathan and Ewans 2007).

As the chosen site is an open sea but not sufficient far away from the coast, the occurrences of the wave direction tend to be higher in some particular direction sectors compared to other. The general angular statistics of the occurrences of wave directions in each season is shown in Table 4.2. There is a large spread of the wave direction for all the four seasons as the resultant vector length is quite small. The skewness of the observed wave directions is quite small which indicates that the occurred wave directions are near symmetric.

Table 4.2 Statistics of wave directions over four defined seasons.

	Mean (°)	Resultant vector length	Angular standard deviation	Angular skewness	Angular kurtosis
1st Season	153.9	0.3644	1.1275	-0.1458	0.3798
2nd Season	165.3	0.3402	1.1488	0.0398	0.3365
3rd Season	191.3	0.3624	1.1293	0.0802	0.2901
4th Season	176.6	0.2786	1.2012	0.0894	0.3178

Notes: The values are calculated based on the directional statistics (Berens 2009): a direction α is first transformed to unit vectors: $r = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$. Basic statistics: mean resultant vector: $\bar{r} = \frac{1}{N} \sum_i r_i$, mean resultant vector length: $R = \|\bar{r}\|$ (a measure of concentration), standard deviation: $s = \sqrt{2(1-R)}$, skewness: $b = \frac{1}{N} \sum_{i=1}^N \sin 2(\alpha_i - \bar{\alpha})$ and kurtosis: $k = \frac{1}{N} \sum_{i=1}^N \cos 2(\alpha_i - \bar{\alpha})$.

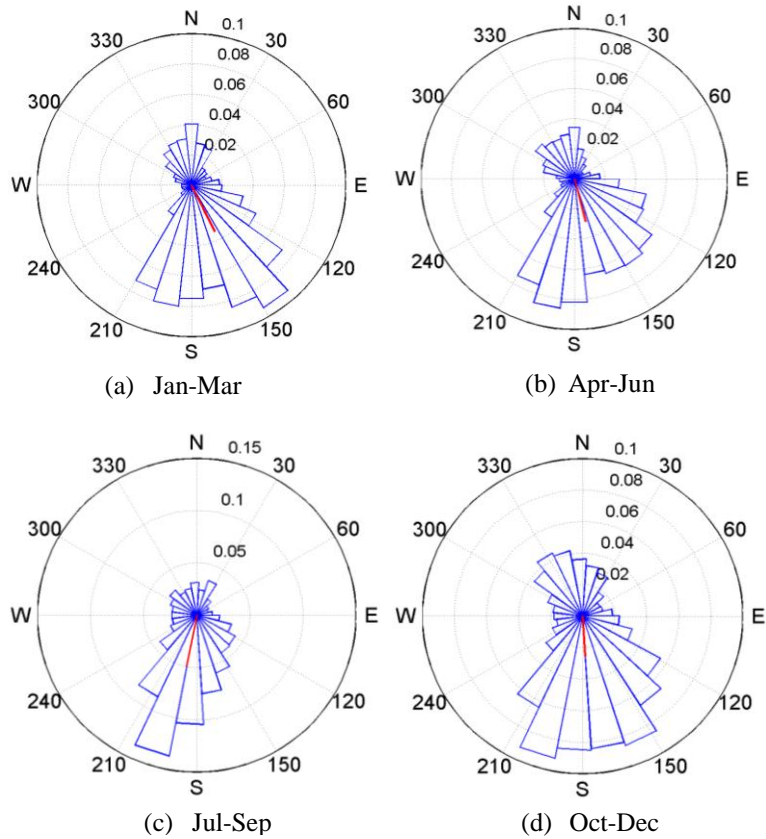


Figure 4.4 Angular histograms of H_S over four different defined seasons (red line represents the resultant vector length).

Figure 4.4 shows the angular histograms summarizing the occurrences of wave directions for the different seasons. Clearly, it shows the wave directions are more likely to occur from the south-east direction. It is convenient to classify the data according to “geological directions” — for example, direction from “sea to land” or “land to sea”. At this site, the wave direction between 90° and 240° belongs to the direction from “sea to land”, while wave direction between 270° to 30° constitute the direction from “land to sea”. It is observed that the directions from the sea, especially for a value ranging from 120° to 210° , occur more frequently than the direction from “land to sea”. For the directions between these

two dominant directions, for example the northeast and southwest directions, the occurrences are relatively much lower.

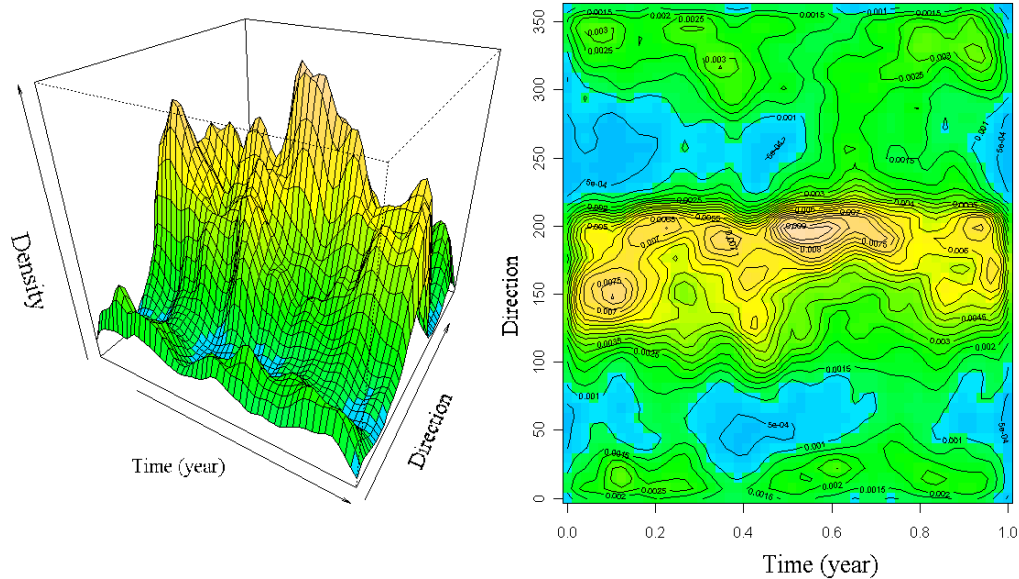


Figure 4.5 2D Kernel density plot of H_S with directionality and seasonality.

Detailed information on the distribution of the wave data with both the directional effect and the seasonal effect is illustrated by the kernel density plot in Fig. 4.5. It is observed that, for the wave directions around $150^{\circ} \sim 200^{\circ}$, the probability density function value is relatively much higher (the yellow region) consistently through the whole year. The seasonality, on the other hand, makes only small fluctuations in the density function values for H_S which implies weak seasonal effect on directionality. This implies that the frequency of occurrences of waves at the chosen site is sensitive to the directionality but not the seasonality. Distribution of H_S is more critical for the “sea to land” direction sector. However, though the waves are infrequent from some directions, extreme H_S can still occur from those directions. Hence, these possible extremes should not be ignored.

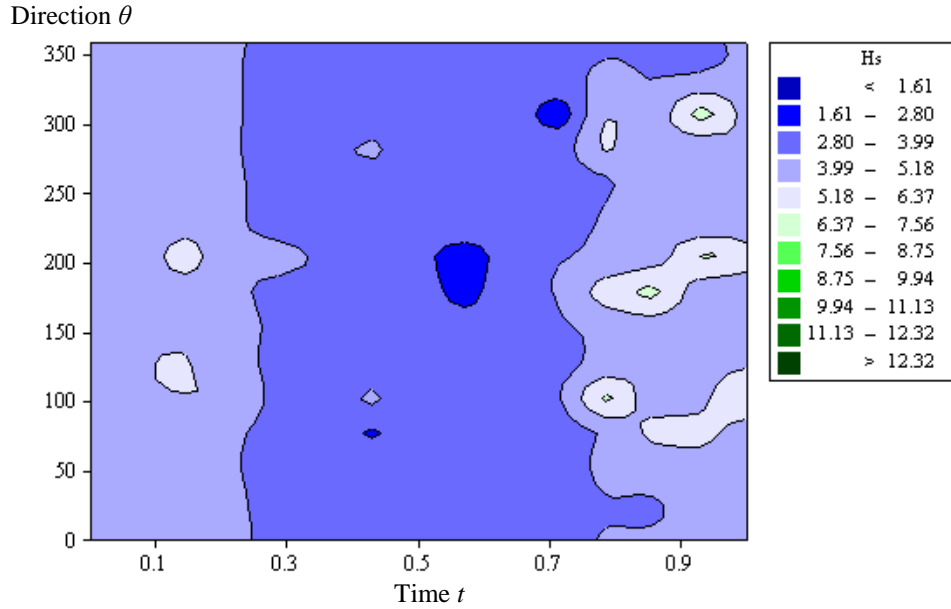


Figure 4.6 Isoline plot of H_s with directionality and seasonality.

To understand the effect of seasonality and directionality on the magnitude of H_s , Fig. 4.6 shows the isoline plot of H_s with these two dimensions. Obviously, the seasonality gives the most significant influence to the values of H_s as the figure shows significant changes along the seasonal (t) dimension. However, the change along the directional axis is less obvious. Generally, the wave record shows that the magnitude tends to be smaller in the middle period of the year, see Fig. 4.6. Whilst a large value of H_s is observed in December may imply a higher likelihood of various types of hurricanes in that month, a low value in July may indicate a relatively calm season. These observations serve to illustrate that one should recognize the importance of seasonal effect in modeling the magnitude of H_s , while the directional effect should be recognized when the occurrences (or frequencies) is to be modeled. Consequently, the variations of magnitude of H_s

with respect to both time and direction should be considered in establishing of a probability model.

4.3 Test for Stationarity of Poisson-GPD model

As the occurrence of extreme H_S affects the performance of offshore structures, an appropriate statistical model is essential. As discussed in Chapter 1, the Pareto distribution model is used to model the exceedances above a specific threshold and the Poisson process is used to model the occurrence times of these exceedances. Thus, the Poisson-GPD models both the intensity and frequency of extreme events. As discussed in the preceding section, from the H_S data, its occurrences is directional dependent whereas the intensity is seasonal dependent. Thus, the parameters in the Pareto distribution (that is, $\{u, \xi, \sigma\}$) are affected by seasonality, whereas the Poisson process parameter (that is, $\{\lambda\}$) is dependent on directionality.

Unfortunately, using Eq. (3.8) in the Poisson-GPD model requires the data to be stationary. The seasonal trend and directional pattern exhibited by the time series records of H_S imply significant non-stationarity. One way is to pre-treat the collected data so that the stationary stochastic model can be adapted for the current application, such as by segmenting the data into portions which are stationary. Several segmentation algorithms are available and examples include the testing of unit root assuming that the time series follow a regression model (Enders 2004), a Bayesian approach assuming that the initial distribution model of

the time series is known (Yang & Kuo 2001), as well as detecting the change point in the time series based on simple statistical properties (Hawkins 2001). There is no unique method to segment the time series, and the selection of an appropriate method is highly dependent on data characteristics and the required homogeneity of data in each segment.

Strictly speaking, an absolute stationary H_S time series data can hardly be obtained. Within a reasonably long period, H_S series always exhibit alternating occurrences of monotonically increasing and decreasing trends due to the influence of storms. A long term statistical model is not meant to characterize a “developing” or “decaying” sea-state. If a segmentation strategy is to be adopted, the appropriate length of data in each segment is a primary consideration. The issue of separating into different “months”, “seasons” or “years” warrants careful consideration. From a statistical point of view, the derived piecewise random sequence must be *sufficient long and stationary* for probabilistic analysis. The next consideration is the *starting position of a stationary segment*. Normally, there is a transition period between two specific environments so that a specific point in time for the change is difficult to ascertain. Both issues to be considered in the segmentation of both direction θ and time t are discussed next.

4.3.1 Segmentation Algorithm for Seasonality

With regards to segmentation in time, the H_S time series is recursively segmented in time until the optimum segmentation is obtained. Here, the mean value is used as a criterion for defining the homogeneity of H_S within a time period. A

statistical measure based on *t-test*, which has been used in other segmentation procedures (Toth et al. 2010), is adopted. The main task is to partition the time series into segments such that each segment has significant differences in the mean from the adjacent segments. However, the original algorithm is only applied to detect the instantaneous changes for a time series without consideration of the length of each discretized segments. For this example, it is necessary to use constant length of segments to discretize the time series in order to guarantee each established model covers the same length of period. The basic reason is to ensure each of the models is established based on the same amount of data. Based on this concern, some essential complementary procedures to identify the starting point of a segment have been added to the original method. The general algorithm is as follows:

- A specified time interval ΔT is selected, for example “seasonal length”, and the corresponding cuts based on this time interval are made over the whole time series wave data based on their t_i values. This lead to a total number of $1/\Delta T$ segments in the time dimension. For example, if ΔT is set to 0.25, the whole year (unit one) will be divided into four different time sectors (e.g. like four seasons). Or if ΔT is set to 1/12, the whole year will be divided into 12 time sectors (e.g. like 12 months).
- Then the difference in the means across the subsets to the left and to the right at each cut are evaluated using the *t-statistic*:

$$t = \left| (\mu_{left} - \mu_{right}) / s_D \right| \quad (4.1)$$

where

$$s_D = \left[\left(s_{left}^2 + s_{right}^2 \right) / \left(N_{left} + N_{right} - 2 \right) \right]^{1/2} \times \left(1 / N_{left} + 1 / N_{right} \right)^{1/2} \quad (4.2)$$

μ_{left} , μ_{right} are the mean values, s_{left} , s_{right} are the standard deviation and N_{left} , N_{right} are the sample size in the left and right segment, respectively. The significance level for the cut at that position is calculated by (Toth et al. 2010):

$$P(t) \approx \left\{ 1 - I_{\left[\frac{v}{v+t^2} \right]} \times (\delta v, \delta) \right\}^\gamma \quad (4.3)$$

where $\gamma = 4.19 \ln N - 11.54$, $\delta = 0.40$, $v = N - 2$,

$I_x(a, b)$ is the incomplete beta function and $N = N_{left} + N_{right}$

The significance level $P(t)$ of a possible cutting point is defined as the probability of obtaining the value t -statistic or lower values within a random process. The adequacy of this measure has been proven and demonstrated in a simulation study by Bernaola-Galván et al. (2001). The t -statistic is tested for each cut along the time series. The minimum statistical significance level calculated from these cuts $P_{overall}$ must exceed the required significance level P_o , usually taken as 0.95. If this is satisfied, the series is segmented at significance level P_o .

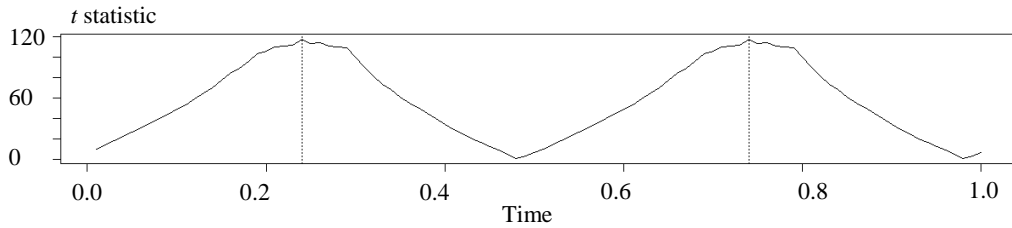
- The segmentation procedure in first and second step is repeated with various choices of the starting points in order to identify the maximum p-value P_{best} over all segmentations. The associated segmentation is considered as the best segmentation for the series representing a segmentation at the significance level P_{best} .

$$P_{overall} = \min P(t_i) \geq P_o \text{ for } i = 1, \dots, n \quad (4.4)$$

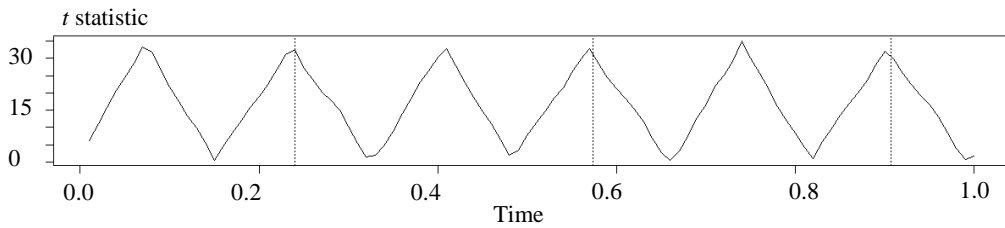
$$P_{best} = \max_{j=1}^m \min_{i=1}^n P_j(t_i) \quad (4.5)$$

where n is the number of cuts, m is the total number of changes of the starting points in the search for the best segmentation.

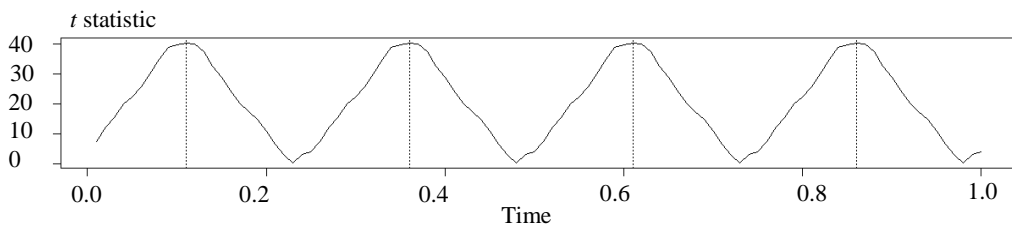
The algorithm is applied to test the appropriateness of the number of segments at the same time. These are 2, 3, 4 and 5 segments in the example. Since this study concerns the yearly cyclic changes instead of short period variations, the weekly mean values are used in this test for identifying the seasonal changes within the time series. In order to search the optimum solution (corresponding to P_{best}) in each of these segmentations, different cut starting positions are selected and compared. Here the first cut position is initially chosen at the starting point of the time series at $t=0$ and then slowly slide to the end of the year to test the performance for different cut scenarios among the year. A number of 100 different starting positions are utilized in this searching (e.g. starting position are chosen at $t = 0.01i$ for $i = 1, \dots, 100$). This is used to keep the search for the optimum point at a precision of 0.01 (year) for the time series. The results of this performed algorithm in each of these segmentations are illustrated in Fig. 4.7. Instead of the significance level value $P(t)$, the values of t -statistic is used in the plotting. The reason is because most of the calculated $P(t)$ are very close to 1 and are not easy to compare in the same figure.



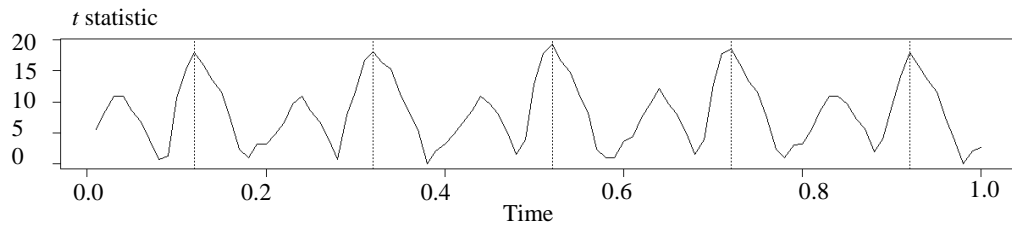
(a) Discretization of two segments in the seasonality



(b) Discretization of three segments in the seasonality



(c) Discretization of four segments in seasonality



(d) Discretization of five segments in seasonality

Figure 4.7 Results of t -statistic for discretized time series (dotted line represents optimal segmentation points).

Besides the mean in the defined segmented time series, another possible statistical property that should also be considered is the seasonal trend. The seasonal trend is a form of predictable pattern of variations in the time series data. Typically, the existence of the seasonal trend would make obvious changes in the statistical properties of the time series for different period which will make the

data non-random. However, the randomness is one of the key assumptions in establishing the statistical model based on the time series data. This requirement must be achieved to guarantee a constant location scale and shape parameter in the constructed GPD model.

To deal with this issue, the runs test is applied here to detect the seasonal trends in the following analysis. The runs test (Bradley 1968) is a non-parametric procedure for the evaluation of statistical independence and underlying trends within a sequence of observed values. It can be used to decide if a data set is from a random process. A run is defined as a series of consecutive increasing or decreasing values. The number of increasing or decreasing values is the length of the run. For a random data set, the probability that the $(i+1)$ th value is larger or smaller than the i th value follows a binomial distribution, which forms the basis of the runs test (Bendat and Piersol 1986).

For the discretized segments, if the seasonal trend is not very obvious within the partitioned time series, the observed values are expected to have equal probabilities of larger or smaller than the sample median. So by coding values above the median as positive and values below the median as negative, if the number of runs is significantly higher or lower than expected, the hypothesis of statistical independence of the elements may be rejected (Mendenhall and Reinmuth 1982):

H_0 : the sequence was produced in a random manner.

H_1 : the sequence was not produced in a random manner.

The runs test is thus therefore tested by the statistic:

$$Z = \frac{R - \bar{R}}{s_R} \quad (4.6)$$

where R is the observed number of runs, \bar{R} is the expected number of runs, and s_R is the standard deviation of the number of runs. The values of \bar{R} and s_R are computed as:

$$\bar{R} = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad (4.7)$$

$$s_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} \quad (4.8)$$

where n_1 and n_2 are the number of positive and negative values in the series. The runs test will reject the null hypothesis if $|Z| > Z_{1-\alpha/2}$ where the Z statistic follows a standard normal distribution. To ensure randomness within the segments, the runs test is also calculated for each of the segments for different scenarios from the segmentation algorithm. The tested results including both the t -*statistic* and runs test are summarized in Table 4.3.

It can be seen from the table that the calculated t -*statistic* for four segmentations are all very large. Among these four different scenarios, the results show that 2 segments and 4 segments are better choices in the segmentations as the t -*statistic* are higher compared to the other two. However, the runs test shows that for the case of 2 segments, the p-value is quite small, even lower than a commonly required level (e.g. 0.05). This indicates that the seasonal trend effect

is very obvious within the segments such that the data within the segment is no longer stationary. Or in the other words, the length of the segment might be too large since a further discretization could help to reduce the trend effect, see Table 4.3. Compared to the 2-segments case, the calculated p-value in runs test for 4 segments is quite large. This ensures the randomness within the segments if the time series is partitioned into four parts with the identified optimum cuts. Therefore, 4 segments is the optimum choice for the segmentation as the tests showed that each of the four segments has mean values that is significantly different from the adjacent ones and the seasonal trend effect is minimal. Thus, this is adopted in this example.

Table 4.3 Summary of tested algorithms in different segmentations of H_S with respect to time.

No. of segments	Discretized segment length ΔT	Maximum <i>t</i> -statistic	Minimum p-value in runs test for all the segments with the optimum cut	Cut starting position
2	0.50	117.6776	0.0005	0.24
3	0.33	35.0320	0.0059	0.23
4	0.25	40.29127	0.0693	0.10
5	0.20	19.35148	0.1797	0.11

After the time sectors with reasonable stationary data sets are identified, the GPD model is next applied to the H_S data in each of the sectors. Here, a varying threshold defined as 1.5 times the mean in each sector is used in the POT approach to accommodate seasonal heterogeneity. To reduce the serial correlations, de-clustering is performed to filter out the independent data. A minimum time interval of 12 hours (time span value) between extremes is considered. Small differences between the empirical and theoretical values in the

estimated L-moment and mean exceedances support the choice of the threshold and time interval. In order to justify the chosen threshold and time span are good enough, the suitability of several other values of threshold and time span are also compared in the model test. These include the consideration of threshold equals to 0.5, 1.0 and 2.0 times the mean and time span equals to 24 and 36 hours. More detailed information of testing the suitability of these two values can be seen from Appendix D. A typical empirical kernel density estimate for the filtered extremes is shown in Fig. 4.8.

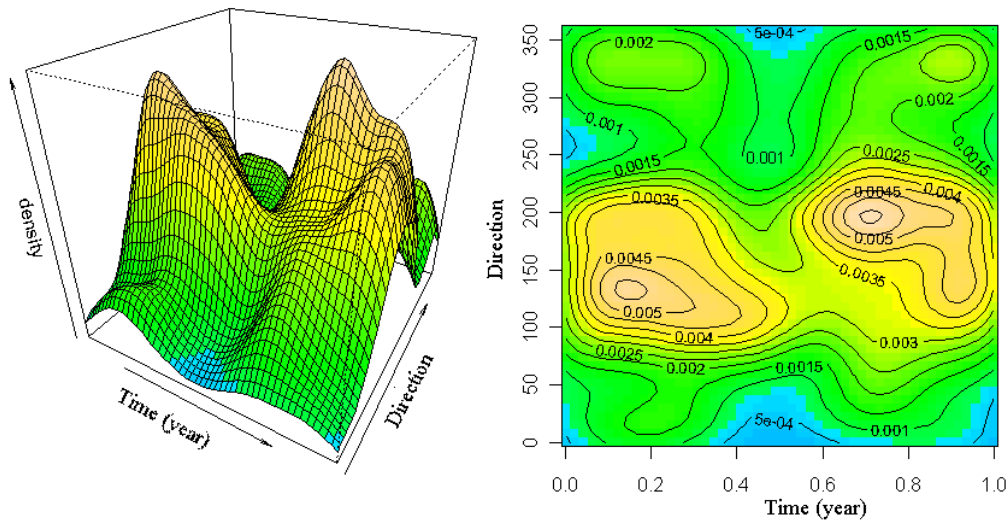


Figure 4.8 2D kernel density plot of extreme H_5 with directionality and time.

The figure clearly shows that the frequency of the identified extreme H_5 varies with the direction as expected. This generally means that the occurrence rate of these exceedances (e.g. the Poisson rate λ in the Poisson-GPD model) will change due to different directions. Therefore, further partitioning is necessary with respect to direction.

4.3.2 Segmentation Algorithm for Directionality

The proposed procedure for directional sector segmentation falls under the topic of change point detection in non-stationary Poisson processes. Analogous to the occurrence times t_1, \dots, t_n of extreme H_S , the directional parameter θ is subdivided into suitable intervals $\Delta\theta$ as directional sectors, so that the extreme H_S occurred at directions $\theta_1, \dots, \theta_n$ within each sector can be considered to have a constant Poisson rate λ_i . The quality of the segmentation is mainly assessed by testing whether each segmented sector has a specific constant Poisson rate for the directional extremes. That is, the presence of change points is tested. A simple cumulative sum statistical method (Galeano 2007) is adopted here for this purpose. For the observed directions θ_i for each of the identified extreme H_S , the statistic C_i is used to detect the changes of Poisson rate for the occurrences within the discretized directional sector (e.g. $\theta_i \in [\underline{\theta}, \bar{\theta}]$), which is defined as

$$C_i = |D_i|, \quad i = 1, \dots, n \quad (4.9)$$

where

$$D_i = \sqrt{n} \left(\frac{\theta_i}{\theta_n} - \frac{i}{n} \right) = \sqrt{n} \left(\frac{\sum_{j=1}^i d_j}{\sum_{j=1}^n d_j} - \frac{i}{n} \right) \quad (4.10)$$

n is the number of observed direction values within the sector, θ_i is the i th largest observed direction value within the directional sector such that $\underline{\theta} \leq \theta_1 < \theta_2 < \dots < \theta_n \leq \bar{\theta}$, and d_i is the differences in directions between two consecutive θ_i , defined as $d_1 = \theta_1$, $d_i = \theta_i - \theta_{i-1}$, for $i = 2, \dots, n$. For a Poisson process, the d -statistic behaves asymptotically like a standard Brownian bridge on $[0, 1]$, see Inclán and Tiao

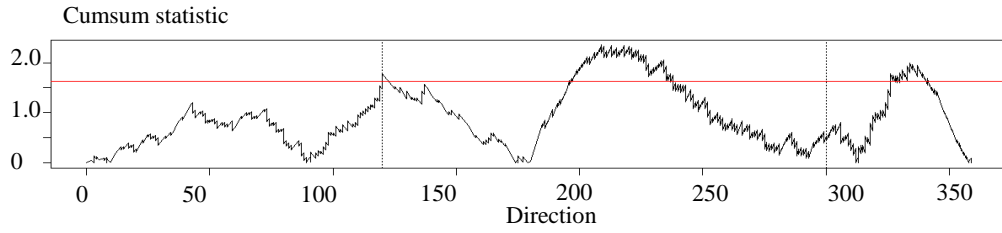
(1994). The limiting distribution will give a critical estimate for the asymptotic C_i value in a homogeneous Poisson process (Galeano et al. 2007). A significant exceedance of the asymptotic value implies the need for finer segmentation.

The steps in performing this test are similar to the time segmentation procedure. The directional segmentation is performed for the data in each time sector separately. Again, the procedure targets at both an optimal starting point for the segments and an optimal segment size. The largest C_i value is retained in each segmentation test as a reference to compare with the theoretical limiting value. The proposed algorithm, for the present case of sectors of equal sizes, is as follows:

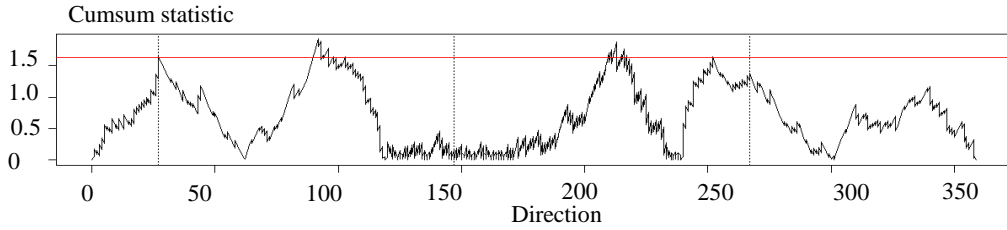
- Define a directional sector size $\Delta\theta$ and define the directional sectors over the range of direction values from 0° to 360° .
- Group the observed directions into different directional sectors according to its value. For each of the directional sector, sort the values of the directions θ_i in an ascending order (e.g. $\theta_1 < \theta_2 < \dots < \theta_n$) and calculate the statistic C_i for each observed direction according to Eqs. (4.6)-(4.7). by comparing the maximum calculated C_i value against the critical value corresponding to a predefined significant level α , if all values are lower than the critical, it may be inferred that the Poisson rate is constant within each segmented directional sectors.
- If all the directional sectors cannot achieve the requirement (e.g. maximum C_i is larger than the limiting value), the procedures in the previous two steps will be repeated with various choices of starting points in the direction

values to determine the optimum segmentations. If the optimum case still cannot meet the requirement, the associated $\Delta\theta$ is deemed too coarse, where the rate is not piecewise constant and a smaller $\Delta\theta$ should be considered.

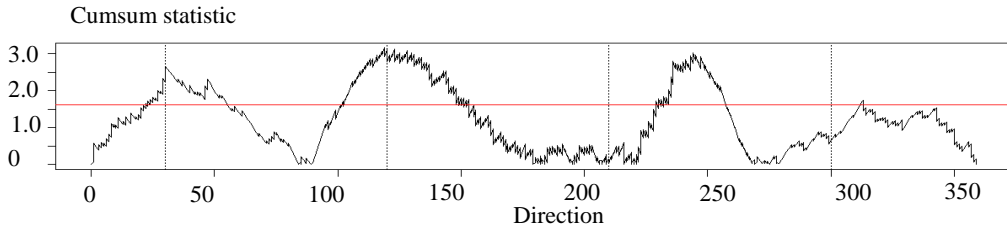
Based on these, the directional segmentation procedure is applied to each of the time sectors. The directional segmentation variants with 2, 3, 4 and 5 segments are tested and the results are summarized in Table 4.4. A critical value $C_{i,\alpha=0.01}=1.624$ is used here at a significance level of $\alpha=0.01$ for the detection of a change in λ within a segment. The segmentation procedures are applied to the data with shift in the starting point to search for the optimum solution in each of the cases. As shown in the table, the maximum C_i values for the segmentation variants with 2, 3 and 4 segments have exceeded the critical value which indicates the potential existence of a change point within segments, whereas the variant with 5 segments shows the best performance and satisfies the requirements. This can be illustrated by comparing the plot of C_i values for the tested segments in Fig. 4.9. It is seen that the computed C_i for all the segments are larger than the limit value even for the optimum discretization scenario. However, for division into 5 segments, if the starting point of the segmentation shifts to 62° (the optimum discretization), all the C_i drop to values lower than the limit value and thus satisfies the requirement. This also proves the importance of starting points in the segmentations which has been emphasized in third step of the algorithm. Therefore, 5 segments are chosen for modeling the stationary Poisson processes.



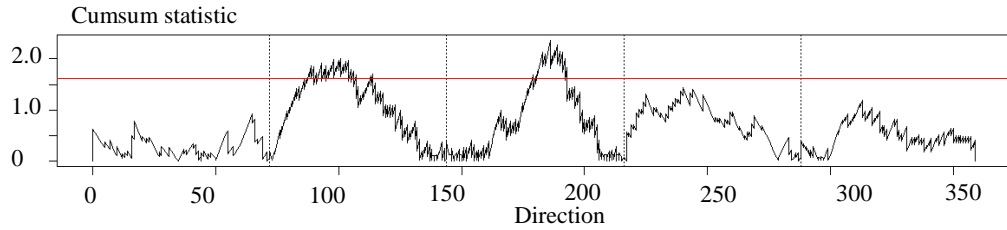
(a) Optimum discretization of two segments in the directionality



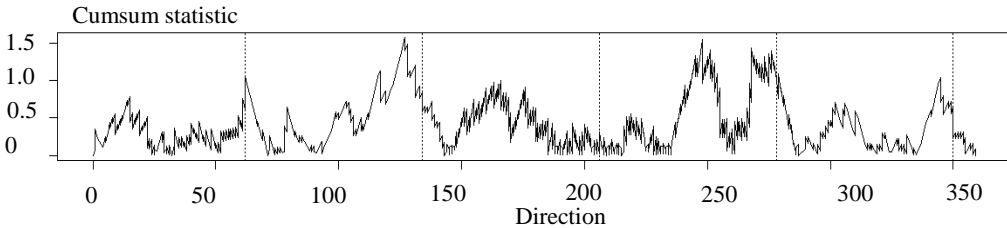
(b) Optimum discretization of three segments in the directionality



(c) Optimum discretization of four segments in the directionality



(d) Discretization of five segments in the directionality without shift



(e) Optimum discretization of five segments in the directionality

Figure 4.9 Comparison of C_i values between different segmentations for time sector $[0.10, 0.35)$. (dotted lines are segmentation points and the red solid line represents limiting value for significant level $\alpha=0.01$).

Table 4.4 Summary of tested statistics for different segmentations of extreme H_S with respect to directionality.

Number of segments	Maximum C_i	$\Delta\theta$	Starting point
2	2.0949	180 °	120 °
3	1.9227	120 °	27 °
4	3.0697	90 °	30 °
5	1.5737	72 °	62 °

Hence, the resulting discretized model involves 20 sectors, namely, 4 time sectors \times 5 directional sectors. The extreme H_S within each sector form a specific GPD model. To determine the parametric values of the GPD model in these sectors, the probability weighted moments (PWM) estimation method is adopted herein since it can estimate the values of the statistical parameters well even for small sample size, as concluded in Chapter 3. The results of the determined model parameters are shown in Tables 4.5 and 4.6.

Table 4.5 Estimated shape parameter ξ in each discretized sector.

		Time t			
		(0.10-0.35]	(0.35-0.60]	(0.60-0.80]	(0.85-1.0] & (0-0.10]
Directionality θ	(-10°~62°]	-0.9514	-0.4736	-0.6944	-1.0260
	(62°~134°]	-0.8823	-0.4055	-0.4449	-0.8829
	(134°~206°]	-0.9538	-0.8313	-0.5716	-1.1393
	(206°~278°]	-0.7682	-0.6712	-0.7738	-0.3506
	(278°~350°]	-1.0520	-0.5246	-0.6808	-0.4538

Table 4.6 Estimated scale parameter σ in each discretized sector.

		Time t			
		(0.10-0.35]	(0.35-0.60]	(0.60-0.80]	(0.85-1.0] & (0-1.0]
Directionality θ	(-10°~62°]	2.9467	1.8014	1.9605	2.7212
	(62°~134°]	3.7969	2.0682	2.0823	4.0230
	(134°~206°]	3.3720	2.3465	2.6957	3.9907
	(206°~278°]	3.0608	2.1123	2.3260	2.1463
	(278°~350°]	3.4300	1.4980	2.0324	2.3934

It could be observed that the differences in parameter values between the sectors are quite obvious which implies large statistical differences between the discretized sectors. This is consistent with the former inference that the seasonal and directional effects can significantly affect the wave height values.

4.4 Time Varying Modeling

The preceding section shows a discrete representation of the time varying effects in the collected wave height data, where the main focus is given to the dependencies with the seasonality and directionality. In this section, it is intended to further improve this discrete model by not having abrupt value changes in the parameters of the GPD model between adjacent sectors.

4.4.1 2D Fourier Series Characterization

For a time varying natural process, it is likely that the statistics of a continuous natural phenomenon will not have abrupt variation. Given the discretized model as discussed in the former section, there will be abrupt variations, hence lacking in continuity along the directionality and seasonality dimensions. Instead of a sudden change between adjacent sectors, it may be more realistic representation of natural process, given the presence of uncertainties, if the changes in the model parameter between adjacent sectors transit smoothly with respect to direction and time.

There were numerous work done on the characterization for extreme H_S . These studies provided a wide range of techniques to characterize the variability of H_S with respect to time and directionality (Méndez et al. 2006; Menendez et al. 2009; Mackay et al. 2010). The straight forward approach is to model each of the GPD model parameters as a smooth time varying function, using maximum likelihood estimation in conjunction with a suitable smooth basis function. For example, a Fourier series (Jonathan and Ewans 2011) and spline (Green and Silverman 1994) bases have been proven to be appropriate. Other alternatives, like Legendre polynomials (Northrop and Jonathan 2011), random fields (Rue and Held, 2005) or spatial splines (Ramsay 2002; Ruppert et al. 2003) are also approaches to model the statistical dependency amongst the model parameters and other factors. Some approaches have been extended to include multiple or multivariate covariates in the modeling (e.g. Marx and Eilers 1998).

Similar to the proposed idea in Jonathan and Ewans (2011), a 2-D Fourier series expansion is introduced herein to model the time and directional dependencies for the shape and scale parameters. Firstly, the estimated parameters within each sector are actually values at discrete points equally-spaced over a 2-D space. For example, the estimated parameters in the sectors (Tables 4.5 and 4.6) are treated as a two-dimensional signals in a 4×5 grid as $\{\zeta(n, m), \sigma(n, m)\}$ $n=0,1,2,3$; $m=0,1,2,3,4$. Then, 2-D discrete Fourier transformation (DFT) is applied to these points to obtain the discrete frequencies along the two directions (m represents the directionality and n represents the time). Here, the number of pairs of frequencies used in the 2-D Fourier expansion is 20, which is the same as

the number of sectors. The value of the discrete spectrum for frequency at $(\omega_k, \omega_l) = 2\pi(k/4, l/5)$ $k=1, \dots, 4; l=1, \dots, 5$ can then be calculated by (Rahman 2011):

$$F_{\xi}(\omega_k, \omega_l) = \frac{1}{20} \sum_{n=0}^3 \sum_{m=0}^4 e^{-i(\omega_k n + \omega_l m)} \xi(n, m) \quad (4.11)$$

$$F_{\sigma}(\omega_k, \omega_l) = \frac{1}{20} \sum_{n=0}^3 \sum_{m=0}^4 e^{-i(\omega_k n + \omega_l m)} \sigma(n, m) \quad (4.12)$$

Next, the discrete spectrum is converted into continuous “signals” in the original space by using the inverse discrete Fourier transformation (IDFT). The shape or scale parameter for a specific time t and direction θ can then be computed as:

$$\xi(t, \theta) = \sum_{n=0}^3 \sum_{m=0}^4 e^{i(\omega_k t + \omega_l \theta)} F_{\xi}(\omega_k, \omega_l) \quad (4.13)$$

$$\sigma(t, \theta) = \sum_{n=0}^3 \sum_{m=0}^4 e^{i(\omega_k t + \omega_l \theta)} F_{\sigma}(\omega_k, \omega_l) \quad (4.14)$$

Note that the domain has been changed from $[0, 1]$ to $[0, 4]$ for the time t and $[0, 360]$ to $[0, 5]$ for the direction value θ in the initial Fourier transformation, hence, t and θ must also be changed according to the scale. Figure 4.10 illustrates the computed values of the discrete spectrum. The characterization of the shape and scale parameters is also illustrated in Fig. 4.11. The cumulative distribution function can be calculated by summing all the information within the sector of interest.

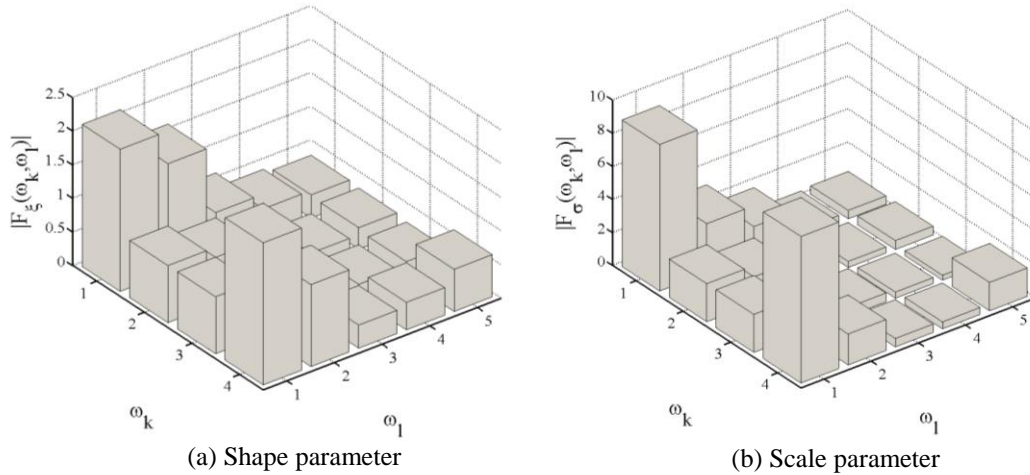


Figure 4.10 Discrete spectrum for shape and scale parameters after DFT.

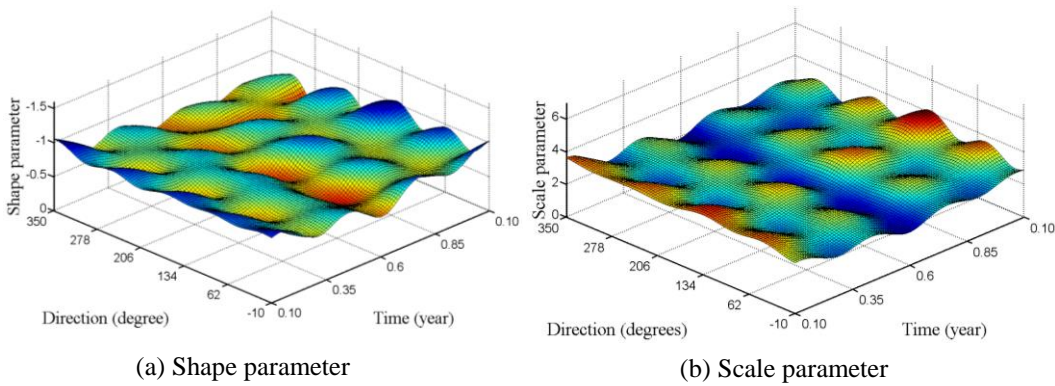


Figure 4.11 2D Fourier characterizations of shape and scale parameter changes along the time and direction axes.

To evaluate the performance resulting from this pair of transformation, the goodness-of-fit to the observed data for both the original discrete model and the smooth transition model are compared through the value of likelihood function. One should note the calculation of the likelihood function requires a slight modification in the probability density function for the derived model. If the varying properties of the parameters are taken into account, the probability

density function for each of the observed exceedances $\{x_j\}_{j=1}^n$ following a Pareto distribution is expressed as:

$$f_{x_j|\xi_j, \sigma_j}(x) = \frac{1}{\sigma_j(t_j, \theta_j)} \left(1 + \xi_j(t_j, \theta_j) \left(\frac{x_j - u_j}{\sigma_j(t_j, \theta_j)} \right) \right)_+^{-\frac{1}{\xi_j(t_j, \theta_j)}} \quad (4.15)$$

where t_j and θ_j are the seasonality and direction of the observed exceedance x_j ; and $\sigma(\cdot)$ and $\xi(\cdot)$ are the scale and shape parameters for the exceedance which can be calculated through the Fourier characterization depending on the observed values of t_j and θ_j . Hence, the likelihood function for the whole set of observed exceedances can be expressed as:

$$\begin{aligned} L\left(\{X_j\}_{j=1}^n; \Theta\right) &= \prod_{j=1}^n f(x_j; \Theta) \\ &= \prod_{j=1}^n \frac{1}{\sigma_j(t_j, \theta_j)} \left(1 + \xi_j(t_j, \theta_j) \left(\frac{x_j - u_j}{\sigma_j(t_j, \theta_j)} \right) \right)_+^{-\frac{1}{\xi_j(t_j, \theta_j)}} \end{aligned} \quad (4.16)$$

For ease of calculation, instead of the likelihood function $L\left(\{X_j\}_{j=1}^n; \Theta\right)$, the log-likelihood function $\log L\left(\{X_j\}_{j=1}^n; \Theta\right)$ is used for the comparison. The computed log-likelihood function of the constant discrete model has a value of -1589.5 whereas with smooth characterization, the computed value is -1509.9. Thus there is improvement to the model when smooth characterization is employed.

4.4.2 Model Validation

With a modification of the basic Poisson-GPD model equation (e.g. Eq. (3.8)), the improved cumulative distribution functions for sector S_i with n_i exceedances, can be derived as:

$$F_{S_i}(x) = \exp \left\{ -\lambda_{S_i} \sum_{j=1}^{n_i} \left(1 + \xi_j(t_j, \theta_j) \left(\frac{x - u_{S_i}}{\sigma_j(t_j, \theta_j)} \right) \right)_+^{\frac{1}{\xi_j(t_j, \theta_j)}} \right\} \quad (4.17)$$

where λ_{S_i} and u_{S_i} are the threshold and Poisson rate for the exceedances in the i th sector. $\xi(t_j, \theta_j)$ and $\sigma(t_j, \theta_j)$ are the shape and scale parameters of the j th exceedances in the sector, both of which are functions of time and direction. Note that n_i is number of exceedances observed from the collected data which covers the period of 25 years, and λ_{S_i} is the expected number of exceedances in a reference period P (years) which can be simply implemented as $\lambda_{S_i} = n_i P / 25$.

Once the cumulative distribution functions F_{S_i} for the sectors are obtained, the overall cumulative distribution function $F_{Overall}$, which is $P(x < X_o)$, can be calculated as the product of F_{S_i} by assuming independence between sectors. That is, $F_{Overall}(x) = \prod_{i=1}^{20} F_{S_i}(x)$.

The cumulative distribution functions are the basis for estimating the design values. Therefore, it is necessary to prove that the proposed model is suitable for use in structural reliability analysis. For this purpose, the cumulative distribution function of the developed model is compared with the empirical cumulative distribution function of the original data. The comparison is performed through a simulation study. For each of the time and directional sector,

a group of data is generated based on the GPD model as constructed in Section 4.3. Each data group is simulated for a period of 25 years (corresponding to the length of the collected data) based on its own Poisson rate. The comparison of the cumulative distribution functions between these simulated data and the original data for all the discretized sectors are shown in Figs. 4.12-4.16. The p-values of the KS test computed for each of the figure are also given. The results show that the cumulative distribution function of the simulated data agrees well with the original data since all the p-values are large (e.g. larger than the critical value at $\alpha=0.05$).

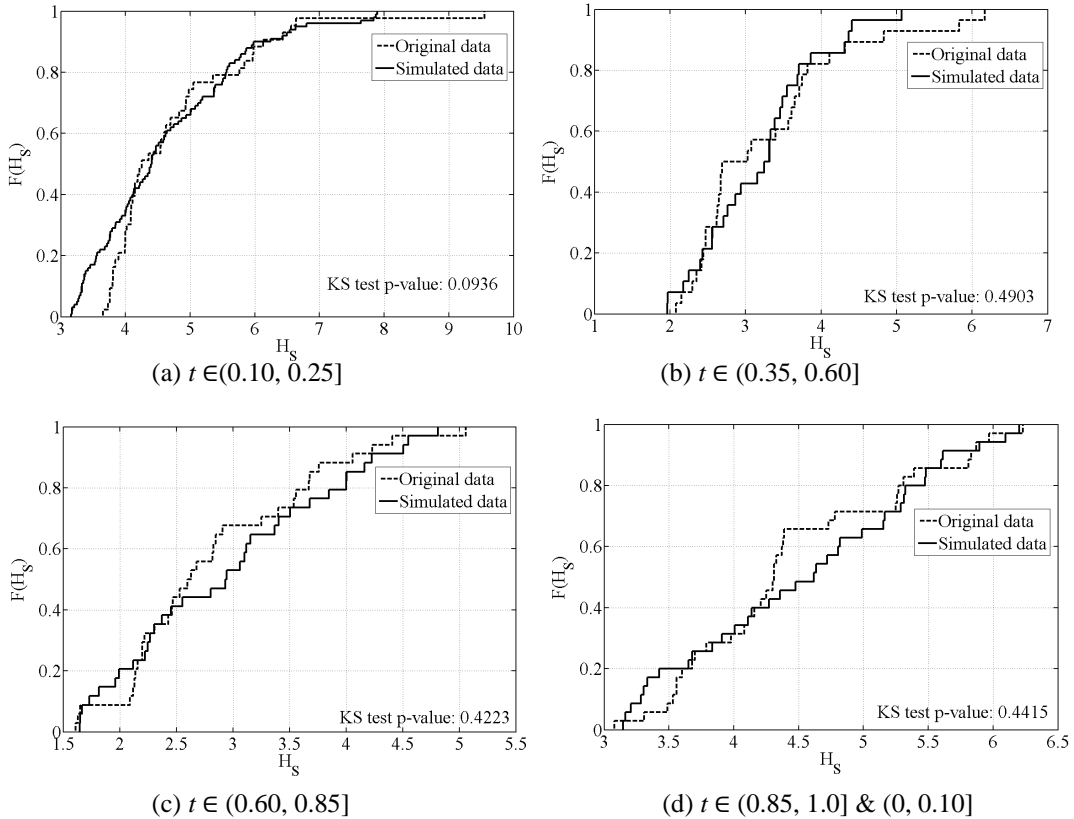


Figure 4.12 CDF of original data and simulated data for $\theta \in (-10^0 \sim 62^0)$.

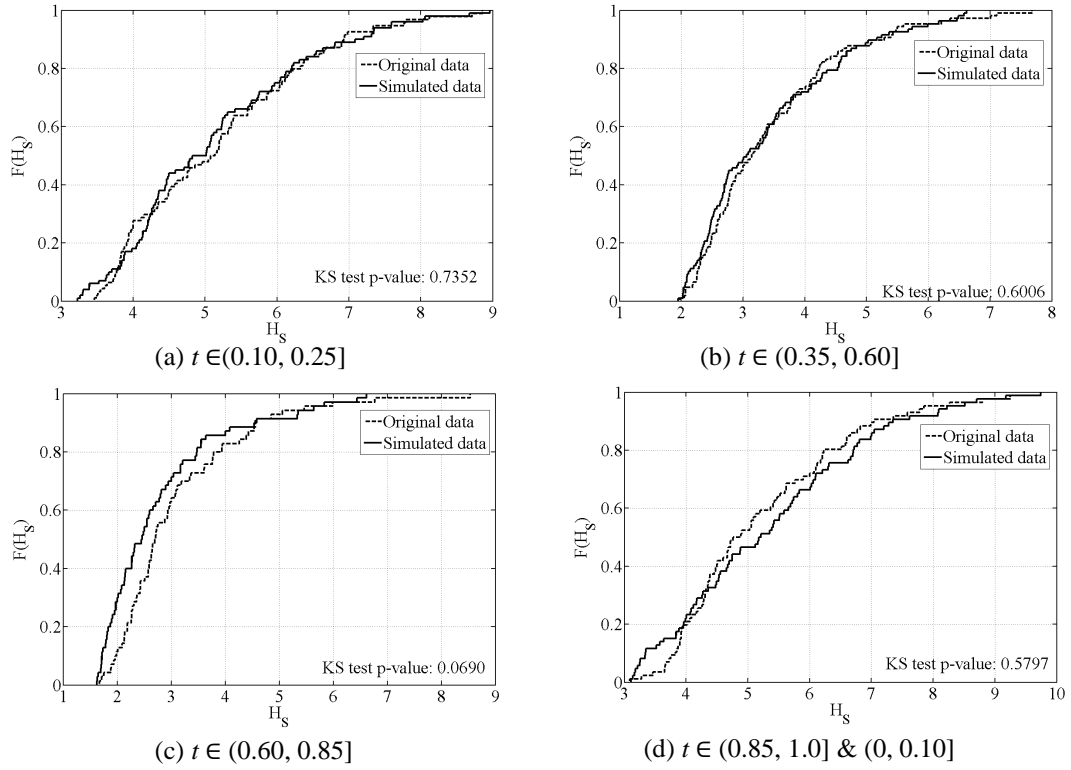


Figure 4.13 CDF of original data and simulated data for $\theta \in (62^\circ \sim 134^\circ)$.

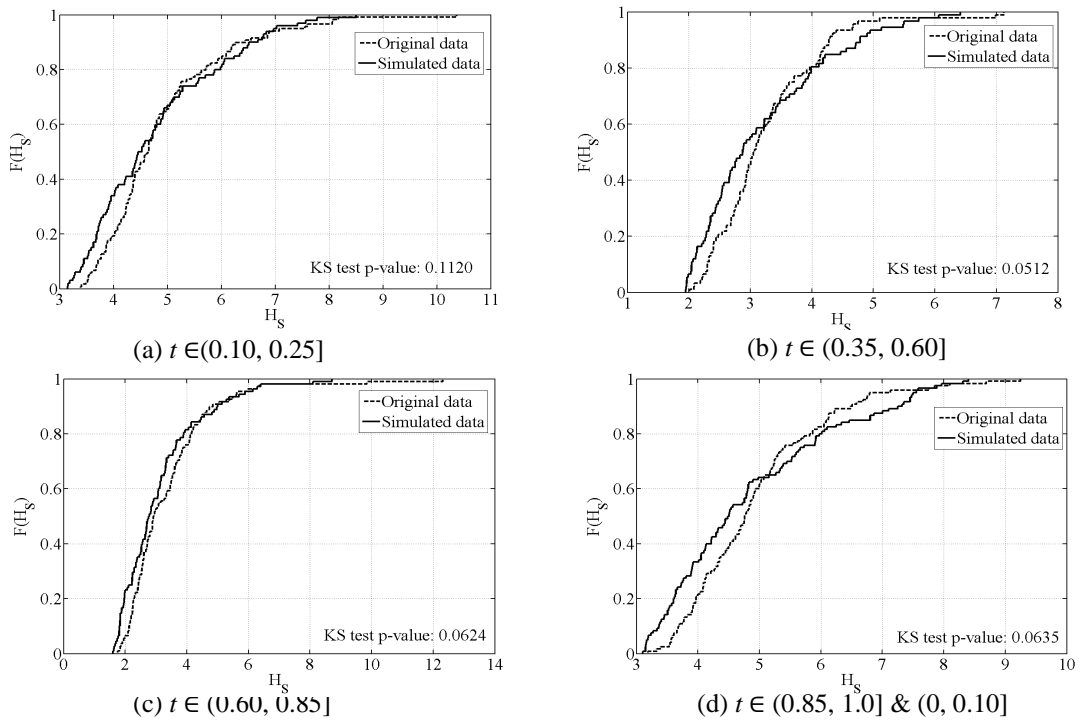


Figure 4.14 CDF of original data and simulated data for $\theta \in (134^\circ \sim 206^\circ)$.

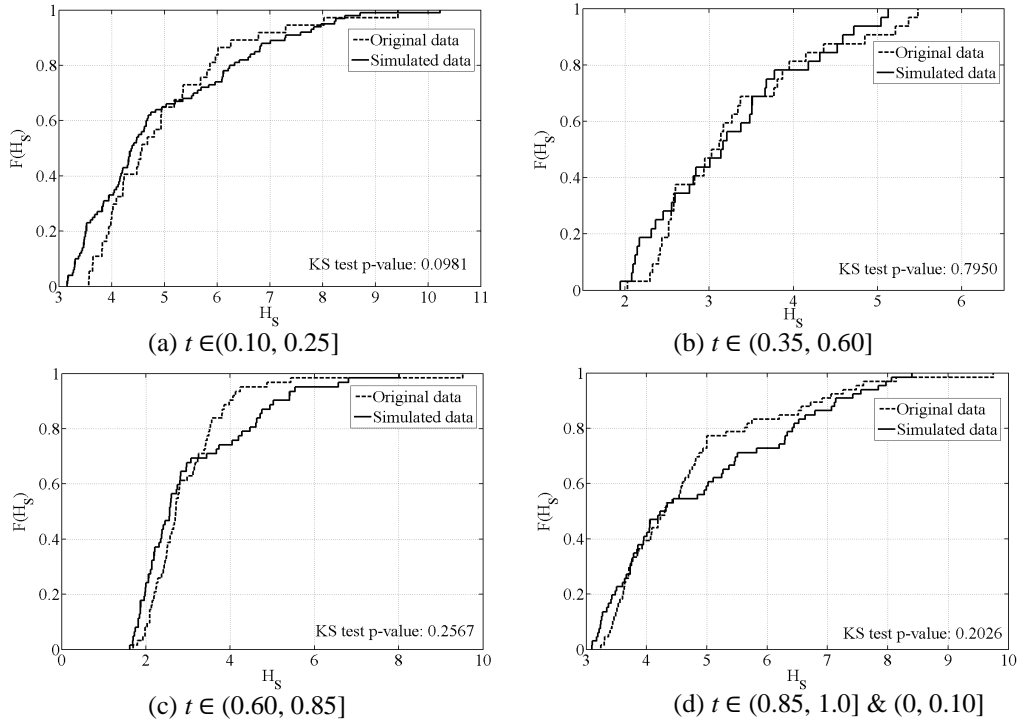


Figure 4.15 CDF of original data and simulated data for $\theta \in (206^\circ \sim 278^\circ)$.

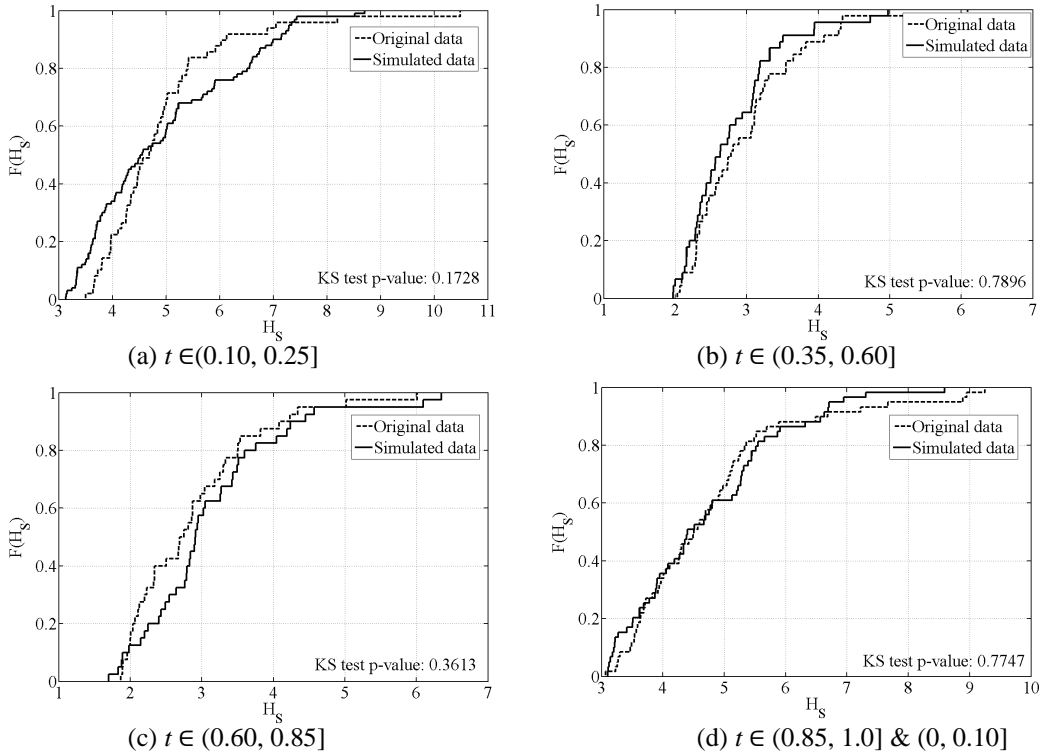


Figure 4.16 CDF of original data and simulated data for $\theta \in (278^\circ \sim 350^\circ)$.

One may notice that there are some deviations in the fittings to the original data for particular sectors (for example, the time sectors in Fig. 4.16). This is actually mainly because of the Fourier transform which tries to take care of the overall accuracy (by smoothing the sudden change in the parameters) instead of the accuracy of the individual sectors. It is a balance between accuracy in fitting the individual sectors and the overall fit. As such, the fitting in a sector's model is not so good, but from an overall design point of view, it may be more important as the return value is derived from the total exceedance probability. The individual is also important but when compared with the overall design criterion (e.g. $F_{Overall}(x) = \prod_{i=1}^{20} F_{S_i}(x)$), the importance is diminished. So from consistency point of view, the design and overall performance of the structure from a long term point of view is considered. However, if the interest is within particular directions or period (e.g. such as temporary operations or installation), then this fitting within that small sectors should be the focus and the analysis adjusted accordingly.

For further validation, the value of time and direction in the simulated data are compared against the original data. As shown in Fig. 4.17, the fluctuation of the mean with time of the simulated data follows similar pattern and magnitude as that of the original data in Fig. 4.17. The seasonal effect is clearly shown in the plot as large values are observed during winter and small values are observed during summer. As expected, the simulated data are quite close to the original data.

The directions are also compared to ascertain the quality of the proposed model. Figure 4.18 compares the cumulative distribution function for the simulated wave direction and the original wave data. A p-value of 0.5548 is computed in a two sample K-S test which implies adequacy of the developed model in characterizing the direction. It may be concluded from this simulation study that the developed model are both realistic and accurate.

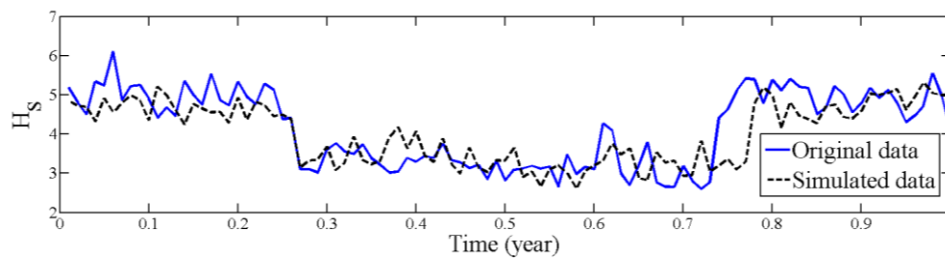


Figure 4.17 Comparison of simulated and original mean of H_S with time.

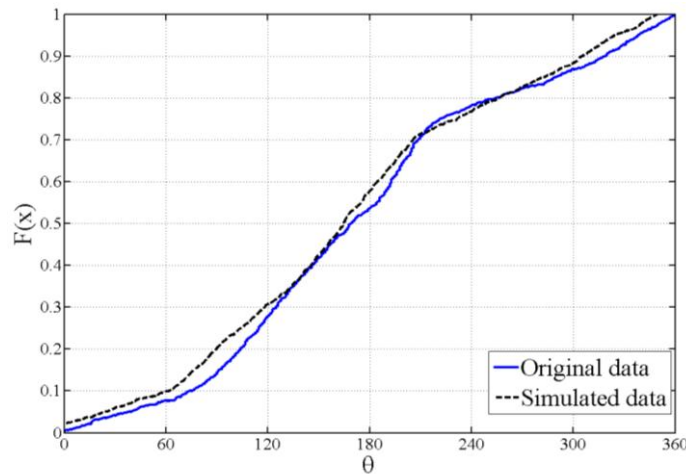


Figure 4.18 Comparison of simulated and original observed wave directions.

4.5 Static Push-Over Analysis

To demonstrate the performance of the proposed model for environmental loads in offshore structural reliability assessment, a fixed jacket structure taken from USFOS (2007) is examined in the following study.

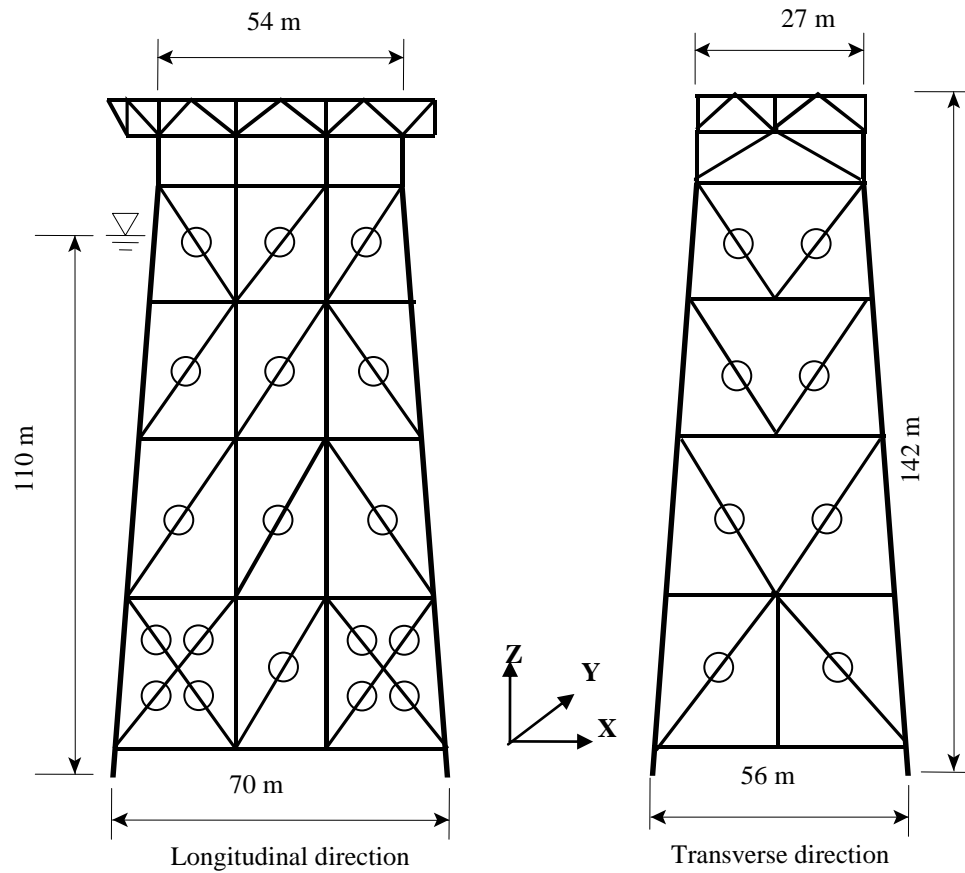


Figure 4.19 Jacket structure framing (circled elements are key structural elements in reliability analysis).

4.5.1 Structural Model Description

The jacket is designed for a water depth of 110 m with 8-legs arranged in a two by four rectangular grid, with the main loads coming from the north (laid

direction). The overall dimensions are 27 m × 54 m at the top and 56 m × 70 m at the bottom and the total height is 142 m. Horizontal bracings are installed at 5 levels. The jacket is resting on four corner clusters with eight skirt piles in each group and no leg piles are used. The longitudinal jacket frames are diagonal-braced, with X-braces between central and corner legs at the bottom bay. Transverse frames are K-braced, with the bottom K inverted to form a double X as shown in Fig. 4.19.

In the structural model, only bearing structure of the jacket is included in the analysis. The topside and the risers are not accounted for directly. The permanent loads and live loads are incorporated through nodal loads at the top level of the structure. Soil-structure interaction is simply modeled as linear springs. The detailed structural properties can be found from the example files in USFOS (2007). The detailed input files for constructing this model are given in Appendix E.

The proposed stochastic model in the preceding section is used to characterize the extreme wave height with the approximation $H_{max}=1.9H_S$ (DNV 2007). The reliability analysis is performed for the maximum wave height within 100 years, H_{100max} . That is, $P=4$ in the calculation of extreme frequency in each sector, and $\lambda_{Si} = n_i P / 25$ in Eq. (4.8). The current is assumed to have a velocity of 2m/s along with the same direction as the wave. The present study will not consider any long term climate change effects in the associated environment. That is, the statistical model of the environment parameter is assumed to remain the same within the 100 years considered in this study.

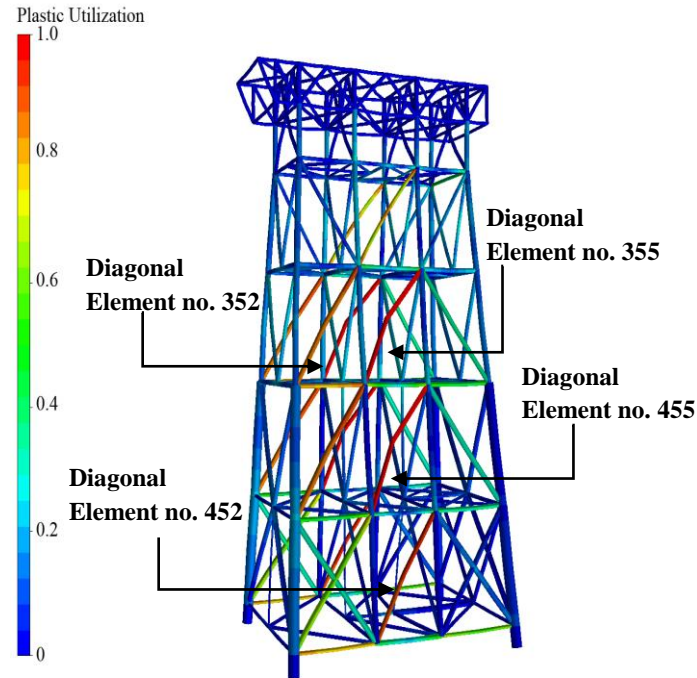


Figure 4.20 Plasticity utilization plot.

Besides the consideration of the wave characteristics, several uncertainties associated with the key structure’s mechanical properties are also considered herein. An initial push-over analysis is conducted on the structure for a longitudinal direction wave loading. Based on the ultimate strength analysis of the platform, it is found that the diagonal members below the sea level showed high degree of plasticity utilization before the base shear of the jacket failed, as shown in Fig. 4.20. It indicates the importance of the diagonal members over other members in contributing to the ultimate strength of the whole structure. Thus in this study, only the key elements of the jacket structure (marked with circles in Fig. 19) are selected where the uncertainties associated with manufacturing and corrosion effect (reduction in thickness) are studied. The yield strength of the steel, BS 968 for high strength tubes, is described by a lognormal distribution

with a suggested coefficient of variation of 0.05~0.08 (Baker 1969). The uncertainty in the thickness associated with corrosion effect is modeled as normally distributed with a coefficient of variation of 0.17 (Zhang 2010). The statistics of the random variables are given in Table 4.7.

Table 4.7 Random variables in reliability analysis.

Variables	Distribution	Mean	c.o.v.
F_y (MPa)	Lognormal	360	0.05
T (mm)	Normal	48	0.17

4.5.2 Reliability Analysis with Importance Sampling

With the uncertainties mentioned in the preceding section, the reliability against total collapse of a jacket structure is analyzed. The performance function associated with the ultimate limit state can be written as

$$G = R_{ultimate} - L_{100max} \quad (4.18)$$

where the ultimate strength of the structure $R_{ultimate}$ and the 100 year maximum load L_{100max} can be computed from the push-over analysis via USFOS. The analysis gives the probability of failure for the jacket in resisting a large wave height that may, on average, be observed once in 100 years. However, the performance function G is not known explicitly and as such, analytical solutions become formidable. It may only be determined point-wise when a push-over analysis is performed using the finite element model. Depending on the number of structural analysis required to evaluate the failure probability, the

computational cost may be exorbitant, especially if direct Monte Carlo simulation is employed. To improve numerical efficiency, the response surface method which utilizes a second-order polynomial to approximate the limit state function from a limit number of discrete numerical analysis is applied to construct an approximate performance function.

In analyzing $R_{ultimate}$ of the jacket structure, all the key tubular structural members beneath the seawater surface are assumed to have the same uncertainties in the thickness (e.g. T) and the yield strength (e.g. F_y). Thus, the equation associated with the ultimate limit state of this jacket can be expressed as:

$$R_{ultimate} = a_0^{\theta_i} + a_1^{\theta_i} F_y + a_2^{\theta_i} T + a_3^{\theta_i} F_y^2 + a_4^{\theta_i} T^2 \quad (4.19)$$

where $R_{ultimate}$ is in [MN], T is in [mm] and F_y is in [MPa]. The coefficients $\{a_0^{\theta_i}, a_1^{\theta_i}, a_2^{\theta_i}, a_3^{\theta_i}, a_4^{\theta_i}\}$ are determined through regression using data generated from Monte Carlo simulation. Note that the values of the coefficients $\{a_0^{\theta_i}, a_1^{\theta_i}, a_2^{\theta_i}, a_3^{\theta_i}, a_4^{\theta_i}\}$ may change with the direction θ_i . These are evaluated in each of the five directional sectors with respect to the structure, while assuming the performance function within the same directional sector does not change. Based on 100 simulated results, the values of the calculated coefficients for each directional sector are obtained and presented in Table 4.8.

Similarly, the response surface method can be applied to construct the function of the environmental loading L on the jacket in terms of the environmental parameters. For simplification, only the wave height H and the

current speed u are considered to be the random variables for predicting the total loading. L is evaluated at a set of discrete values of H and u using the push-over

Table 4.8 Determined coefficients of Eq. (4.19) for different directions.

	a_0	a_1	a_2	a_3	a_4
-10 °~62 °	-297.7612	63.4862	142.1718	-2.5144	-11.5600
62 °~134 °	-381.5760	69.2461	170.6110	-2.8295	-14.0079
134 °~206 °	-448.2937	73.4087	194.1363	-2.9642	-16.1396
206 °~278 °	-427.7499	70.7961	188.4236	-2.6406	-15.7409
278 °~350 °	-350.4651	81.9595	154.5041	-5.1180	-13.0788

analysis and approximated as an equation using the following simple form (Heideman 1980):

$$L = c_1 (H + c_3 u)^{c_2} \quad (4.20)$$

where u is the speed of the current in [m/s] and c_1 , c_2 and c_3 are constants. As the speed of current is assumed not changing (e.g. 2m/s), the factors related to the current speed can be considered as constants. The determination of these constants can be obtained from curve fitting, see Fig. 4.21. This gives the approximated equation for the response base shear as:

$$L = 0.0315(H + 8.2630)^{2.1058} \quad (4.21)$$

where L is the response base shear in MN, and H is the wave height in meters. The wave period is taken to be 16 seconds in the analysis. The estimating of L_{100max} will simply take the 100 year maximum wave height value H_{100max} .

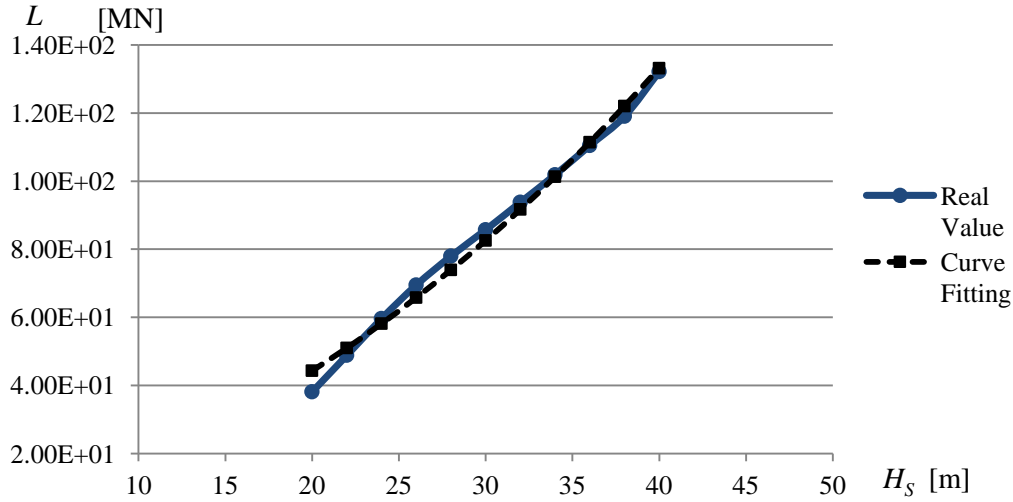


Figure 4.21 Comparison between calculated base shear and curve fit.

Hence, the performance function in Eq. (4.18) can be expressed in terms of F_y , T and H_{100max} as:

$$G = a_0^{\theta_i} + a_1^{\theta_i} F_y + a_2^{\theta_i} T + a_3^{\theta_i} F_y^2 + a_4^{\theta_i} T^2 - 0.0315 (H_{100max} + 8.2630)^{2.1058} \quad (4.22)$$

In order to calculate P_f efficiently, importance sampling is introduced and implemented in the numerical procedure as presented in Chapter 2. For this example, the three random variables in Table 4.7, denoted by $x_1=F_y$, $x_2=T$ and $x_3=H_{100max}$, are assumed to be independent. A sample size $N_{Pf} = 10000$ is used. The design point is determined by FORM (Ang and Tang 1984) where the independent random variables $\underline{X}=(x_1, x_2, x_3)$ are transformed to the equivalent standardized normally distributed random variables $\underline{U}=(u_1, u_2, u_3)$ using the transformation techniques as discussed in Section 2.2.3. The failure probability would then be calculated in the standard Gaussian space with a shift to the design point (u^*_1, u^*_2, u^*_3) as:

$$P_f = \frac{1}{N} \sum_{j=1}^N I_{[G(\underline{u}_j)]} \frac{\phi(u_1^j) \cdot \phi(u_2^j) \cdot \phi(u_3^j)}{\phi(u_1^j - u_1^*) \cdot \phi(u_2^j - u_2^*) \cdot \phi(u_3^j - u_3^*)} \quad (4.23)$$

where $\phi(\cdot)$ is the standard normal probability density function. The results for different time and direction sectors are summarized in Table 4.9. The result associated with the traditional extreme model which does not consider any of the seasonal and directional effects is also included.

The reason to compare the traditional approach and the current model is to show the importance of the proposed procedures in the safety assessment. However, it should be recognized that there may not be an “absolute” correct answer for the failure probability. It is not possible to find a large number of real structures in reality and use the records of failure to confirm the correctness of the results. The study and comparison performed here is only used to demonstrate the difference between the results if the seasonality and directionality are ignored in the wave height model.

Table 4.9 Failure probability of jacket structure for different time and directional sector of wave model.

		t_i				Total P_f (10^{-6})
		(0.10, 0.35]	(0.35, 0.60]	(0.60, 0.85]	(0.85, 1.0] & (0, 0.10]	
θ_i	(-10°, 62°]	1.41	1.41	0.567	0.714	4.10
	(62°, 134°]	4.72	4.21	3.43	4.16	16.5
	(134°, 206°]	13.5	11.4	11.3	12.3	48.4
	(206°, 278°]	9.72	8.10	7.78	8.71	34.3
	(278°, 350°]	1.30	1.02	0.928	1.27	4.52
Total P_f (10^{-6})		30.7	26.1	24.0	27.1	
Overall=						107.9
Traditional Extreme Model=						36.8

The calculated overall failure probability using the derived model is higher than that from the traditional model. In addition, the influence of the seasonal and directional effects is clearly shown. The first time sector has the largest failure probability of $P_f=3.07\times 10^{-5}$ and the third time sector has the lowest failure probability of $P_f=2.40\times 10^{-5}$. This difference between the time sectors is smaller compared to the differences between the directional sectors with the largest $P_f=4.84\times 10^{-5}$ at $(134^\circ, 206^\circ]$ and lowest $P_f=4.10\times 10^{-6}$ at $(-10^\circ, 62^\circ]$. The influence of the environmental loads on the structural reliability is not only associated with the load intensity but also with the frequency of occurrences. For example, the time segment $(0.35, 0.60]$ has a lower extreme wave height than the time segment $(0.85, 1.0]$ and $(0, 0.10]$, whereas the higher frequency of occurrence in directional sector $(62^\circ, 134^\circ]$ within the time segment $(0.35, 0.60]$ leads to a larger failure probability compared to time sector $(0.85, 1.0]$ and $(0, 0.10]$. This effect is primarily due to the frequency of occurrence of the extremes.

An extension to the above analysis is to investigate the influence of the assumed direction of the loads on the structural reliability and ultimately the design. For example, assume the main loads are coming from the transverse direction of the structure. In other words, instead of putting the structure along the north direction, the laid direction is now changed to along the east direction. This leads to a change in the performance function for the structural strength in different directions. Similarly, the coefficients $\{a_0^{\theta_i}, a_1^{\theta_i}, a_2^{\theta_i}, a_3^{\theta_i}, a_4^{\theta_i}\}$ in Eq. (4.19) are determined from regression analysis and the results recorded in Table 4.10.

Table 4.10 Determined coefficients of Eq. (4.19) for rotated structure.

	a_0	a_1	a_2	a_3	a_4
-10 °~62 °	-369.7863	79.1633	162.9958	-4.4986	-13.7395
62 °~134 °	-310.8566	68.2146	145.1774	-3.1826	-11.9147
134 °~206 °	-348.4256	65.4774	159.8829	-2.4828	-13.0418
206 °~278 °	-449.0569	75.1380	193.3730	-3.2631	-16.0473
278 °~350 °	-446.0514	68.2682	196.3785	-2.0717	-16.36398

Repetition of the analysis with the same environmental loads and structural specification leads to the results shown in Table 4.11. As expected, the most critical direction sector has changed from (134°, 278°] to (206°, 350°]. The overall failure probabilities for all the time segments has increased which is expected since the structure is weaker in direction where the main loads are acting. The rotated structure exposes the direction with its smallest resistance to the most adverse environment. In the example, although the overall failure probability of the rotated structure ($P_f=1.18 \times 10^{-4}$) differs only slightly from that of the original structure ($P_f=1.08 \times 10^{-4}$), the difference may be larger for a different structure. This indicates that the performance of the jacket structure within this environment is sensitive with respect to its “laid direction” when the directional wave load is the major concern.

The conclusions regarding the time and directional effects of the wave on the safety of the platform analyzed may not be generalized to other structures and ocean site. But the proposed method provides a general basis for the safety assessment of any newly designed or existing platforms with respect to the considered environment.

Table 4.11 Failure probability of rotated jacket structure with respect to wave model.

		t_i				Total P_f (10^{-6})
		(0.10, 0.35]	(0.35, 0.60]	(0.60, 0.85]	(0.85, 1.0] & (0, 0.10]	
θ_i	(-10°, 62°]	3.76	1.86	1.71	2.08	9.41
	(62°, 134°]	0.853	0.748	0.581	0.731	2.91
	(134°, 206°]	2.81	2.27	2.25	2.49	9.82
	(206°, 278°]	12.8	10.8	10.4	11.6	45.6
	(278°, 350°]	15.0	12.3	11.5	11.6	50.5
Total P_f (10^{-6})		35.3	28.0	26.4	28.5	
					Overall=	118.2

To sum up, the above analysis shows the needs to account for time and direction in the statistical modeling of the wave loads for a more complete reliability assessment of the structure. The influence of the time varying environment on the performance of the fixed jacket structure is more comprehensively reflected using the proposed approach.

4.6 Concluding Remarks

The effect of time and spatial variant characteristics of the significant wave height on the reliability of an offshore structure is studied. A stochastic model to account for both time and directional effects of wave loads is proposed, where a segmentation algorithm is used to treat the data to ensure that each segment satisfies the weak stationarity requirement of both the intensity and frequency of extreme waves. To model the parameters as not abruptly changing between the divided segments, 2-D Fourier transform is employed. The advantage of the proposed technique is demonstrated through the reliability analysis of an offshore

platform subjected to such time-variant environment and compared with that using the traditional reliability method. Based on the example analysis, the conclusions are summarized as follows:

- The characteristics of the wave, in terms of the magnitude and direction, were markedly different for different time periods. Issues related to the modeling of the occurrence rate and magnitude of extreme wave height from a stationary time series data have been addressed for using the Poisson-GPD model. It was found that the occurrence rate of the extremes is highly affected by the directional effect whereas the magnitude of the extremes is quite dependent on seasonal effect.
- The statistical analysis of time varying environment data requires a prior step of data grouping to approximately satisfy the assumption of stationarity of the data before it can be assessed. The use of discrete statistical model for the modeling of time varying extreme wave height has been examined using actual data. It was found that by incorporating such segmentation steps in the statistical model can help to model the underlying physical variable more accurately. The dependencies between the wave height and the time or direction can be largely reduced by the discretization step.
- The influence of modeling the seasonal effect is examined in a reliability analysis for an offshore platform. It was found that by incorporating such seasonal and directional effects in the statistical model can more adequately reflect the underlying physical processes and will help to detect the most critical environmental conditions for engineering structures in a marine

environment. This offers the engineer the opportunity to design structures via a full consideration of varying marine load.

It is important to note that all of these conclusions were drawn from the use of the WIS wave data and the structural analysis for the example jacket structure. It may be possible in other studies that the structure may not exhibit a significant time varying characteristic. However, the developed new ideas in the current work serves to aid the practicing engineers to assess the performance of an offshore structure within a time varying environment. The procedures in establishing a robust extreme model for the load related environment parameter remains the same.

Chapter 5 Modeling the Multivariate Environmental Condition for the Offshore Structural Analysis

In this chapter, the long term performance of offshore structures under multivariate environmental conditions is investigated. The study focuses on the modeling of multivariate random variables associated with ocean climate conditions experienced by offshore structures. The traditional joint distribution model may only be applicable to the multivariate problem to a certain extent. To extend to slightly more complex relationship such as non-constant correlation between pairs of random variables, the copula model is introduced. A comparative study will be done between the copula and some other available approaches in solving the same problem. Based on the constructed statistical model, a discretized copula method is proposed and demonstrated through determining the long term load associated with a specific level of reliability for a jacket structure problem. The accuracy and efficiency of this approach problem is also discussed at the end of this chapter.

5.1 Introduction

While addressing different load cases for the offshore structures, designers are usually required to estimate the environmental conditions at the ocean site, and usually a multivariate analysis is performed. The safety of the structure must be

ensured at least over the design life of the structure, which can amount to many years of exposure. Normally, the design of offshore structure employs a 50 or 100 years return value for the environmental parameter (DNV 2006). The long term assessment of offshore structure considering the dependency of wave height with the time and directionality has been discussed in Chapter 4. However, the stochastic nature of the environmental loads encountered by the structure can be complicated. The current practice in analyzing the seasonal and directional effects of wave load is limited to structures whose safety is most affected by the significant wave height H_S . For structures that are sensitive to more ocean parameters, for example wind speed or wave period, the static push-over approach may no longer be accurate enough to predict the response of the structure. In the actual situation, the number of significant parameters may be more and the influence of various multivariate models in predicting the long term performance of structures under such complex environment is of interest.

Besides the marginal distribution of the individual parameters, such as wave heights and wave periods, the dependency structure between various ocean parameters also affects the estimates of the response statistics and eventually the structural reliability. If the actual dependency is nonlinear, then the models currently used may be inadequate, as revealed by several researchers (Guedes Soares et al. 2001; IPCC 2007; Vanem 2011). A number of approaches have been developed in the context of multivariate analysis in wave climate studies and other areas, as will be presented in the following section. Despite this, it is recognized that they are not widely applied in engineering practice. This could be

due to lack of clear criteria for selecting the appropriate models compounded by the fact that different groups of data may have their own characteristics. A second challenge is that the processing of the established environment parameter model in the long term assessment requires an effective technique in the direct integration method, as shown in Eq. (2.2). As the stochastic analysis for the offshore structure within a given environment condition is very time consuming, the direct integration based on environmental parameters simulated from the constructed multivariate model is cumbersome and impractical. Several recent works have emphasized the complexity of these issues in offshore environment and have addressed comparison of alternative numerical methods and model test (Cheng et al. 2003; Agarwal and Manuel 2009; Baarholm et al. 2010), the possible reduction in simulation efforts by selecting the critical sea states (Norton 2004), the use of environmental contour and inverse first order reliability method (IFORM) (Saranyasontorn and Manuel 2005), and the suitable use of bootstrap method in deriving the confidence interval from the structural analysis (Efron and Tibshirani 1993).

This chapter aims to answer two questions. Among the available multivariate models, which of these give the best performance in modeling the ocean parameters; and how to implement the multivariate models into the structural analysis in order to assess the long term safety of the offshore structures. To address these issues, this chapter will adapt the copula approach in constructing a multivariate statistical model using measured data taken from a

buoy at the southern coast of Alaska. The basic time domain wave simulations form the basis for the structural statistical analysis in this study.

5.2 Bivariate Models for Sea State Parameters

5.2.1 Conditional Joint Distribution Model

Among the probabilistic models available in the literature, the most commonly recommended model adopted by the design code is the conditional joint distribution model which is widely applied to a variety of problems (Burton et al. 2001; Jonathan and Ewns et al. 2011; Ernst and Seume 2012). The joint distribution model that is most pertinent in marine engineering is the distribution of significant wave height (H_S) and peak period (T_P) which characterizes the occurrences of sea states. Ochi (1992) introduced a bivariate log-normal distribution in modeling the behavior of sea states (H_S , T_P). Haver (1985), Guedes Soares et al. (1988), and Bitner-Gregersen and Haver (1989) have applied a joint environmental model to different engineering calculations based on the combination of marginal distribution of H_S and various conditional distributions of T_P . The basic reason of such conditional distribution model is generally because H_S is the parameter that affects most design conditions of ocean structures while T_P has less influence. According to the practice design code (DNV 2012), the conditional bivariate distribution model generally assumes H_S follows a Weibull distribution while T_P follows a lognormal distribution whose model parameters are conditional on H_S which can be expressed by:

$$f_{T_p|H_s}(t|h) = \frac{1}{\sqrt{2\pi\sigma t}} \exp\left\{-\frac{(\ln t - \mu)^2}{2\sigma^2}\right\} \quad (5.1)$$

where t and h are variables representing the peak period and significant wave height, and parameters μ and σ are functions of h .

$$\mu = E[\ln T_p] = a_1 + a_2 h^{a_3} \quad (5.2)$$

$$\sigma^2 = \text{Var}[\ln T_p] = b_1 + b_2 e^{b_3 h} \quad (5.3)$$

and $a_1, a_2, a_3, b_1, b_2, b_3$ are coefficients which could be determined from the fitting of the parametric model to the data through the regression analysis.

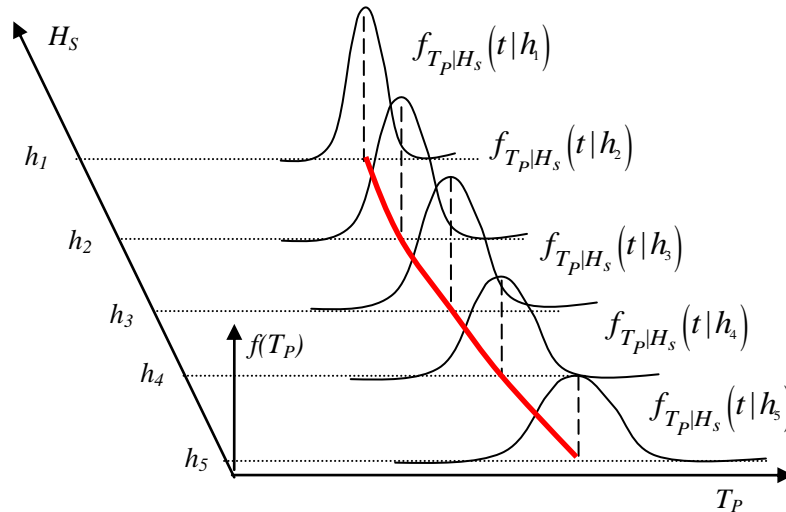


Figure 5.1 Sketch of several conditional probability density curves of T_p at various values of H_s .

This concept of conditional distribution models can also be found for the other ocean parameters from the design code (DNV 2012). The conditional

distributions for selected significant load governing parameters are listed in Table 5.1.

Table 5.1 Conditional joint distribution models for selected stochastic ocean parameters (DNV 2012).

Parameter	Description	Distribution
$F_{H H_s}$	Wave height distribution in a stationary sea state (storm)	Conditional Weibull Distribution $F_{H H_s}(h h_s) = 1 - \exp\left\{-\delta\left(\frac{h}{h_s}\right)^\alpha\right\}$ $\alpha=2.13 \quad \delta=2.26$
$F_{H_{\max} H_s}$	Largest wave height in a stationary sea state	$F_{H_{\max} H_s}(h_{\max} h_s) = [F_{H H_s}(h_{\max} h_s)]^{N_{\text{wave}}}$ $\approx \exp\left\{-N_{\text{wave}} \exp\left[-\delta\left(\frac{h_{\max}}{h_s}\right)^\alpha\right]\right\}$
$F_{V_w H_s}$	1-hour mean wind speed	Conditional Normal Distribution $\mu=5.5+1.8h_s \quad \sigma=4.9-0.26h_s$
$F_{V_c H_s}$	Current speed at mean sea level	Conditional Normal Distribution $\mu=0.2+0.04h_s \quad \sigma=0.1$

Other than these pioneering research works, many of the current studies demonstrate several points that could be further developed. Prince-Wright (1995) showed that a maximum likelihood model using the Box and Cox (1964) transformation of joint sea state parameter data to a Gaussian model can model the collected data well, especially under the presence of non-stationarities. This has been proven through a comparative study, in which the analysis has been

applied to both wave and wind data, by Bitner-Gregersen et al. (1998). A similar approach can also be seen in the work by Ferreira and Guedes Soares (2002) where the transformation technique is well applied. When considering the ocean characteristics, Bitner-Gregersen (2005) investigated the probabilistic description of the sea states for including a wind and a swell component which could give a more realistic characterization. Repko et al. (2004) developed a bivariate model of significant wave height and peak period based on a given independent value of wave steepness. A comparison of several of these approaches with a bivariate maximum entropy distribution, including clarifying the differences and performance in estimating a return value, by utilizing a specific data set can also be found in (Dong et al. 2013). Generally, under certain conditions each of the provided model has their own advantages.

5.2.2 Nataf Model

If limited information are available, such as insufficient collected data, the establishing of joint probabilistic model using the Nataf approach attracts considerable attention in engineering. In applying Nataf method, the data are transformed from the original space to mutually independent standard normal variates. The joint density probability distribution for the multiple variables can be expressed as a function of the marginal distributions and the correlation coefficients amongst all pairs of variables (Kiureghian & Liu 1986):

$$f_{x_1, \dots, x_n}(x) = f_1(x_1) f_2(x_2) \cdots f_n(x_n) \frac{\phi_n(\mathbf{y}, \boldsymbol{\rho}_\theta)}{\phi(y_1) \phi(y_2) \cdots \phi(y_n)} \quad (5.4)$$

$$\begin{cases} \Phi(y_i) = F_i(x_i) \\ y_i = \Phi^{-1}(F_i(x_i)) \end{cases} \quad i = 1, 2, \dots, n \quad (5.5)$$

where $F_i(\cdot)$ is the marginal distribution of the i th variable, $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative and density functions of standard normal variables. Each coefficient ρ_{ij}^{θ} in the correlation matrix $\boldsymbol{\rho}_\theta$ can be determined by evaluating the integral:

$$\rho_{ij} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\frac{F_i^{-1}(\Phi(y_i)) - \mu_i}{\sigma_i} \right) \left(\frac{F_j^{-1}(\Phi(y_j)) - \mu_j}{\sigma_j} \right) \phi(y_i, y_j, \boldsymbol{\rho}_{ij}^{\theta}) dy_i dy_j \quad (5.6)$$

where ρ_{ij} is the correlation coefficient between i th variable and j th variable in the original sample set. The Nataf approach is most appropriate for dealing with problems where only partial information is available, for example, only marginal distribution and correlation coefficients are known. It only requires the first two statistical moments and the marginal distributions of each of the marginal variables in the joint distribution as well as the correlation between the variables. Unlike the rigid conditional bivariate distribution model as described in Section 5.2.1, Nataf model can also be easily extended to multivariate models.

The Nataf approach has been successfully applied to model the ocean parameters. Wist et al. (2004) used the Nataf model for characterizing successive wave heights and successive wave periods. A more complicated methodology has been presented by Sagrilo et al. (2011) for creating a Nataf model which includes

the wave, wind and current parameters. It is also applied in a structural reliability analysis while an environmental contour is estimated from the Nataf model (Silva-González et al. 2013). The feature of the approximation for the distribution of the physical variables depends on whether the vector of the transformed standard normal variables is close to being multi-normal.

5.3 Copula Theory

Copula is another alternative tool for dealing with multivariate data problems, and has become very popular in the past two decades, where broad applications emerged in economics (Patton 2002), science (Embrechts 2009), hydrology (Zhang and Singh 2005; Genest and Favre 2007; Salvadori et al. 2007), finance (Cherubini et al., 2004; Genest et al. 2009) etc. The word copula originates from Latin, meaning “tie, connection or link”. Hence, from this basic meaning behind this word, one can know that copula is a model that is able to “couple” univariate marginal distributions to a multivariate distribution. In practice, the copula model is used to construct a multivariate distribution function by combining the marginal distributions and specific dependence structures between the variables.

5.3.1 Definition and Basic Properties

The original definition of copula is introduced by Sklar (1959). However, a later work done by Nelson (2006) gives a more detailed and easier to understand

introduction to the basic theories of copula. The basic definition of two dimensional copula is given by the following:

Definition of copula: A 2-dimensional copula is a function $C: [0, 1]^2 \rightarrow [0, 1]$ with the following properties:

1. For every u in $[0, 1]$,

$$C(u, 0) = C(0, u) = 0 \quad (5.7)$$

and

$$C(u, 1) = C(1, u) = u \quad (5.8)$$

2. For every u_1, u_2, v_1, v_2 in $[0, 1]$ with $u_1 \leq v_1$ and $u_2 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \quad (5.9)$$

These properties are the basic requirements for a copula function which actually parallels the essential important properties of multivariate distribution model. One may interpret property 1 as the *grounded* property of copula and property 2 as the *non-decreasing* feature of the copula function. The extension to a multivariate case is theoretically straightforward and convenient while the domain is changed to higher dimensions. Hence, an n -dimensional copula function, or simply *n-copula*, has the domain and the range values as:

$$C: [0, 1]^n \rightarrow [0, 1] \quad (5.10)$$

The constructed *n-copula* must also obey the rules in the basic definition. For example, the marginal uniformity implies that:

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i \quad \forall i = 1, \dots, n \quad (5.11)$$

With these formal definition and properties of copula, a multivariate distribution with specified marginal distributions can be constructed according to Sklar's theorem:

Sklar's Theorem: Let F be an n -dimensional distribution function with marginal distributions F_1, \dots, F_n . There exists an n -dimensional copula C such that for all $x \in \mathbb{R}^n$

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (5.12)$$

If F_1, \dots, F_n are all continuous, then C is unique. Otherwise, C is uniquely determined on $\text{Ran}(F_1) \times \dots \times \text{Ran}(F_n)$, where $\text{Ran}(F_i)$ is the range of the function F_i . Conversely, if C is a copula and F_1, \dots, F_n are distribution functions, then the function defined in Eq. (5.12) is a multivariate distribution function with marginal distributions F_1, \dots, F_n .

On account of Sklar's theorem, one could see that the copula does not cater to the individual behavior in the multivariate problem. Compared to the traditional joint model, the main advantages of copula approach is that the selection of an appropriate model for the dependence among the variables can proceed independently from the choice of the marginal distribution.

5.3.2 Examples of Copula

Various kinds of copula families and classes are available in the literature (Nelson 2006). Each family or class of copula has its own characteristics that could show its advantages for certain types of data. Most of these copulas are built for a bivariate case which may be expanded to a multivariate model quite easily for the same copula family or class. In this section, the Gaussian copula and the Archimedean copula, which are frequently used in practice, are discussed.

5.3.2.1 Gaussian Copula

Gaussian copula appears quite often in risk analysis solutions where a transformation step for the original variables is needed. The dependence structure could generally be written in a function with multivariate normality. For a bivariate case, it is expressed as:

$$C_{Gaussian}(u, v, \rho_{uv}) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho_{uv}^2}} \exp\left(\frac{2\rho_{uv}uv - u^2 - v^2}{2(1-\rho_{uv}^2)}\right) dudv \quad (5.13)$$

where ρ_{uv} is the parameter for the bi-dimensional normal vector. The Gaussian copula is a member of the elliptical copula family which is of elliptical contour distributions. The dependence is modeled by using a symmetric and positive definite matrix which describes the dependence between pairs of variables. Because of its tractability, it is widely applied to different risk analysis problems (Renard and Lang 2007). A shortcoming of Gaussian copula is that all the data within the domain shares the same linear dependency. If the dependence structure is not consistent over the entire domain of the variables, it may not be possible to

accurately characterize the data. More comprehensive reviews of Gaussian copula can be seen in Lebrun and Dutfoy (2009).

5.3.2.2 Archimedean Copula

The Gaussian copula mentioned above belongs to the family of elliptical copulas which only have limited features of symmetric dependence. However, in modeling the ocean parameters, it is usually common to observe large dependencies in the extreme values. For example, a hurricane or storm may induce both large values in wave height and wind speed. Under this situation, the Archimedean copulas could be introduced to model such asymmetries in the ocean data. The Archimedean copula is easy to construct and is established based on an algebraic method using a generating function $\varphi(\cdot)$:

$$C_{Archimedean}(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (5.14)$$

where $\varphi: (0, 1] \rightarrow [0, \infty)$ is a strictly decreasing convex function with $\varphi(1)=0$, and it is assumed that $\varphi^{-1}(t) = 0$ for all $t > \lim_{u \rightarrow 0} \varphi(u)$. Several well known copula families belong to the Archimedean copulas, and have been applied widely in the literature (Vandenberghe et al. 2011; Ariff et al. 2012; Corbella and Stretch 2013). The formulas of the most applicable one parameter Archimedean copulas such as Gumbel, Frank, Clayton and Ali-Mikhail-Haq are listed in Table 5.2.

Table 5.2 Examples of Archimedean Copulas (Frees and Valdez 1999; Cherubini et al., 2004; Nelson 2006).

Copula	Bivariate Formula $C_{\theta}(u, v)$	Generating Function $\varphi_{\theta}(t)$	$\theta \in$
Gumbel	$\exp\left\{-\left[(-\ln u)^{\theta} + (-\ln v)^{\theta}\right]^{\frac{1}{\theta}}\right\}$	$(-\ln t)^{\theta}$	$[1, +\infty)$
Frank	$\frac{1}{\theta} \ln\left(1 + \frac{(e^{u\theta} - 1)(e^{v\theta} - 1)}{e^{\theta} - 1}\right)$	$\ln \frac{e^{\theta t} - 1}{e^{\theta} - 1}$	$(-\infty, +\infty)$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$	$t^{-\theta} - 1$	$(1, +\infty)$
Ali-Mikhail-Haq	$\frac{uv}{1 - \theta(1-u)(1-v)}$	$\ln \frac{1 - \theta(1-t)}{t}$	$[-1, 1)$

An illustration of the differences in the characteristics of these famous copulas can be seen in the scatter plot shown in Fig. 5.2 where all of them have the same marginal distributions. Compared with Gaussian copula, one could easily see that the Gumbel copula is more appropriate for data which has stronger correlations at high values (less spread) but less correlated at low values while the Clayton copula better describes data exhibiting strong left tail dependence. The Frank copula is most appropriate for data that have relatively weaker dependencies at the tails. In practice, the change in dependencies with the values in the collected ocean data is quite common and hence a suitable model should be used to capture such characteristics shown in the observed data. More comprehensive reviews regarding Archimedean copula and Gaussian copula can be found in Genest and MacKay (1986) and Nelson (2006).

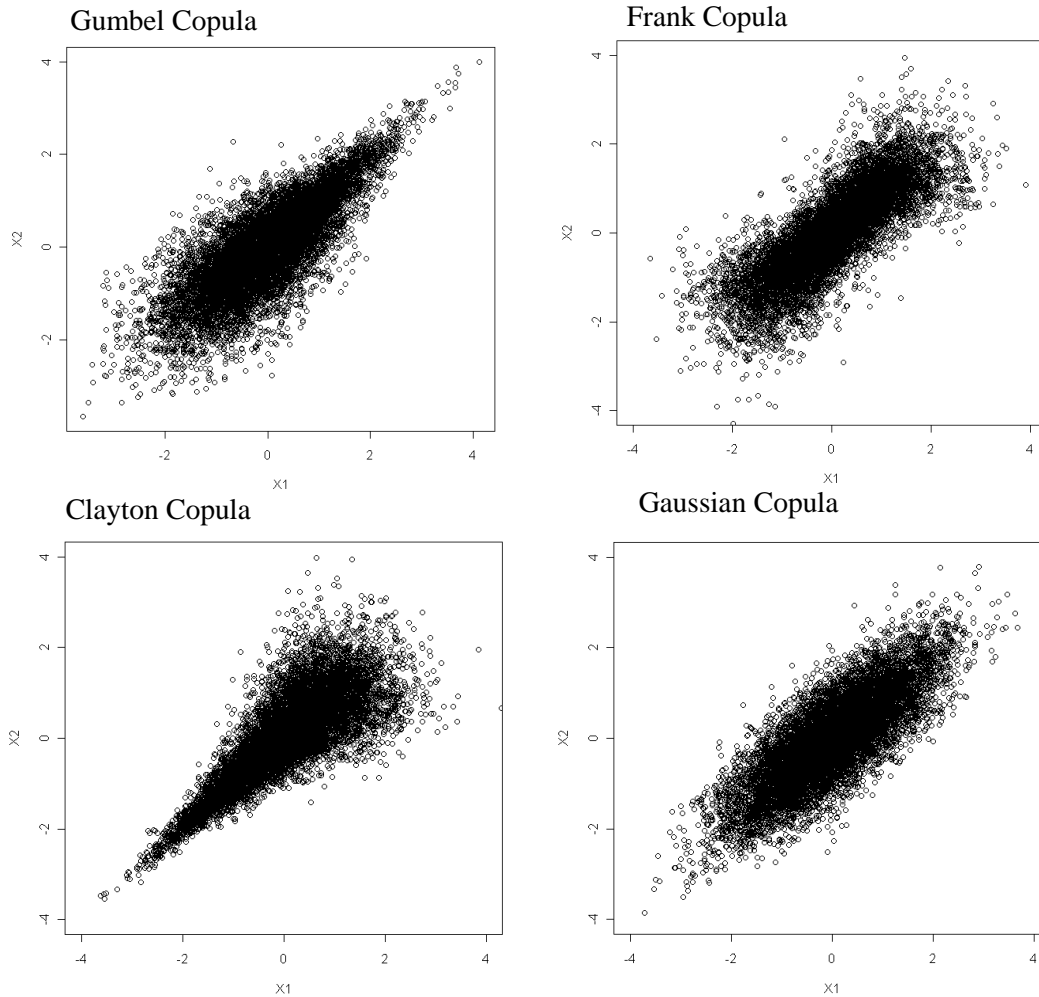


Figure 5.2 Comparison of different bivariate copulas for marginal distributions following standard normal distributions for correlation coefficient equals to 0.8.

5.3.2.3 Copula Parameter Estimation

The above well-developed copulas belong to the parametric copula multivariate models. In estimating the parameters in the copula, the asymptotic maximum likelihood estimation (MLE) may be employed. Based on Sklar's theorem, the density of the multivariate distribution function can be obtained as:

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \times \prod_{i=1}^n f_i(x_i) \quad (5.15)$$

where $f_i(x_i)$ is the marginal probability density function of the i th variable, $c(\cdot)$ is the density function of the copula which can be derived as the n th partial derivative of copula function C :

$$c(F_1(x_1), \dots, F_n(x_n)) = \frac{\partial^n (C(F_1(x_1), \dots, F_n(x_n)))}{\partial F_1(x_1) \partial F_2(x_2) \dots \partial F_n(x_n)} \quad (5.16)$$

This canonical representation of the multivariate probability density function generally separates the likelihood function into two parts and the statistical fitting of a copula model is decomposed into two steps:

- Fitting the marginal distributions to each of the variable.
- Fitting a copula function for the transformed multivariate data.

Therefore, let the observed data set to be $\{x_{1j}, \dots, x_{nj}\}_{j=1}^m$, the expression for the log-likelihood function can be written as:

$$l(\Theta) = \sum_{j=1}^m \ln c(F_1(x_{1j}), \dots, F_n(x_{nj})) + \sum_{j=1}^m \sum_{i=1}^n \ln f_i(x_{ij}) \quad (5.17)$$

where Θ includes all the parameters in both marginal distribution and copula function. The maximization of this equation would give the estimated values for all the parameters. However, in most of the cases, the number of parameters is quite large, which may lead to difficulties in obtaining the solutions. It is thus more common to estimate the parameters separately based on the maximization of the marginal distribution likelihood function and copula likelihood function:

$$\hat{\theta}_{MLE} = \begin{cases} \max_{\theta_1 \in \Theta} \sum_{j=1}^m \ln c(F_1(x_{1j}), \dots, F_n(x_{nj})) \\ \max_{\theta_2 \in \Theta} \sum_{j=1}^m \sum_{i=1}^n \ln f_i(x_{ij}) \end{cases} \quad (5.18)$$

where θ_1 are the parameters in the copula function and θ_2 are the parameters in the marginal distribution function. Such an adaptation of the original approach is called the maximum pseudo-likelihood method which is widely used (Genest and Favre 2007).

5.3.3 Dependence Concepts

To highlight the significance of the copula approach, a detailed interpretation of dependence concepts is given in the following. The easiest to comprehend and convenient concept of dependency is Pearson's correlation coefficient ρ , which is a measure of linear dependence between random variables and is defined as:

$$\rho(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sigma_{X_1} \sigma_{X_2}} \quad (5.19)$$

where $\text{cov}(X_1, X_2) = E(X_1, X_2) - E(X_1)E(X_2)$ is the covariance between X_1 and X_2 , $E(\cdot)$ is the expectation function and σ_x, σ_y are the standard deviations of X_1 and X_2 . The correlation coefficient is adopted in most of the multivariate analysis because it is a very common statistics that can be straightforwardly estimated from the data. However, several shortcomings are associated with this definition:

1. The correlation coefficient basically reflects the linear relationship between the coupled variables. In the case of perfect linear dependence where $\rho = 1$, the dependency is well represented. In the case of imperfect linear dependence where $-1 < \rho < 1$, the correlation coefficient is at best a “first-order” representation of the dependency.
2. The linear correlation coefficient is invariant with respect to linear transformations of the variables. However, it is not invariant to strictly increasing nonlinear transformations. The property of linear dependency may not be preserved through such transformations.
3. Moreover, the correlation coefficient is not scale invariant. For example, if X_1 have relatively much larger values compared to X_2 , X_1 would have larger influence on the value of $\rho(X_1, X_2)$.

In view of these limitations, the rank correlation coefficient corresponding to the measure of rankings between two variables may be more appropriate. The two most well-established concordant measures of such rankings are Kendall’s τ_k and Spearman’s ρ_s . Both coefficients will give value of 1 for variables that have perfect monotonic relationship and -1 for variables that have perfect counter monotonic relationship. For any bivariate copula, these two coefficients can be directly linked to the copula function as (Nelson 2006):

$$\tau_k(u, v) = 4 \int_0^1 \int_0^1 C(u, v) dC \quad (5.20)$$

$$\rho_s(u, v) = 12 \int_0^1 \int_0^1 C(u, v) dC - 3 \quad (5.21)$$

The most proper copula should give the best characterization in the dependence between the ocean parameters. And by combining with the marginal distribution, the parametric model should give the best fit to the observed data.

5.4 Comparative Study in Multivariate Modeling

In this study, the investigation of the long term performance of offshore structure will primarily rely on the statistical models that have been built for the ocean parameters. In order to have a proper perspective of the versatility of copula approach, it is useful to compare the performances of the available approaches in modeling the same multivariate ocean data. Therefore, for the comparison purpose, this section will go through the construction of bivariate model by using the aforementioned approaches, namely, (a) conditional joint distribution model, (b) Nataf model, and (c) copula model. The analysis of interest concerns the modeling of wave height and wave period which will be further utilized in the structural dynamic analysis to obtain the conditional response distribution.

5.4.1 Data Pre-treatment

The collected data set chosen for this analysis is also taken from the Wave Information Studies (WIS 2012) as mentioned in the Chapter 3 and 4. The location of the collected data is at the south coast of Alaska (52.00 °N, 172.00 °E; Buoy No. 82442) and the water depth at the location is 150m. The data set spans

26 years (1985/1/1 01:00 -2011/1/1 01:00) of hourly records. Several ocean parameters are contained in this data set. These include: significant wave height H_S (m), peak wave period T_P (s), wave direction θ_S ($^\circ$), wind speed V_W (m/s) and wind direction θ_W ($^\circ$). Values of θ_S and θ_W (angle in degrees) refer to the direction measured clockwise from the North. The long term record shows variations in both the magnitude and direction of wave and wind at this location. A clear seasonal change pattern could be observed from the mean H_S and V_W in Fig. 5.3, which indicate more severe condition in winter compared with summer.

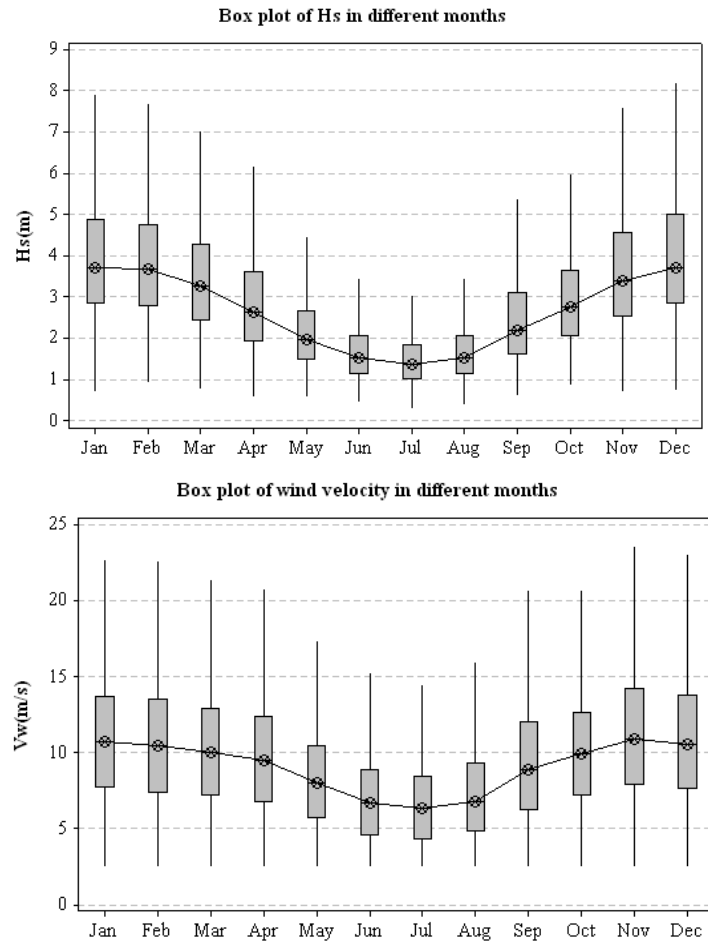


Figure 5.3 Box plot of H_S and V_W over different months.

As discussed in Chapter 4, the time varying effect may induce large uncertainties in the parameter estimates. Thus, it is not possible to apply the statistical analysis directly to the multivariate data sample. In general, a pretreatment based on statistical criteria is needed to separate out different groups of data for the analysis. For such a multivariate ocean data set, this procedure even needs more efforts. Several criteria for defining a “good” multivariate data set include:

- (1) Stable values of individual statistical parameters of interest (e.g. the mean, variance of H_S , T_P or V_W).
- (2) Constant likelihood of occurrence of the physical variables (e.g. occurrences rate of θ_S and θ_W).
- (3) Stable relationship between the parameters.

It is observed that besides property 1, which has been highlighted in Fig 5.3, the other two properties also have time-varying features. Figure 5.4 shows the observed values of (H_S, T_P) for two different months. The scatter diagram illustrates that there is greater dependence during November than during August. This may be explained by the fact that severe season would lead to more extreme events such that large values in the ocean parameters will tend to occur simultaneously. This is also reflected in the respective directions of the wave and the wind speed, see Fig. 5.5. In the present study, interest is paid to the evaluation of structural response associated with the critical environmental conditions during operation, and such low values in the wave height or wind speed will be of less

concern. Therefore in the following study, a specific group of data covering the most severe period of four months (Nov-Feb) is chosen for analysis. The variations of the mean and standard deviation of H_S , T_P and V_W over these months are relatively small, as illustrated in Fig. 5.3, and can be accounted within the same statistical model. Of course the data for the period selected may not be perfectly homogeneous, but the time varying nature especially for the two major load governing factors H_S and V_W has been removed to a sufficient extent to improve the accuracy for further statistical modeling.

As discussed in Chapter 4, another important issue regarding the directional effects must be considered. It is necessary to divide the data according to different occurrences along the directional sectors. Figure 5.6 shows a rose plot summarizing the wave height magnitude H_S and its direction θ_S for the selected period of data. Wave directions between 180° and 270° occur more often than the others, and a relative greater frequency of occurrence of large wave height is also associated with this wave directional sector. It is therefore important to focus on constructing a good probabilistic model to describe the wave behavior within a directional sector $\theta_S \in [180^\circ, 270^\circ]$. For the wind loading, since it is relatively much less important compared to wave loading, the direction θ_W will also be studied for this directional sector. To check the dependency between θ_S and θ_W , the histogram of the difference ($\theta_S - \theta_W$) within this direction sector is plotted in Figure 5.7. The angular statistical results show a small deviation between θ_S and θ_W . For simplicity, the wind and the waves may

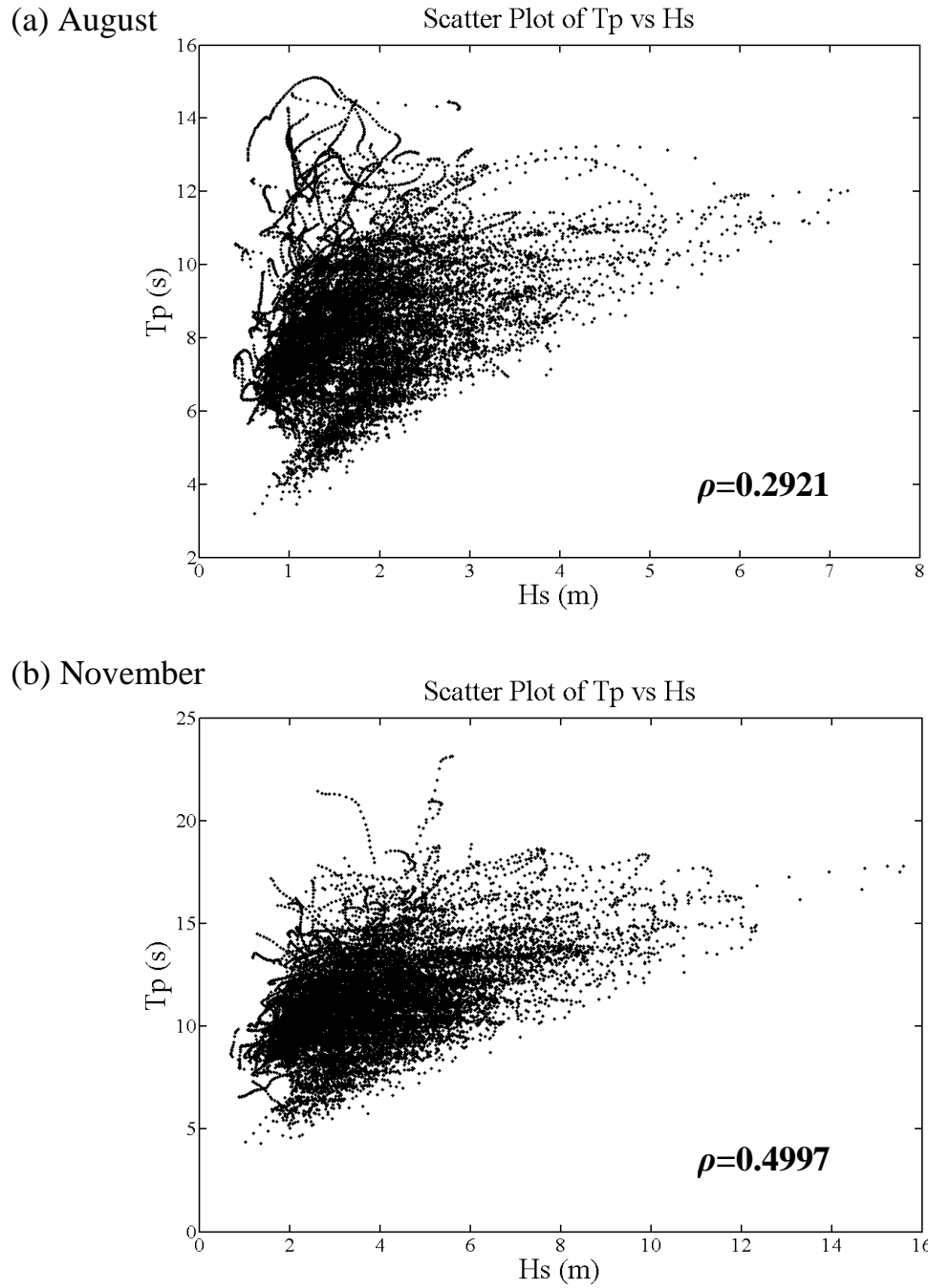
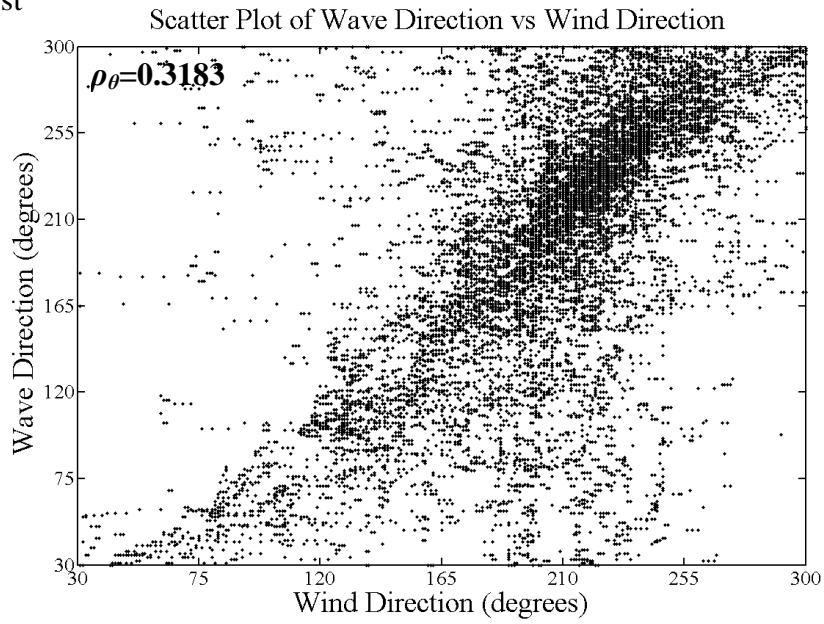


Figure 5.4 Comparison of dependencies between H_s and T_p for two different months (ρ is linear correlation coefficient).

(a) August



(b) November

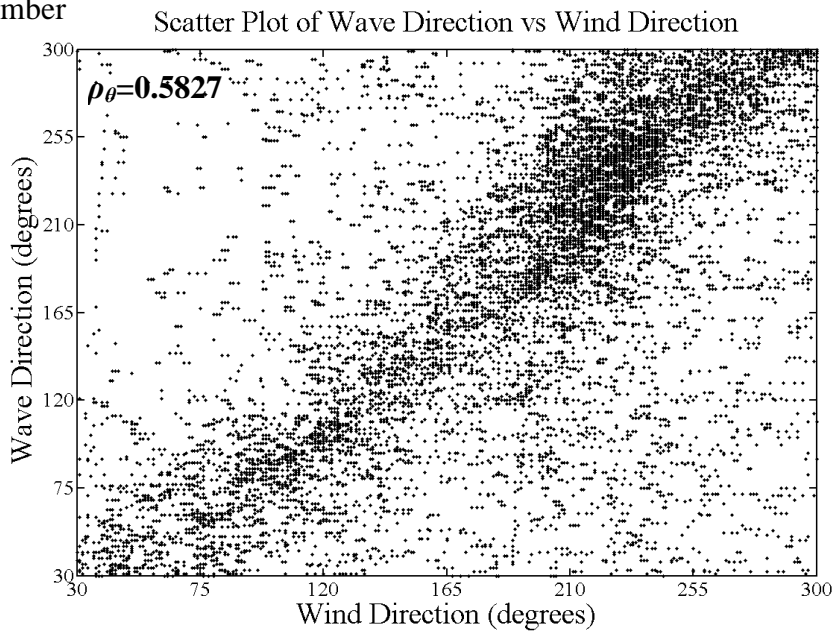


Figure 5.5 Comparison of dependence between θ_S and θ_W over two different months (ρ_θ is angular correlation coefficient).

be assumed to have a coinciding direction given such a small difference in ($\theta_S - \theta_W$) as suggested in the design code (DNV 2007). As a result, the data set ($H_S, T_P, V_W, \theta_S, \theta_W$) covering the period of November to February and with value θ_S lies in the interval $[180^\circ, 270^\circ]$ will be extracted for subsequent statistical analysis.

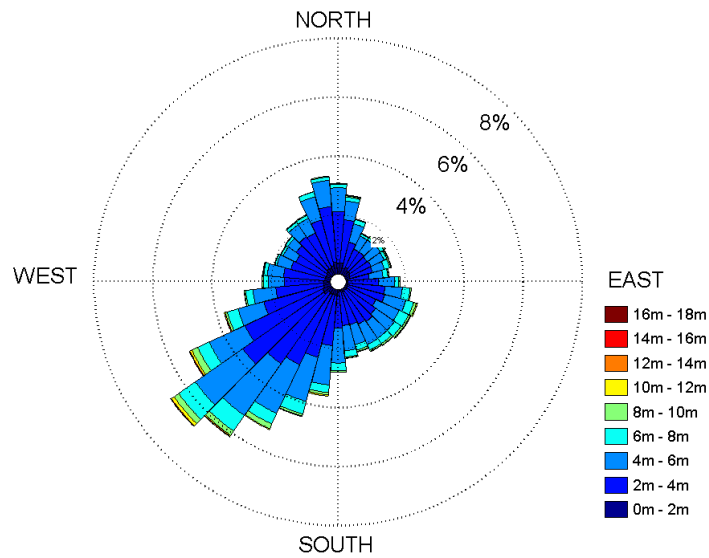


Figure 5.6 Rose plot of wave direction in Nov-Feb.

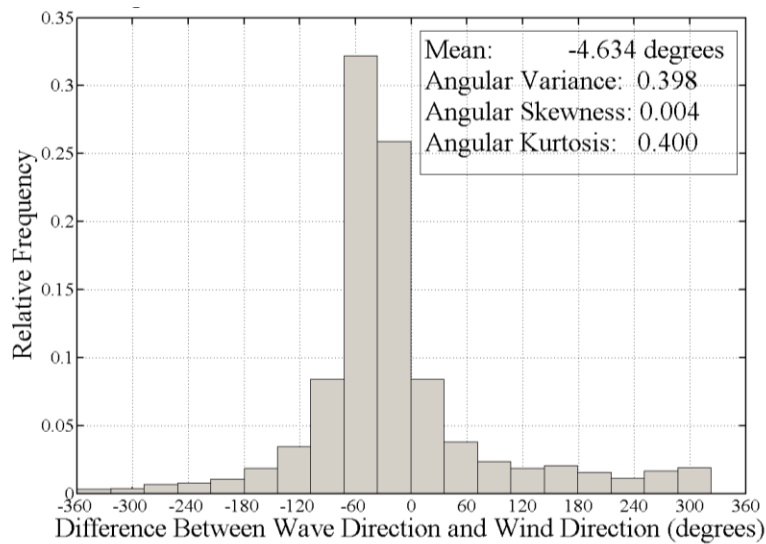


Figure 5.7 Histogram of difference between wind direction and wave direction ($\theta_S - \theta_W$).

5.4.2 Application of Bivariate Models

Based on the extracted data set, the three different bivariate modeling approaches as introduced in Section 5.2 are applied. However, in implementing three statistical models, the steps in estimating the model parameters are different. For the conditional joint distribution model, the maximum likelihood method may not be practical. For example, in constructing the bivariate model (H_S, T_P) , there are 6 parameters in the conditional lognormal distribution model for T_P as given in Eq. (5.1). Together with the marginal distribution parameters of H_S , the likelihood function, formulated by using the conditional probability density function in Eq. (5.1)-(5.3), turns out to be quite complicated as there are too many parameters to estimate. The maximization of the likelihood function in determining the model parameters is not easy requiring an efficient algorithm in view of the numerous parameters to be determined. Moreover, if the available data sample size is large, the computation becomes very cumbersome as the likelihood function value tends to be quite large.

For this reason, a regression analysis is employed as an easy approximation of the parameter values in the conditional joint distribution model. The basic assumption is that, by dividing the domain of H_S into small intervals from the smallest value to the largest value the conditional lognormal parameters of T_P are nearly constant for the “local” H_S within the same interval. The conditional parameters could then be determined by applying the nonlinear fit to the observed data according to Eq. (5.1)-(5.3). Figure 5.8 shows a general nonlinear fit for the lognormal parameters of T_P conditional on H_S . Each point in

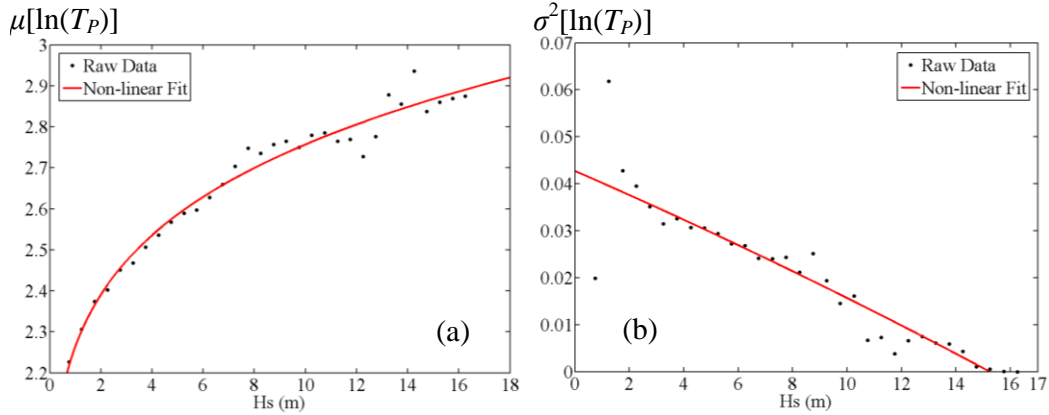


Figure 5.8 Nonlinear fit of (a) $\mu[\ln(T_p)]$ vs H_S and (b) $\sigma^2[\ln(T_p)]$ vs H_S based on Eq. (5.1) and Eq. (5.3) for points representing H_S within interval values of 0.5.

Table 5.3 Regression results for points representing H_S within interval values of 0.5.

Regression fitting equation: $\mu[\ln(T_p)] = a_1 + a_2 \times H_S^{a_3}$	Regression fitting equation: $\sigma^2[\ln(T_p)] = b_1 + b_2 \times \exp(b_3 \times H_S)$
Coefficients (with 95% confidence bounds): $a_1 = 1.319$ (0.250, 2.387) $a_2 = 0.945$ (-0.096, 1.986) $a_3 = 0.183$ (0.027, 0.338)	Coefficients (with 95% confidence bounds): $b_1 = 0.233$ (-1.129, 1.596) $b_2 = -0.191$ (-1.547, 1.166) $b_3 = 0.013$ (-0.071, 0.098)
Goodness of fit: SSE: 0.030 R-square: 0.971 Adjusted R-square: 0.969 RMSE: 0.032	Goodness of fit: SSE: 0.001 R-square: 0.830 Adjusted R-square: 0.819 RMSE: 0.006

Fig 5.8 is the mid-point over an interval of 0.5m, and the points range from 1.0m to 16.5m in the domain of H_S . The results of the fitted values and the associated marginal distribution model of H_S are also presented in Table 5.3. Overall, the deviation from the fitted line is small. Only some “jumps” are observed at the tails

of the curve. One should note that the number of points (every point represent a local “ H_S ” in the divided interval) used in the fit can affect the estimates of the model parameters. This is an issue of trade-off between accuracy in the regression and accuracy in the statistical parameters such as mean and standard deviation associated with each interval. To overcome this requires a reasonable amount of data.

To demonstrate the importance of this point, another regression model, which divides the domain of H_S into intervals of 0.10m ranging from 0.7m to 16.2m, is applied here. The results in terms of the regression parameter estimate are given in Table 5.4 and illustrated in Fig. 5.9. The regression becomes poorer compared with the former case as shown in the goodness-of-fit values. The estimated parameter values $\{a_1, a_2, a_3, b_1, b_2, b_3\}$ deviate quite significantly, especially for $\sigma^2[\ln(T_P)]$. The reason is because the divided intervals of H_S are too small such that the value of $\mu[\ln(T_P)]$ and $\sigma^2[\ln(T_P)]$ calculated over each point contains significant uncertainties due to the limited data set. This causes significant errors in the modeling of the conditional relationship between H_S and T_P which directly affects the quality of the established bivariate model. It is found that by using this set of parameters, when the conditional model is combined with a marginal lognormal distribution model of H_S , the value of the log-likelihood function is decreased from previous case of -3019300 to -3666500. Based on these comparisons, the results obtained in Table 5.3 are adopted in the following multivariate model.

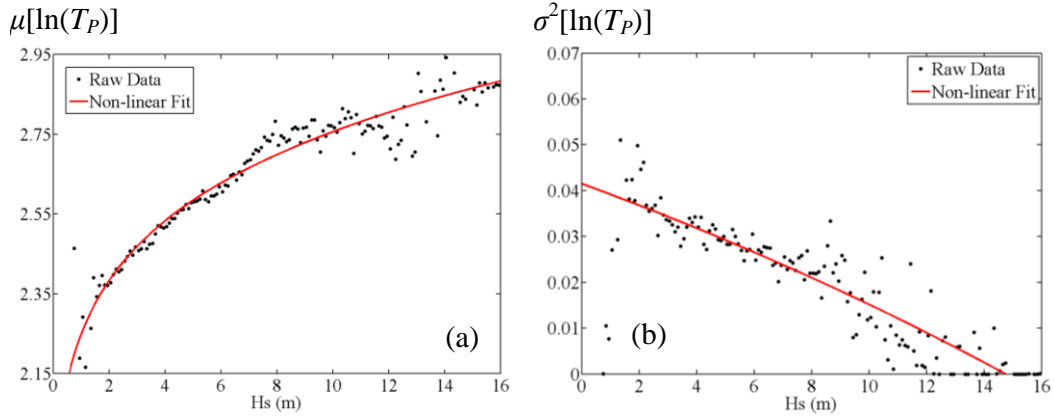


Figure 5.9 Nonlinear fit of (a) $\mu[\ln(T_p)]$ vs H_S and (b) $\sigma^2[\ln(T_p)]$ vs H_S based on Eq. (5.1) and Eq. (5.3) for points representing H_S within interval values of 0.1.

Table 5.4 Regression results for points representing H_S within interval values of 0.1.

Regression fitting equation: $\mu[\ln(T_p)] = a_1 + a_2 \times H_S^{a_3}$	Regression fitting equation: $\sigma^2[\ln(T_p)] = b_1 + b_2 \times \exp(b_3 \times H_S)$
Coefficients (with 95% confidence bounds): $a_1 = 1.039$ (-0.162, 2.240) $a_2 = 1.210$ (0.032, 2.388) $a_3 = 0.152$ (0.033, 0.271)	Coefficients (with 95% confidence bounds): $b_1 = 0.1282$ (-0.086, 0.342) $b_2 = -0.0867$ (-0.297, 0.124) $b_3 = 0.02653$ (-0.026, 0.079)
Goodness of fit: SSE: 0.348 R-square: 0.925 Adjusted R-square: 0.924 RMSE: 0.049	Goodness of fit: SSE: 0.010 R-square: 0.709 Adjusted R-square: 0.705 RMSE: 0.008

The bivariate conditional joint modeling of (H_S, V_w) is approached in quite a similar way. Several former research works related to this can be found in (Dong et al., 2004; Dong, 2007; Dong et al., 2008; Liu et al. 2002). Compared with the bivariate model of (H_S, T_p) , a slight change in the model structure, as

suggested in prior works (Ditlevsen 2002), is made to Eqs. (5.1)-(5.3). The model of V_W for a given H_S is modeled by a 2-parameter Weibull distribution:

$$f_{V_W|H_S}(V_W | H_S) = \frac{k}{\lambda} \left(\frac{V_W}{\lambda} \right)^{k-1} \exp \left[- \left(\frac{V_W}{\lambda} \right)^k \right] \quad (5.22)$$

where the shape parameter k and scale parameter λ are approximated by a nonlinear relationship with the given H_S :

$$k = c_1 + c_2 H_S^{c_3} \quad (5.23)$$

$$\lambda = c_4 + c_5 H_S^{c_6} \quad (5.24)$$

The fitted results to the original data (H_S , V_W) can be illustrated in Fig.

5.10 and shown in Table 5.5.

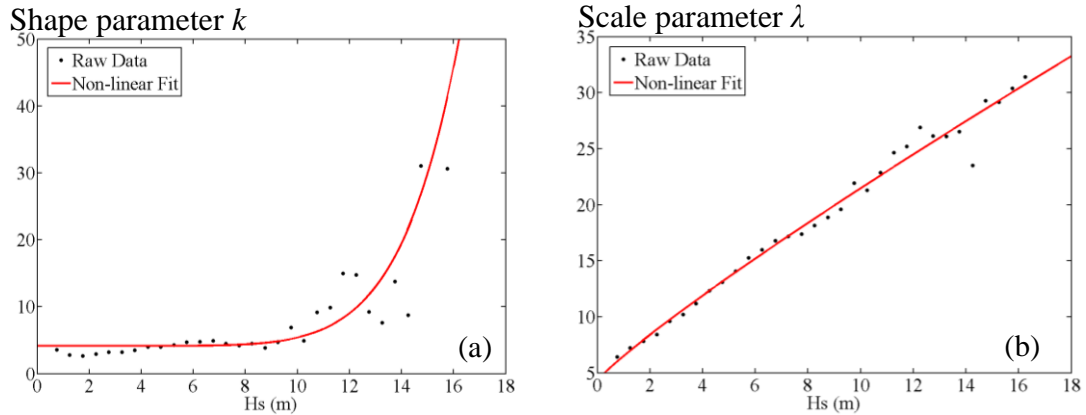


Figure 5.10 Nonlinear fit of (a) k vs H_S and (b) λ vs H_S based on Eqs. (5.23) & (5.24).

Table 5.5 Regression results for k vs H_S and λ vs H_S .

Regression fitting equation: $k = c_1 + c_2 \times H_S^{c_3}$	Regression fitting equation: $\lambda = c_4 + c_5 \times H_S^{c_6}$
Coefficients (with 95% confidence bounds): $c_1 = 4.145$ (1.454, 6.836) $c_2 = 4.364 \times 10^{-8}$ (-3.833×10^{-7} , 4.706×10^{-7}) $c_3 = 7.458$ (3.869, 11.050)	Coefficients (with 95% confidence bounds): $c_4 = 4.284$ (2.109, 6.459) $c_5 = 2.220$ (1.073, 3.367) $c_6 = 0.889$ (0.7234, 1.054)
Goodness of fit: SSE: 883 R-square: 0.7511 Adjusted R-square: 0.733 RMSE: 5.619	Goodness of fit: SSE: 31.4 R-square: 0.982 Adjusted R-square: 0.981 RMSE: 1.041

Compared with the conditional joint distribution model, the procedures of the Nataf and copula approaches require the determination of marginal distributions for the variables. In seeking the best representation of the data, three theoretically available marginal distribution models are adopted for H_S , T_P and V_W . These are Weibull, Gamma and Lognormal distribution models, which are most frequently used models in prior studies (Jaspers 1956, Battjes 1972, Ochi 2011). The model which gives the maximum likelihood function value for the data is considered as the best parametric model. Table 5.6 summarizes the results of each model for H_S , T_P and V_W , according to the maximum likelihood method. A general view of the quality of fit for the one-dimensional marginal density can also be seen in Fig. 5.11. The quality of the selected marginal distribution model can also be observed in the fitting of the tail values as shown in Fig. 5.12. The best models based on the results are Lognormal, Gamma and Weibull for H_S , T_P

and V_W respectively. These are adopted when performing the multivariate fit in both the Nataf and copula approach.

When constructing the copula model, a group of well known families of one parameter copulas, featuring a wide range of dependence, are considered in this study (Nelson 2006). These are the Gumbel, Gaussian, Frank, Clayton, Plackett and Ali-Mikhail-Haq families. The corresponding parameter value in each copula model is estimated by maximizing the pseudo-likelihood function value according to Eq. (5.18). In the Nataf approach, the parameter is evaluated using numerical integration according to Eq. (5.6). The fitted results for both (H_S , T_P) and (H_S , V_W) are given in Table 5.7.

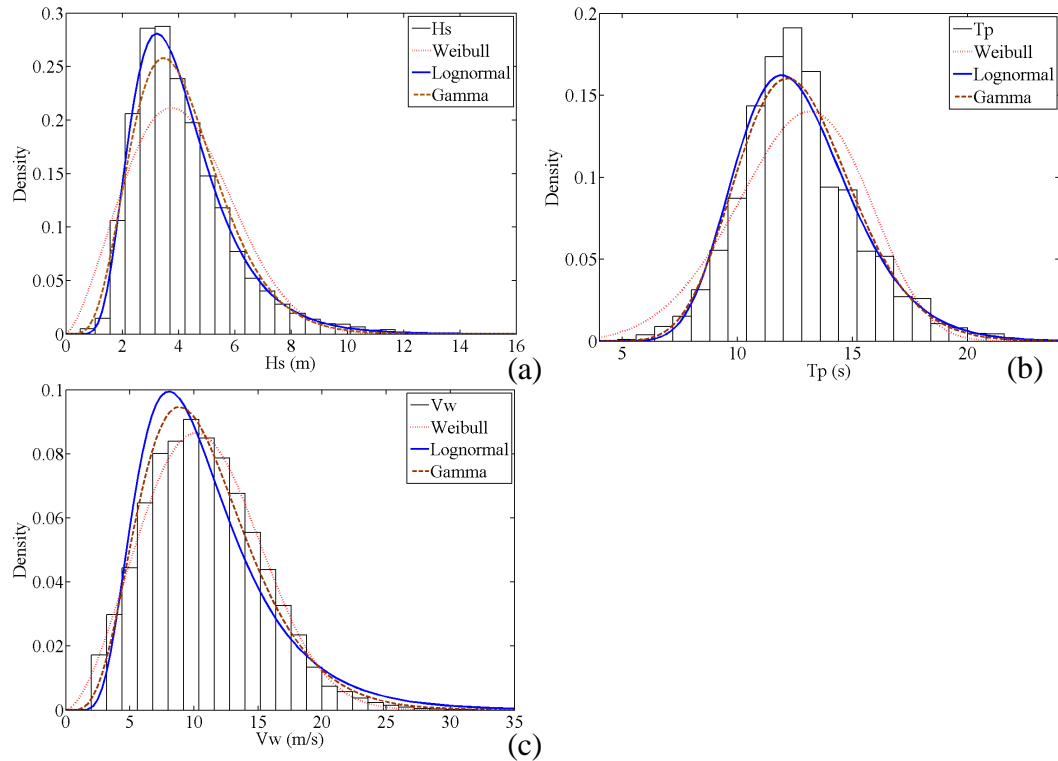


Figure 5.11 Marginal parametric model fit for (a) H_S , (b) T_P , and (c) V_W .

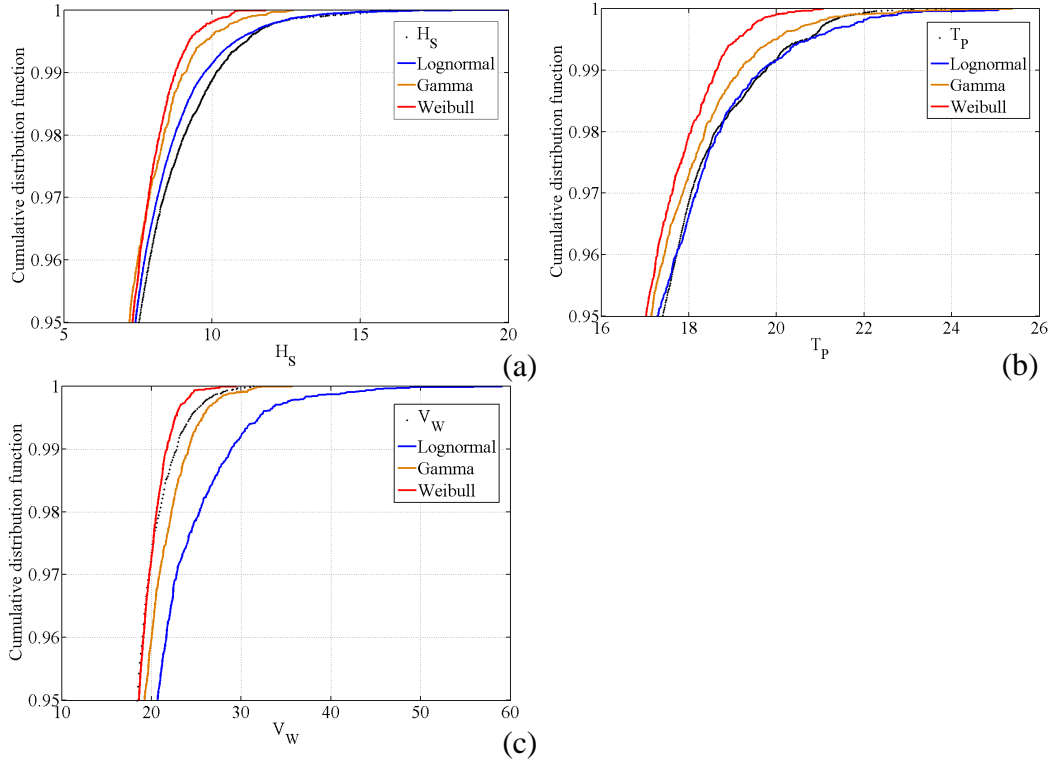


Figure 5.12 Tail fittings of marginal parametric model for (a) H_S , (b) T_P , and (c) V_W .

Table 5.6 Results of marginal distribution model parameter estimates (with estimated standard error in bracket).

	Significant wave height (H_S)	Peak period (T_P)	Wind velocity (V_W)
Lognormal Model PDF: $f(x \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$	$\mu=1.336$ (0.002) $\sigma=0.405$ (0.002) Loglikelihood = -59897 *	$\mu=2.519$ (0.001) $\sigma=0.202$ (0.001) Loglikelihood = -75650	$\mu=2.288$ (0.003) $\sigma=0.450$ (0.002) Loglikelihood = -94095
Gamma Model PDF: $f(x a, b) = \frac{1}{b^a \Gamma(a)} x^{a-1} \exp\left(-\frac{x}{b}\right)$	$a=6.192$ (0.047) $b=0.668$ (0.005) Loglikelihood = -60502	$a=25.069$ (0.196) $b=0.505$ (0.004) Loglikelihood = -75501 *	$a=5.590$ (0.043) $b=1.930$ (0.015) Loglikelihood = -93023
Weibull Model PDF: $f(x k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^k\right]$	$k=2.428$ (0.010) $\lambda=4.666$ (0.011) Loglikelihood = -62907	$k=5.121$ (0.020) $\lambda=13.715$ (0.016) Loglikelihood = -77447	$k=2.632$ (0.011) $\lambda=12.166$ (0.027) Loglikelihood = -93017 *

Notes: * maximum loglikelihood value indicates the best model.

In testing the quality of the three potential bivariate models, which consist different formulations and number of parameters, the standard Akaike Information Criteria (AIC) is adopted, given by (Akaike 1973):

$$AIC = -2l(p) + 2p \quad (5.25)$$

where p is the number of parameters used in each model, and $l(p)$ is the maximum log-likelihood resulting from each model. The AIC takes into account the simplicity of the model and the goodness-of-fit, and a smaller AIC value indicates a better model.

5.4.3 Results and Discussions

The parameter estimates, AIC statistics and log-likelihood function values for each of the models are presented in Table 5.7. The results show that the conditional joint distribution approach gives an undesirable model, as it gives the largest AIC value in both (H_S, T_P) and (H_S, V_W) . The basic reason is that the uncertainty contained in the conditional joint distribution model does not only cover the uncertainty in the parameter estimates of the marginal distribution (for example, μ and σ for the marginal lognormal distribution of H_S) but also the uncertainty in determining the values of $\{a_1, a_2, a_3, b_1, b_2, b_3\}$ through regression analyses, see Table 5.7. Clearly, the statistical uncertainty is much larger compared to the other two approaches as could be observed in the comparison of the estimated standard errors with the associated parameters, see Table 5.7.

Another major reason is that the form of the function describing the relationship between the variables is fixed in the conditional joint distribution approach.

Table 5.7 Comparison of parameter estimates and goodness-of-fit to the data for the 3 approaches (with estimated standard error in bracket).

		Parameter Estimate	Total Log-likelihood	No. of Parameters	AIC
(H_s, T_p)	<i>Copula families:</i>				
	Gumbel	$\theta=1.396$ (0.0062)	-131741	5	263493
	Gaussian	$\theta= 0.467$ (0.004)	-131428	5	262867*
	Frank	$\theta= 3.076$ (0.039)	-131655	5	263320
	Clayton	$\theta= 0.537$ (0.008)	-132949	5	265907
	Plackett	$\theta= 4.070$ (0.067)	-131666	5	263342
	Ali-Mikhail-Haq	$\theta= 0.834$ (0.006)	-132265	5	264541
	<i>Conditional joint distribution model:</i>	$a_1= 1.039$ (0.613) $a_2= 1.210$ (0.601) $a_3= 0.152$ (0.061) $b_1= 0.128$ (0.109) $b_2=-0.087$ (0.107) $b_3= 0.027$ (0.027)	-3019300	8	6038616
	<i>Nataf model:</i>	$\rho_\theta = 0.487$	-131442	5	262895
	(H_s, V_w)	<i>Copula families:</i>			
Gumbel		$\theta= 1.809$ (0.010)	-143340	5	286690*
Gaussian		$\theta= 0.622$ (0.003)	-145034	5	290079
Frank		$\theta= 4.809$ (0.042)	-145226	5	290463
Clayton		$\theta= 0.780$ (0.007)	-148841	5	297692
Plackett		$\theta= 8.367$ (0.132)	-144818	5	289647
Ali-Mikhail-Haq		$\theta= 0.970$ (N.A.)	-148180	5	296371
<i>Conditional joint distribution model:</i>		$c_1= 4.145$ (1.373) $c_2= 4.364 \times 10^{-8}$ (2.178×10^{-7}) $c_3=7.458$ (1.832) $c_4=4.284$ (1.110) $c_5= 2.221$ (0.585) $c_6=0.889$ (0.084)	-2061200	8	4122416
<i>Nataf model:</i>		$\rho_\theta = 0.687$	-145354	5	290719

Notes: the lowest AIC value indicates the best model.

On the other hand, both the Nataf and copula approaches show their suitability in fitting the bivariate data and are even easier to apply as fewer parameters need to be estimated, incurring less uncertainty associated with the estimated parameters. The copula approach shows that it is less restrictive by accommodating various structure of dependence, where the Gaussian copula family gives the best fit to (H_S, T_P) , while the Gumbel family shows its best performance for (H_S, V_W) . The Nataf model implicitly assumes the Gaussian family copula which is one form of dependence structure in the copula approach. In this sense, the copula approach can give more desirable results compared to Nataf approach. For example, the Gumbel copula family characterizes data that have more correlations at the two extremes of the dependent distributions but has its highest correlation in the maximas. As mentioned earlier, the ocean storm could induce large wave height H_S and wind speed V_W . Hence, a higher dependency between H_S and V_W could be observed for large values in (H_S, V_W) . However, there are also limitations in some families of copula in modeling data. For instance, the Ali-Mikhail-Haq family can only model a weak dependency structure in bivariate data for which has a correlation coefficient covers a value range from $-2/9$ to $2/9$. This largely reduces the flexibility in the copula approach and may give difficulties in the parameter estimate since the maximum likelihood method may not converge easily.

A general illustration of the suitability in each approach can be seen from the probably contour plots comparing the fitted models against the empirical data for (H_S, T_P) in Fig. 5.12. The near-perfect agreement in the contour line in the

copula model supports its suitability in fitting the available data. Similar results are also obtained for (H_S, V_W) in Fig. 5.13. However, there still exists weakness in all the 3 approaches. For example, a large value of H_S is unlikely to be accompanied with a small T_P because of breaking wave limit. This implicit physical phenomenon is not easy to incorporate in the joint statistical model in a natural manner, although Repko (2004) has shown that the model can be improved by introducing the wave steepness parameter. In real practical applications, like the long term dynamic assessment of structures, the physical aspects of the environmental parameters should be paid attention to.

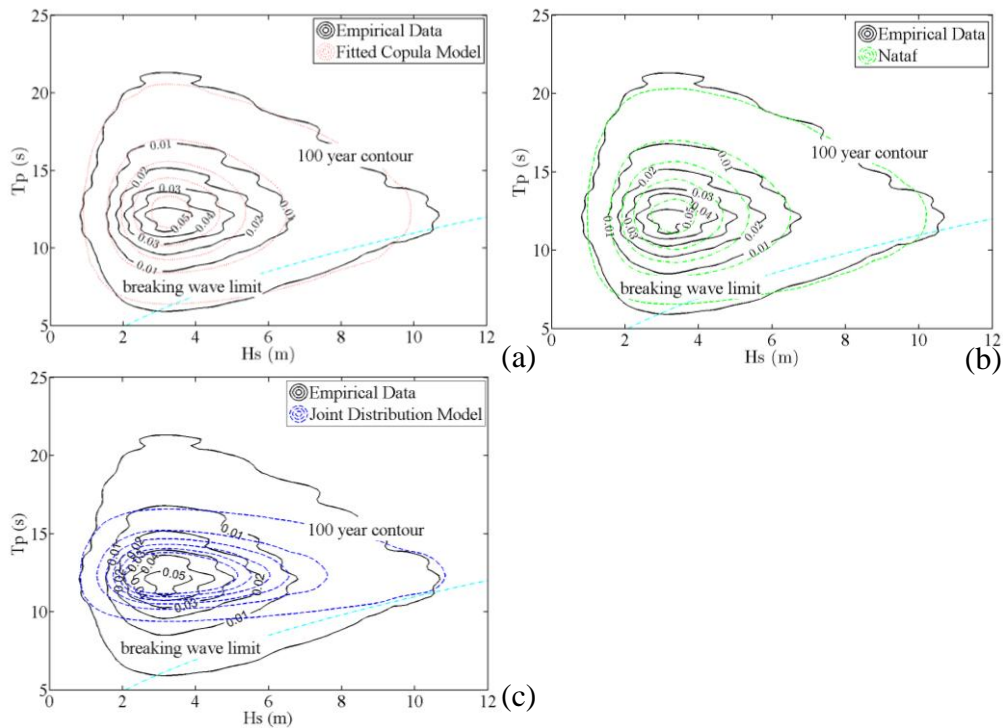


Figure 5.13 Comparison of contour plot between original data and (a) copula approach, (b) Nataf model, (c) conditional joint model for H_S and T_P .

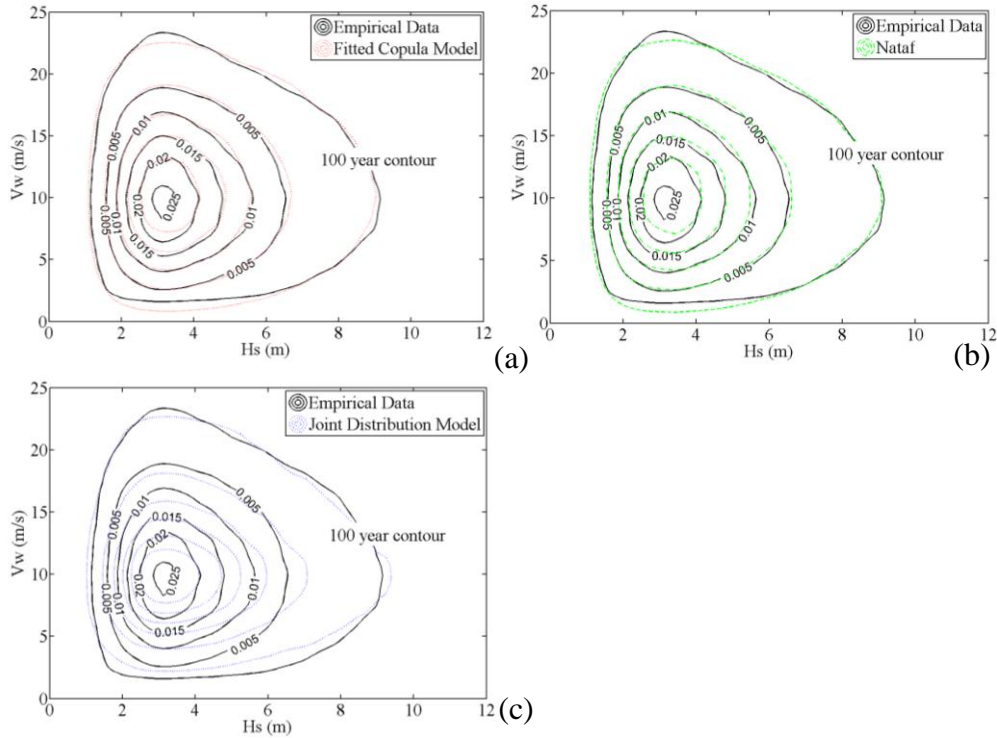


Figure 5.14 Comparison of contour plot between the original data and (a) copula approach (b) Nataf model (c) conditional joint model for H_s and V_w .

The results presented here could also help to explain the difference between the Gaussian copula and Nataf approaches. As aforementioned, although the Nataf assumes a Gaussian copula to describe the dependency structure (Lebrun and Dutfoy 2009), there are several points of deviation in the two approaches. First, the parameter estimation from Nataf model is different from Gaussian copula. The Nataf model estimates the parameter in the Gaussian function, as shown in Eq. (5.6), by preserving the value of correlation coefficient in the original data. This is different from the Gaussian copula applied here where the parameter estimate is based on maximizing the likelihood. The use of correlation coefficient in representing the dependencies between the variables may not be accurate in some cases. The value of ρ could be quite sensitive to

extraordinary large values in the data sample. For example, if there is a presence of outliers in the data, the value of ρ could be greatly affected (Li et al. 2012). In addition, the correlation coefficient ρ is only a measure of linear dependence between random variables (Joe 1997). The initial assumption of linear relationship may not be applicable as the dependency between the variables may not be perfectly linear. Especially in the copula model, there is a transformation of variables from marginal distribution to standard uniform distribution. The dependency could be quite different between the transformed and non-transformed variables. Another minor difference is the uncertainty with the associated estimated parameter. In Gaussian copula, the standard error is estimated based on Fisher's information matrix which is determined from the maximum likelihood function. But the parameter estimated in Nataf is from an integration function which may not give a standard error directly.

5.5 Time Domain Structural Analysis

There are two approaches to perform dynamic structural analysis, namely through analytical solution, or numerical time marching computation.

The analytical solution approach (such as through frequency domain analysis) is only applicable for very simple offshore structures. Complexities such as the effect of nonlinearities in the loading and/or structural elements, and their interactions, are difficult to handle or interpret. For example, the nonlinear soil-spudcan interaction and inter-element joint connections may result in large

amplification of the transfer function in the upper and lower frequency range (Lu et al. 2002). Although it is still possible to perform the analysis in the frequency domain, like using a linearized transfer function, the accuracy of the results may need to be further investigated, especially for a non-Gaussian response.

The alternative numerical time marching approach is only feasible to perform over a limited time span in practice. The results from the short duration analysis can only at best be projected to long duration (such as the design life) response statistics through the asymptotic theory of statistical extremes, assuming certain conditions are fulfilled as discussed in Chapter 2. However, complexities including nonlinear effects can be incorporated, the limitation being availability of fast computing power.

In this work, the time marching approach is adopted in the calculation of the long term load for the example structure. The present study considers the wave loading as the major loading on the offshore structure. Therefore, the emphasis in the current work is given to the analysis of structural response for considering the joint model of (H_S, T_P) . Before the dynamic analysis is to be conducted, one important issue is to ensure an accurate simulation of the wave stochastic process (e.g. the random wave simulation) which is required in the calculation of the conditional response distribution $Q(X(t)|\theta)$ in Eq. (2.12). Realizing this, the fundamental knowledge regarding the basic linear wave theory and wave characterizations is given in Appendix G.

5.5.1 Proposed Discretized Copula Approach

While the joint distribution provides the basic information for characterizing the environmental load, its implementation into the structural analysis still needs further work. As discussed in Chapter 2, the direct integration in Eq. (2.12) is cumbersome and impractical. To achieve an efficient way of calculating it, several approaches have been investigated in the literature (Baarholm 2010, Cheng et al. 2003). Most works followed two commonly used approaches:

- Sea state block method
- Environmental contour method

In the environmental contour approach, a contour line in the parameter space is found from an inverse first order reliability method (IFORM) to represent an extreme sea state condition. The estimate for the long term load is then obtained by searching along the contour for the parameters giving the maximum characteristic extreme response. The advantage of applying this concept is that the calculations are largely reduced (e.g. the structural analysis is only performed along the contour line) compared to the direct integration method (e.g. the computation will cover the whole domain instead of only the contour). But the predictions through the IFORM are only based on limited consideration of the variability of the given environmental parameters, the contour line, and can be inaccurate and unconservative.

For this reason, the sea state block method which utilizes the property of scatter diagram is preferred. As the number of bins in a scatter diagram is

typically excessive, the sea state block approach is to condense the data into a smaller quantity of representative blocks that are manageable. The structural analysis in each of these blocks is then performed to determine an overall long term load based on the integration according to Eq. (2.2). This does not need a numerical integration as required in the direct integration and considers the variability of the environmental parameters over its domain (e.g. the evaluation of the response distribution $Q(.)$ conditional on the given environmental state).

However, the sea state block method is still quite tedious as the conditional response distribution in Eq. (2.12) needs structural analysis to be repeatedly conducted in each of the blocks in the scatter diagram to consider all the possible environmental conditions. For some of the sea state blocks, they may have important contributions to the integration in Eq. (2.2). Inaccurate calculations in these blocks may lead to large errors in the estimated long term value as it may be quite sensitive to these blocks. This is particular important for floating structures which may be quite sensitive to some specific sea state parameter, for example, the wave period (Kawano & Venkataramana 1999). An alternative way of improving this approach is to increase the number of blocks in order to reduce the importance of the critical blocks. But it becomes impractical to perform computationally expensive simulations for all these created blocks since the number of blocks may be too large. As a result, it is thus of interest to explore another efficient alternative approach in handling this problem.

In this work, the analysis incorporates the copula model in describing the environmental parameters. This implies that in determining the long term load,

the calculations in Eq. (2.2) needs to be modified slightly. Instead of the direct integration in the parameter space, the integration domain would be changed to the copula domain in the computation which is expressed as:

$$\begin{aligned}
P_E &= \int_{\theta_1, \dots, \theta_n} Q(X(t) > l | \theta_1, \dots, \theta_n) f(\theta_1, \dots, \theta_n) d\theta_1, \dots, d\theta_n \\
&\text{for } \theta_1, \dots, \theta_n \in (-\infty, +\infty) \\
&= \int_{F_1(\theta_1), \dots, F_n(\theta_n)} Q(X(t) > l | \theta_1, \dots, \theta_n) c(F_1(\theta_1), \dots, F_n(\theta_n)) dC \\
&\text{for } F_1(\theta_1), \dots, F_n(\theta_n) \in [0, 1]
\end{aligned} \tag{5.26}$$

where $F_i(.)$ is the marginal distribution functions of the i th environmental parameter θ_i , and $c(.)$ is the copula density function. Since the computation is taken over the range of the cumulative distribution functions $F_i(\theta_i)$ instead of the original parameter space, the integration becomes much easier as the domain is reduced from $[0, +\infty)^n$ to $[0, 1]^n$. However, to obtain the conditional extreme response distribution $Q(X(t)|\theta_1, \dots, \theta_n)$ still demands large amount of calculations and simulations. This is also the key drawback of the sea state block approach which cannot identify the critical sea state conditions from the parameter space, and thus leads to unnecessary computations for the unimportant environmental condition parameters. To overcome this major difficulty, a concept of discretized subcopula based method is proposed in this investigation.

The concept of a subcopula function C' is a subset of the copula function C where the domain of each marginal variable is bounded. By satisfying the basic properties of copula, the subcopula also follows the mapping of marginal variables to the joint function value:

$$C': S_1 \times \dots \times S_n \rightarrow [0, 1] \quad \text{and} \quad H(x_1, \dots, x_n) = C'(F_1(x_1), \dots, F_n(x_n)) \quad (5.27)$$

where S_i are subsets of $[0, 1]$ which is a subdomain of the original variable. Therefore, a copula could be discretized into many subcopulas that have the same function value mapping as shown in the above equation. By adopting this concept, the computation of the integration equation in Eq. (5.26) can be approximated by a summation equation which transforms the copula into discretized subcopulas:

$$\begin{aligned} P_E &= \int_{F_1(\theta_1), \dots, F_n(\theta_n)} Q(X(t) > l | \theta_1, \dots, \theta_n) c(F_1(\theta_1), \dots, F_n(\theta_n)) dC \\ &\approx \sum_{i=1}^k \bar{Q}_i(X(t) > l | \theta_1^i, \dots, \theta_n^i) \Delta C_i(F_1(\theta_1^i), \dots, F_n(\theta_n^i)) \end{aligned} \quad (5.28)$$

where $\Delta P_E^i = \bar{Q}_i(X(t) > l | \theta_1^i, \dots, \theta_n^i) \Delta C_i(F_1(\theta_1^i), \dots, F_n(\theta_n^i))$ is the exceedance probability in each subcopula such that $P_E = \sum \Delta P_E^i$, $\bar{Q}_i(\cdot)$ represents the conditional response distribution for the environmental parameter values in the i th subcopula which could be determined by assuming an average environmental condition within the subcopula. A simple way is to use the environmental parameter values at the center point in the discretized subcopula for the evaluation of the structural analysis which is similar with the sea state block method. ΔC_i is the probability of the i th subcopula which can be deduced from the copula function by using the i th order difference. For example, for a bivariate copula, ΔC for a subcopula lying in a subdomain $[x_1, x_2] \times [y_1, y_2]$ can be calculated by using the copula function $C(\cdot)$ as follow:

$$\Delta C_{[x_1, x_2] \times [y_1, y_2]} = \Delta_{y_1}^{y_2} \Delta_{x_1}^{x_2} C(x, y) = C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \quad (5.29)$$

Therefore, with these applied procedures, one could process the calculations in the integrations with much easier steps. The simulation and structural analysis required will depend on the number of subcopulas needed in the integration of the copula domain. Thus, in order to improve the efficiency in the calculation, one possible way is to use less subcopulas for the integration in Eq. (5.28). To achieve this, a procedure which only discretizes the “important” subcopulas that contribute most to the integration in Eq. (5.28) is proposed in the computation. For example, the exceedance probability ΔP_E^i within each subcopula is an indicator of its contribution to the overall exceedance probability. The accuracy of the determined long term value is highly depending on the calculations within the subcopulas that have large ΔP_E^i . However, as shown in Eq. (5.28), the approximation of ΔP_E^i within these subcopulas uses an average environmental condition which may not be accurate enough. It is therefore necessary to split these subcopulas into smaller subcopulas to account for the changes in the environmental parameters within the subcopula. This makes the integration relying on finer subdomains and helps to obtain more reliable results. Adopting this concept and using the above equations, the general algorithm in calculating the long term load can be described as follows:

- I. Make an initial discretization of the whole copula domain and estimate the long term load according to Eq. (5.28). Based on the estimated load value l , calculate the individual exceedance probability ΔP_E^i for each of the subcopula as:

$$\Delta P_E^i = \bar{Q}_i \left(X(t) > l \mid \theta_1^i, \dots, \theta_n^i \right) \Delta C_i \left(F_1(\theta_1^i), \dots, F_n(\theta_n^i) \right) \quad (5.30)$$

where $i = 1, \dots, k$ denotes the numbering of the discretized subcopulas.

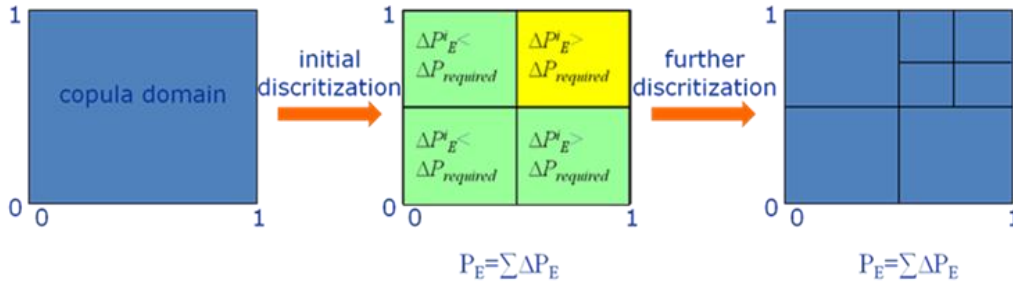


Figure 5.15 Schematic showing of proposed discretization procedure.

- II. For the subcopula exceedance probability that is higher than a required level $\Delta P_{required}$ (e.g. $\Delta P_{E}^i > \Delta P_{required}$), the subcopula will be further discretized into finer subcopulas by dividing each subcopula dimensions' domain into two equal subdomains. Based on these new discretized subcopulas, calculate the long term load and ΔP_{E}^i for all the created subcopulas.
- III. The discretization procedure in step (II) is repeated until all the ΔP_{E}^i calculated for the subcopula is smaller than the required $\Delta P_{required}$. Then the long term load can be calculated by using the discretized subcopulas based on Eq. (5.28).

To demonstrate the proposed method, an offshore structure is assessed for its long term performance in the next section.

5.5.2 Structural Analysis of a Fixed Offshore Platform

5.5.2.1 Example

To demonstrate the performance of the proposed approach, the fixed jacket structure introduced in Chapter 4 is used. The base shear corresponding to a long

term (be specific how long) exceedance probability P_E of 0.01 will be investigated, considering only the wave and current load. The two-dimensional Gaussian copula of (H_S, T_P) which has been constructed in Section 5.4 is applied to model the irregular random waves. The JONSWAP spectrum is utilized for characterizing the sea state condition for a given wave condition of (H_S, T_P) . For simplicity, the current is assumed to have a constant velocity of 1.5m/s along with the wave in the same longitudinal direction. The simplest linear random wave model which has been discussed in Appendix G is used to model the random waves in this work. The random waves are first simulated and then converted to sea load in the dynamic analysis for the fixed structure through USFOS.

5.5.2.2 Simulation requirements

In directly applying the time domain analysis, the application of the standard simulation of stochastic process in representing a random wave condition in the offshore structural dynamic analysis requires several considerations. Foremost is that every sea state or condition can be described by a specific energy spectrum, such as the *JONSWAP* wave spectrum which is defined for specific value of (H_S, T_P) . The energy dispersion and centralization are quite dependent on these values, as illustrated in Fig. 5.16. The stochastic random waves to be generated must be sufficiently representative to capture the probabilistic characteristics of the corresponding sea state, from a calm to severe environment condition.

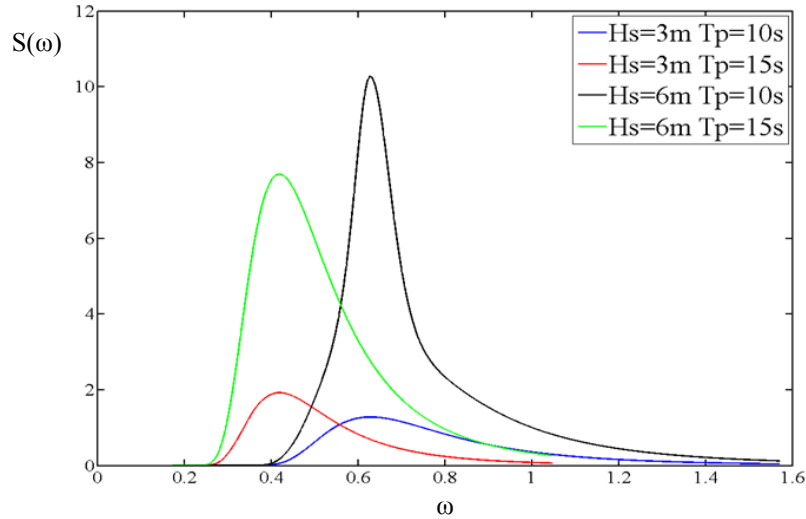


Figure 5.16 Illustration of *JONSWAP* spectrum for different sea states.

Due to the change of parameter values in the spectrum, the statistical properties of the time series and the validity of the simulation scheme must be ascertained. This requires several changes in the simulation procedures for the method presented in the previous section. Specifically, the following factors should be considered (Yamazaki and Shinozuka 1988):

- Cut-off frequency ω_{cut} : the maximum frequency beyond which the spectral density function $S(\omega)$ may be assumed to be zero for either mathematical or physical reasons.
- Number of simulated frequencies N : the least number of simulated waves in representing a stochastic process which has a Gaussian characteristic.
- Length of simulated time T : the minimum simulation time needed in establishing an accurate probabilistic model for the corresponding wave spectrum.

- Time step Δt_{time_step} : a value used in separating the generated value from Eq. (G.10) in order to avoid aliasing according to the sampling theorem (Bracewell 1986).

Depending on the situations and accuracy requirement, the above parameters should take different values (Shinozuka 1987). As can be seen in Fig. 5.12, the sea state parameters H_S and T_P as provided in the established statistical model covers a wide range of values. This implies that the constructed wave spectra would be quite different amongst one other, especially those with (H_S, T_P) values far apart.

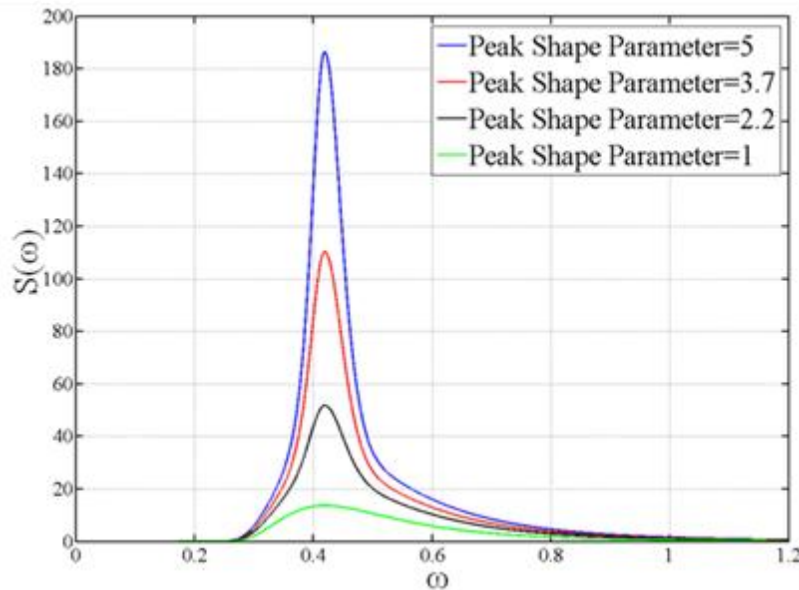


Figure 5.17 Illustration of *JONSWAP* spectrum for different peak shape parameters.

In the simulation of the stationary process for the water elevation $\eta(x, t)$, the cut off frequency ω_{cut} denotes the largest frequency that is used in the random wave representation. The value of ω_{cut} should not be too small as some structures

are quite sensitive to high frequency loadings. A low value of ω_{cut} may also not represent the bulk of the wave energy for a sea state (as given by the area under the spectrum function). The following criterion is usually used to estimate the value of ω_{cut} (Deodatis and Shinozuka 1989):

$$\int_0^{\omega_{cut}} S(\omega) d\omega = (1 - \varepsilon) \int_0^{\infty} S(\omega) d\omega \quad (5.31)$$

where $\varepsilon \ll 1$ (e.g. 0.01 or 0.001) depending on the accuracy required. The chosen ω_{cut} is highly dependent on the value of T_p which generally shows where the energy (the spectral value) is concentrated. Another parameter that could govern the decision in choosing the ω_{cut} is the peak shape parameter γ according to Eq. (G.9). A lower value of γ would mean a large spread of energy in the frequency domain, see Fig. 5.17.

A proper value of ω_{cut} could be determined by using the peak shape parameter and the required ε from the ratio of ω_{cut}/ω_p ($\omega_p = 2\pi/T_p$), see Fig. 5.18. The figure shows that for the same accuracy (ε) in the simulated process, a smaller peak shape parameter would require a larger ω_{cut} . As the peak shape parameter has a value only between 1 and 5, Fig. 5.18 is sufficient for estimating the minimum ω_{cut} needed for simulating different sea state conditions. A value of 4 for the ratio ω_{cut}/ω_p may be adequate for most simulations (Deodatis and Shinozuka 1989) and adopted for the subsequent analysis to achieve an accuracy of $\varepsilon < 0.01$.

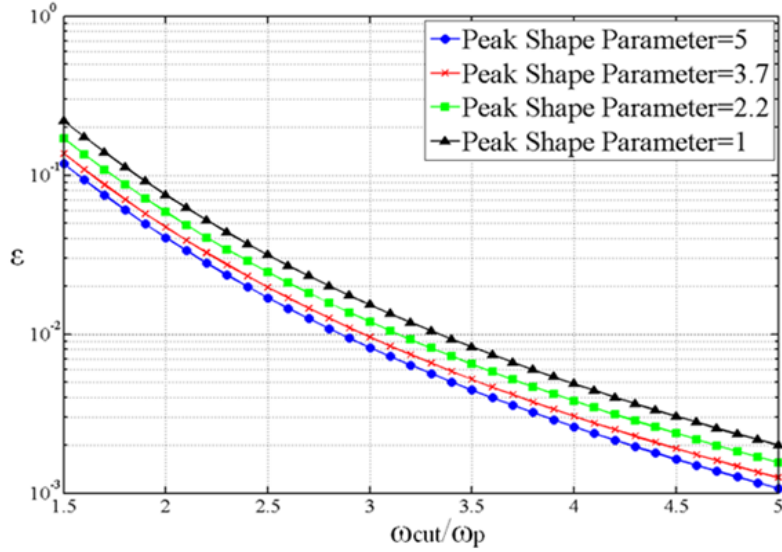


Figure 5.18 Cut off frequency ω_{cut} for different peak shape parameters.

According to Eq. (G.10), when ω_{cut} is fixed, the larger the number of simulated waves N , or equivalently the smaller the $\Delta\omega$, the longer will be the period of the simulated process. The simulated process from Eq. (G.10) would have a period T_O :

$$T_O = \frac{2\pi}{\Delta\omega} \quad (5.32)$$

As such, the value of N , or $\Delta\omega$, is determined according to the required longest period. Utilizing the relationship $N\Delta\omega = \omega_{cut}$, the above equation could be directly written into a function of T_O and ω_{cut} for N :

$$N = \frac{T_O \omega_{cut}}{2\pi} \quad (5.33)$$

In simulating the irregular random waves, since the collected wave data is hourly records, the simulated period $T_0 = 1$ hour is adopted.

According to the sampling theorem to avoid aliasing (Bracewell 1986):

$$\Delta t_{time_step} \leq \frac{2\pi}{2\omega_{cut}} \quad (5.34)$$

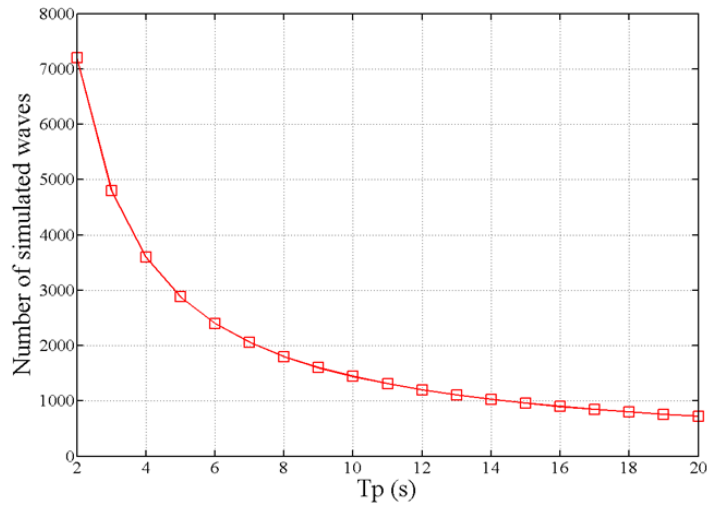


Figure 5.19 Determination of number of frequencies N for a given peak period.

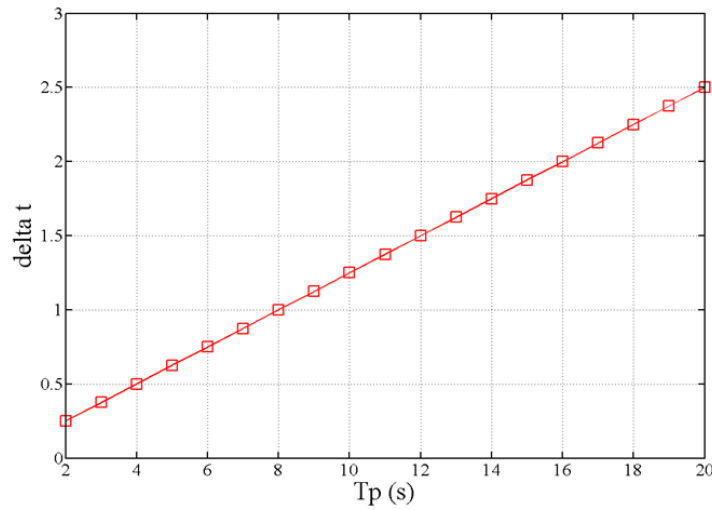


Figure 5.20 Determination of time interval Δt_{time_step} for a given peak period.

Therefore, by using the general value 4 for ω_{cut}/ω_p and Eq. (5.32)-(5.34), the number of frequencies and the required time interval can be determined from the plotted curves as shown in Figs. 5.19 and 5.20.

The accuracy inherent from the above relationships is investigated. Four combinations of (H_S, T_P) in Table 5.8 are chosen, where the H_S and T_P values correspond to the 95 percentile and 5 percentile from its empirical quantile function. The tested critical sea states include a short period with: large wave height and wave period; large wave height and small wave period; small wave height and large wave period; small wave height and small wave period. Other sea states would mostly have the values in between these four cases. The values of the parameters based on the relationships to be adopted for the simulations are provided in the table. General illustrations of the simulated wave process, according to Eq. (G.10), are presented in Fig. 5.21.

The graphs reflect the characteristics of different wave states where spectrum centralizes in the high frequency regions (small T_P) gives fast oscillations in the elevation and large wave spectrum energy (large H_S) would lead to big amplitude in the wave height. The recorded simulated time series data are compared with a Gaussian distribution (e.g. the simulated stochastic process should have a zero mean and standard deviation equals to $H_S/4$) through a two sample K-S test. A large p-value in the Gaussian test has also proved the correctness of the entire process for each of the case. Hence, the above procedures are adequate enough to be applied for all the structural analysis.

The generalized extreme value model is used to fit the maximum response base shear $\bar{Q}(\cdot)$ from the simulation. Ten stochastic processes are simulated and used as the realizations for the construction of each conditional response distribution model $\bar{Q}(\cdot)$.

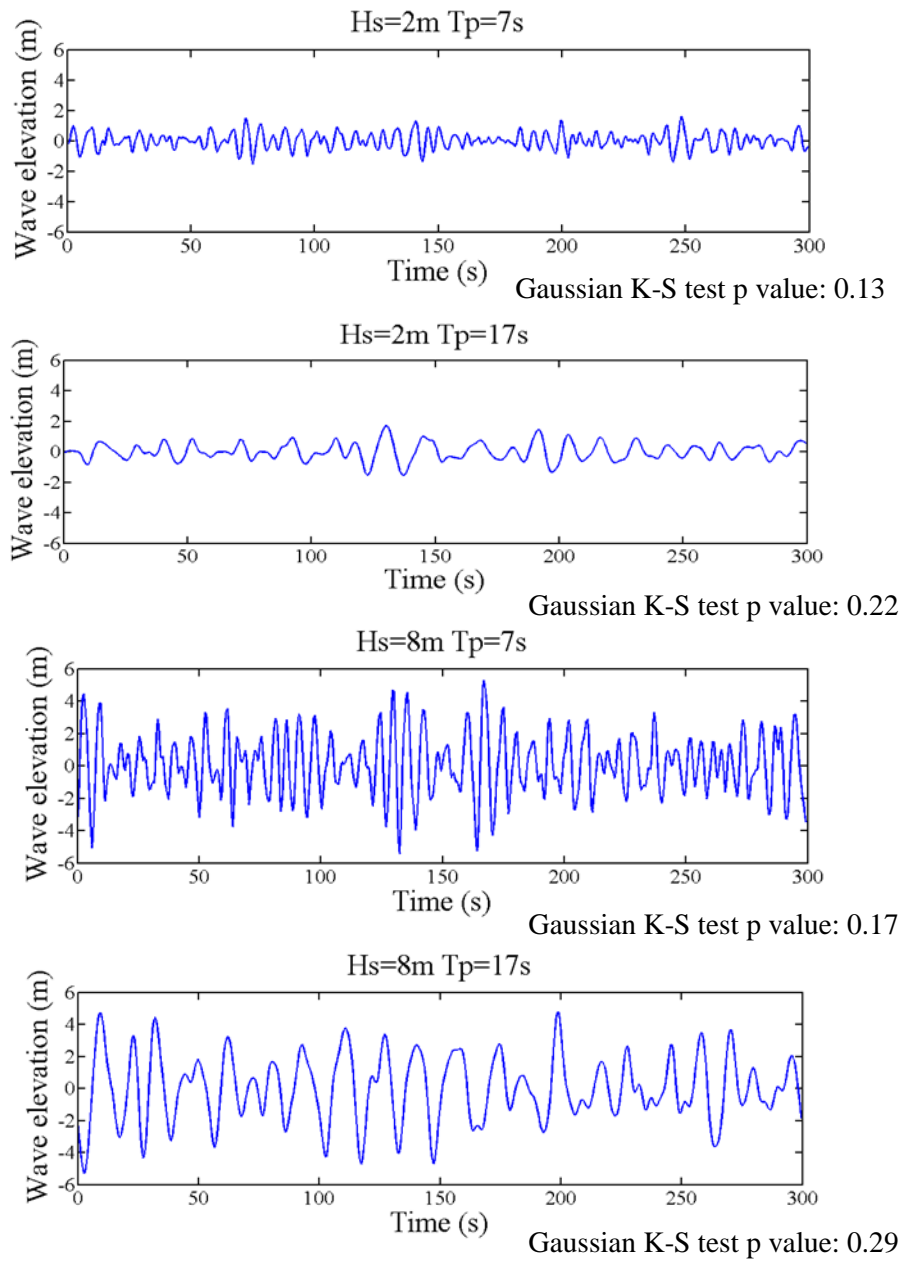


Figure 5.21 Typical simulated wave elevations.

Table 5.8 Sea states and values of simulation parameters for the same accuracy.

	ω_{cut}	T_O (s)	N	Δt_{time_step}
$(H_S=2m \ T_P=7s)$	3.5904	3600	2057	0.875
$(H_S=2m \ T_P=17s)$	1.4784	3600	847	2.125
$(H_S=8m \ T_P=7s)$	3.5904	3600	2057	0.875
$(H_S=8m \ T_P=17s)$	1.4784	3600	847	2.125

5.5.3 Results and Discussions

The accuracy of the proposed approach against the number of discretization steps applied to the copula is examined. The base shear is estimated by taking an overall exceedance probability of 0.01 according to Eq. (5.28). Following the procedures stated in Section 5.5.1, the method is initially performed with a discretization of 16 subcopulas. Both H_S and T_p in the copula space are equally divided into four subdomains, each covering one quartile of the associated variable. The structural analysis is then performed in each subcopula by using the sea state value (H_S, T_P) at the center point, as shown in Fig. 5.22. Based on structural analysis, the conditional extreme response probabilistic model $\bar{Q}(\cdot)$ is constructed for each of the subcopula. Here, three values of $\Delta P_{required}$ are chosen and used in the discretization steps to compare the results. These are 0.005, 0.003 and 0.0015 which correspond to 50%, 30% and 15% of the total exceedance probability of 0.01. For the subcopulas's ΔP_E , if it is larger than $\Delta P_{required}$, a further discretization will be carried out. This is performed by dividing both dimensions of the subcopula domain into four equal size subcopulas. Then the long term base shear is computed based on the discretized model. After the results are obtained, the individual long term exceedance probability ΔP_E in each

subcopula is calculated. The whole calculation is repeated until all the ΔP_E in the subcopula is smaller than $\Delta P_{required}$. The detailed steps can be seen from Fig. 5.22.

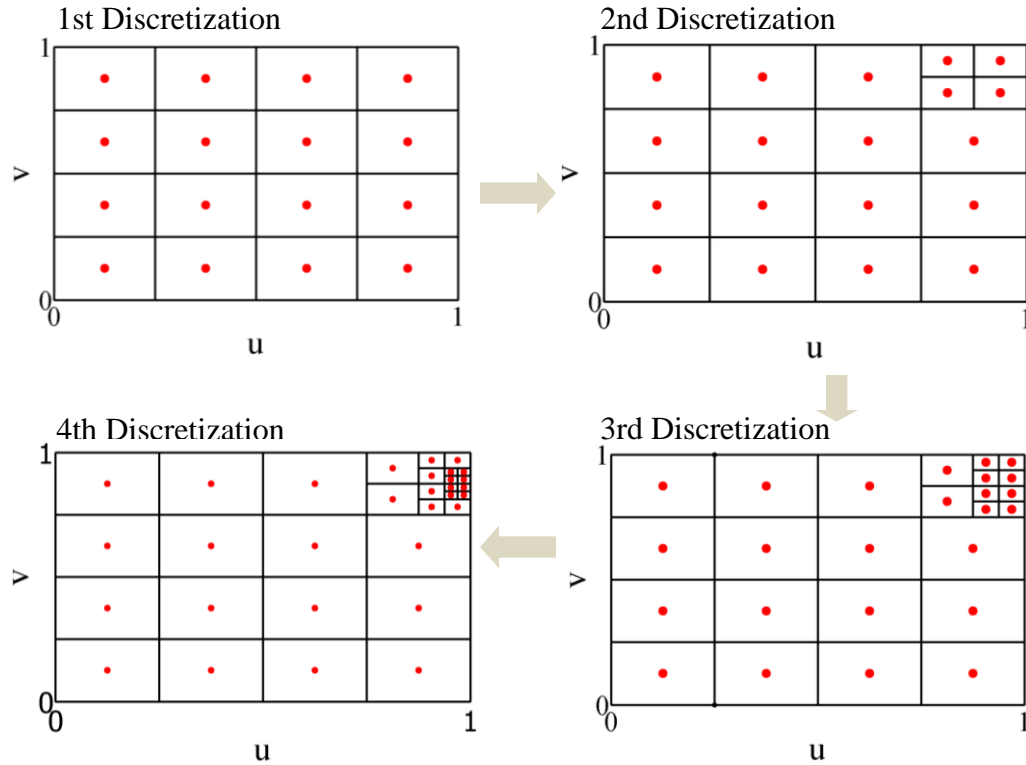


Figure 5.22 Schematic showing of discretization steps ($u=F_1(H_S)$, $v=F_2(T_P)$) and evaluation points (red dots).

As shown in Table 5.9, there are totally four discretization steps performed in the calculation with a final result for the determined long term base shear of $1.1 \cdot 10^7 \text{N}$. For the tested $\Delta P_{required}$ of 0.005, the computation stops at the third step, whereas for $\Delta P_{required} = 0.003$ and $\Delta P_{required} = 0.0015$, the computation stops at the fourth step. The number of subcopulas is increased from 16 to 31 through the four discretization steps. The value of the maximum ΔP_E of all the subcopulas within each step has decreased from 0.0081 to 0.0014 from the first step to the fourth step. Apparently, with the applied discretization, it is expected

that the accuracy of the estimated long term value is improved as more discretized blocks are created. This can be seen from the value of the exceedance probability in each subcopula's conditional extreme response distribution $\bar{Q}(X(t) > l|H_S, T_P)$. It is observed that $\bar{Q}(X(t) > l|H_S, T_P)$ is increasing as the discretization increases from 0.0701 to 0.8235. This generally implies that the sea state values (H_S, T_P) which may result in the most significant characteristic extreme response of the structure are identified in the copula space. As more blocks around these identified regions are created through the discretization steps, the estimated long term value becomes more accurate since the integration of Eq. (2.12) is quite sensitive to these sea state values. In addition the accuracy of the result could also be inspected from the change of the determined long term value in each step. The change of the estimated long term value at each discretization step diminishes from the second discretization step (+10.6%) to the fourth (+1.6%), see Table 5.9. This can also be observed from the value of the maximum subcopula ΔP_E which does not change too much from the third discretization step to the fourth (-17.6%). It gives an indication that a further discretization step may seem unnecessary as this may not cause much change to the results for the determined long term value. In fact, the long term value determined by any further discretization may tend to vary only slightly and converge to the exact solution.

Generally, the proposed approach is able to reflect the importance of the environmental parameter values to the determined long term value. This can be observed by the ΔP_E of each subcopula. For example, as given in Appendix A, at

Table 5.9 Results of proposed approach in determining long term base shear for $P_E=0.01$ (change of estimated long term value is given in bracket).

	No. of Subcopulas	Estimated Long Term Base Shear (N)	Maximum ΔP_E of Subcopulas	Maximum $\bar{Q}(\cdot)$ of Subcopulas
1st Discretization	16	8937629	0.0081	0.0701
2nd Discretization	19	9886673 (+10.6%)	0.0054	0.2116
3rd Discretization	25	10806707 (+9.3%)	0.0044	0.4688
4th Discretization	31	10976878 (+1.6%)	0.0014	0.8235

Notes: $\Delta P_E = \bar{Q}(X(t) > l|H_s, Tp)\Delta C$ represents value of long term exceedance probability in subcopulas, total exceedance probability is calculated by sum of ΔP_E in each subcopula as $P_E = \sum \Delta P_E$.

the 4th discretization step, the values of ΔP_E for subcopulas No. 30, 32, 34 and 36 are 0.0012, 0.0014, 0.0014 and 0.0013, which are much larger compared to the other subcopulas. This means that in other subcopulas, either the conditional response distribution gives a small exceedance probability (e.g. $\bar{Q}(X(t) > l|H_s, Tp)$ is quite small) or the probability of having that sea state values are small (e.g. ΔC is quite small). The discretization to these subcopulas is thus less important as the values of ΔP_E are even smaller if it is discretized into finer subdomains. Additionally, with these procedures in identifying the critical sea states, the numerical efforts have been largely saved. The number of simulations needed are only 360 (each subcopula requires 10 realizations) in the derived model which is far less compared to 2560 (based on 10 realizations for each smallest subcopula $16 \times 16 \times 10=2560$) in the sea state block method. As each simulation requires one hour's time (duration for a stationary sea state), the proposed procedure could save 85.9% computational time compared to the sea state block method. Moreover, if any further accuracy level is required (e.g. a

precision of 0.0001 is required), the results can be achieved by performing repeated discretization steps in the copula space by using the current approach (e.g. the computation can be performed until the change of the estimated long term value between two discretization steps is less than 0.0001). This has been demonstrated in the example where the result of determined long term value is more accurate for using $\Delta P_{required} = 0.0015$ compared to $\Delta P_{required} = 0.005$. However, the accuracy can hardly be controlled in the sea state block method.

5.6 Concluding Remarks

The copula approach has been introduced to model the ocean parameters in a multivariate setting. The conditional probability distribution approach and the Nataf approach in solving multivariate statistical problems are compared with the copula approach, specifically for bivariate data (H_S, T_P) and (H_S, V_W) . Considering that the collected data possesses certain degree of non-stationarity, some segmentations of the data are performed to obtain homogeneous subsets. A discretized subcopula approach is proposed for the long term performance analysis. This technique is demonstrated through the dynamic analysis of a jacket structure using the desired constructed copula model of (H_S, T_P) . The performance of each discretization step in determining the calculated long term value is checked by looking at the convergence rate in the results. This is used to decide whether a further discretization is needed or not. It is concluded from the current study that:

- The copula model can more accurately describe the statistical relationship between the ocean parameters. This is clearly shown in the comparative study where copula gives the most desirable model in characterizing the bivariate data (H_S, T_P) and (H_S, V_W) . The reason lies in the fact that copula is more flexible in modeling the dependency between the physical parameters compared to traditional approaches.
- The proposed discretized copula approach can be implemented in the calculations for estimating the long term value. The derived procedure in discretizing the copula domain into subcopulas based on the exceedance probability is proven to be quite effective in the calculation compared to the sea state block method. This reduces the number of simulations in the computation. The convergence rate in the results regarding to each discretization step provides the information about whether the required level of accuracy is achieved. This provides a numerically efficient and accurate way for the practical long term assessment of an offshore structure within a multivariate environment.

Chapter 6 Conclusions

6.1 Summary of Thesis

The main objective of this study is to propose a framework for the reliability assessment of fixed jacket offshore structures under the wave loading arising from a time varying or multivariate environment. This study focused on establishing a robust and accurate statistical model for the wave height in order to minimize the risks that may be involved during the analysis of an existing fixed jacket structure. Selected ocean data from Wave Information Studies (WIS) data base are used in the study to demonstrate the suitability of the proposed procedures in handling the offshore reliability problems associated with time varying and multivariate environment.

From the studies carried out in this work, the conclusions can be summarized in the following:

- **Establishing a Robust Extreme Value Model** — In establishing a robust extreme model for the assessment of structure's long term performance, the selection of an appropriate method should be carefully considered. Based on the limited amount of data used, the MOM, PWMB and PWMU proved to be the better parameter estimate methods compared to MLE, ADR and KS. Given the presence of random noise, the constructed extreme model may be unreliable and may need to be improved. Thus, data pretreatment may be required to remove the non-stationarity inherent

in the raw data before it can be used in a statistical analysis. For collected time series data that contains high serial correlations, the POT method is found to be the most appropriate method compared to the annual maximum method and the r largest order statistic method. It is also a better choice if a limited time series data is available. However, the performance of the POT method is largely dependent on the appropriate choice of time span and threshold. It is found that a random set based imprecise probability model is a feasible approach that is able to describe the uncertainties associated with the selection of threshold and time span in the POT method. As such, the bounds in the imprecise probability model can provide a consistent means to present the uncertainties in a quantitative form.

- **Reliability Analysis of Offshore Structures within a Time Varying Environment** — Based on the analysis of the selected ocean data, it was found that the seasonality has great influence to the observed wave height magnitude whereas the wave occurrence rate is largely affected by the directionality. The proposed segmentation algorithm is able to divide the data into appropriate sets for constructing the Poisson-GPD model. The use of discrete statistical model together with the Fourier characterization in representing the time varying effect in the extreme wave height was found particularly effective and accurate. The discrete model provides the information of failure probability for an offshore structure within different periods in a reliability analysis. By incorporating such covariate effect in

the statistical model can more adequately reflect the underlying physical processes and help to identify the most critical environmental conditions for engineering structures under the varying environment.

- **Reliability Analysis of Offshore Structures within a Multivariate Environment** — In the modeling of multiple variables, the copula model was shown to be more comprehensive than the traditional conditional joint model and Nataf model. This is due to the fact in the copula model, fewer coefficients are used and a number of copula functions are available to characterize the nonlinear statistical dependencies between the variables. The implementation of the copula model in the calculation of long term value of an offshore structure is conducted. A discretized copula approach has been proposed in the calculation to estimate the long term value. It was found such development reduces the number of computations in the structural analysis for considering all the multivariate environmental conditions in deriving the long term response distribution.

6.2 Recommendation of Future Works

The illustrated methodology for solving the time varying reliability problem may be used to include more factors in the long term reliability analysis. In particular, the effect of deterioration on a structure (for example, a decrease in the structure's strength caused by the corrosion) may be included in the investigation. An algorithm may be developed to discretize the process of deterioration into different time intervals and assess the reliability for the whole operation period of

the structure from each piecewise evaluation. The modeling of the deterioration must be carefully studied. It may also be necessary to develop an efficient technique to accurately estimate the failure probability with these issues. Another important issue is to study the effect of climatic change. The present study analyzed the reliability of the offshore structure based on a wave model which assumes that the environment does not change significantly over the period of interest. There may be a need to carry out a sensitivity analysis based on an appropriate climate numerical model on the predicted design value.

For the multivariate reliability analysis, the copula-based approach is easy to generalize to a multi-dimensional setting. The procedures in calculating the exceedance probability P_E in Eq. (5.41) would not be different except for the higher dimension. For example, if the environmental condition concerns a trivariate case (H_S, T_P, V_W) , the proposed discretization steps remain the same. The only difference is that the short term structural analysis is performed with a discretized hypercubic subcopula in three-dimensional space. The probability ΔC within each subcopula should be calculated based on the third-order difference from the three-dimensional copula function. The results in terms of accuracy and numerical efforts can be controlled through finer discretizations in selected sub-domains as described in the two-dimensional case in Chapter 5. Figure 6.1 gives a general schematic showing this computation process. The structural analysis would become more tedious due to the greater number of sub-domains.

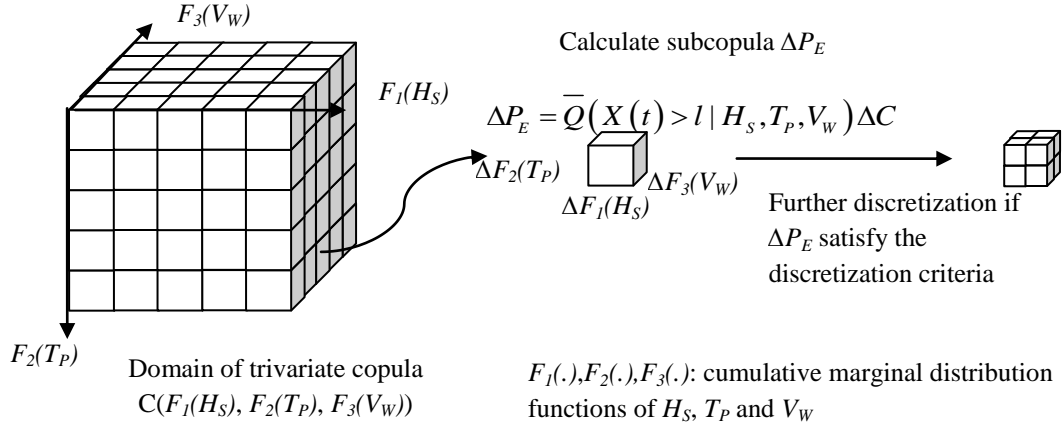


Figure 6.1 Schematic showing of discretization procedures for three-dimensional copula model.

Key to this extension is establishing a three-dimensional copula model. There are various ways of constructing a trivariate copula (Salvadori & De Michele 2004, Cherubini et al. 2004, Agha Kouchak et al. 2010). Generally, it is most convenient to adopt the concepts of nested copula approach and vine copula approach in the construction of multivariate copula.

The nested copula is built by using bivariate copulas as the basis and linked them together to form the multivariate copula. For instance, the construction of a trivariate copula for (H_S, T_P, V_W) is given by

$$C(F_1(H_S), F_2(T_P), F_3(V_W)) = C_2(C_1(F_1(H_S), F_2(T_P)), F_3(V_W)) \quad (6.1)$$

where the C_1 and C_2 are bivariate copulas that are nested to construct a trivariate copula. From Eq. (6.1), one can argue that alternative combinations of variables can be employed. In addition, different copula functions can be applied. For instance, the two-dimensional copula can first be constructed between T_P and V_W

instead of H_S and T_P . An important aspect of this approach is that the C_2 is constructed considering the dependency between $F_3(V_W)$ and C_1 instead of $F_1(H_S)$ or $F_2(T_P)$.

The vine copula approach on the other hand utilizes the conditional distribution models (Kurowicka & Joe 2010). The basic idea is to decompose the multivariate copulas into a cascade of bivariate copulas. For example, if the conditional cumulative marginal distribution functions, $F_{H_S|V_W}(H_S/V_W)$ and $F_{T_P|V_W}(T_P/V_W)$, of H_S and T_P for a given value of V_W are known, the trivariate copula can be constructed through the following:

$$C(F_1(H_S), F_2(T_P), F_3(V_W)) = \int_0^{V_W} C_3(F_{H_S|V_W}(H_S|t), F_{T_P|V_W}(T_P|t)) F_3(t) dt \quad (6.2)$$

where the C_3 is the bivariate copula for the conditional cumulative distribution functions. The conditional distribution function $F_{H_S|V_W}(H_S/V_W)$ and $F_{T_P|V_W}(T_P/V_W)$ can be further expressed in terms of partial derivatives of copulas as:

$$F_{H_S|V_W}(H_S|t) = \frac{\partial C_4(F_1(H_S), F_3(t))}{\partial V_W} \quad (6.3)$$

$$F_{T_P|V_W}(T_P|t) = \frac{\partial C_5(F_2(T_P), F_3(t))}{\partial V_W} \quad (6.4)$$

where C_4 and C_5 are the bivariate copulas for (H_S, V_W) and (T_P, V_W) . As a result, the dependencies between each pair of the three parameters have been characterized through a copula. Therefore, it may be more physically appealing as more copulas are included compared with the nested approach. Another advantage is that the conditional distribution functions are sometimes

theoretically available which could make the construction of copula much easier, such as Eqs. (6.3)-(6.4) obtained from the literature (DNV 2006). The method of constructing the copula for H_S , T_P and V_W can also be extended to model an environmental condition that has more than three parameters. This is achieved by extending recursively the approach outlined above.

Although the copula based multivariate analysis appears conceptually simple, there are certain considerations which need to be taken care of. Compatibility problems may arise (Nelson 2006) when establishing Eq. (6.1), which necessitates further study to develop a way of constructing a reliable copula model. In addition, there is a need to reduce the discretization steps in the proposed approach to make the solution tractable. The accuracy of the final results is largely dependent on the discretized sector. It becomes inevitable that there is a need to develop efficient discretization and computation algorithms for higher dimension copulas.

REFERENCES

- Agarwal, P. and Manuel, L. Incorporating irregular nonlinear waves in coupled simulation and reliability studies of offshore wind turbines, *Applied Ocean Research*, 33(3): 215-227. 2011.
- Agarwal, P. and Manuel, L. Simulation of offshore wind turbine response for long-term extreme load prediction, *Engineering Structures*, 31(10): 2236-2246. 2009.
- Agha Kouchak, A., Bárdossy, A. and Habib, E. Copula-based uncertainty modelling: application to multisensor precipitation estimates, *Hydrological Processes*, 24(15): 2111-2124. 2010.
- Akaike, H. Information theory and an extension of the maximum likelihood principle. In Petrov, B.N. & Csáki, F. (ed.) *Proceedings of the 2nd international symposium on information theory*, Akad Kiadó Budapest. 1973.
- Allan, J.C. and Komar, P.D. Are ocean wave heights increasing in the eastern North Pacific? *EOS, Transaction of the American Geophysical Union*, 81 (47), 561-567. 2000.
- An, Y. and Pandey, M.D. The r largest order statistics model for extreme wind speed estimation, *Journal of Wind Engineering and Industrial Aerodynamics*, 95 (3): 165-182. 2007.
- Anderson, C.W., Carter, D.J.T. and Cotton, P.D. Wave climate variability and impact on offshore design 4mxtremes, Report for Shell International and the Organization of Oil & Gas Producers, 90. 2001.
- Andrade, C. and Castillo, A. Evolution of reinforcement corrosion due to climatic variations. *Material Corrosion* 54(6): 379-386. 2003.
- Ang, A. H-S. and Tang, W.H. *Probability Concepts in Engineering Planning and Design*, vol. 2, Decision, Risk, and Reliability. John Wiley & Sons, New York. 1984.
- Ang, A. H-S. and Tang, W.H. *Probability Concepts in Engineering Planning and Design*, vol. 1, Basic Principles. New York: John Wiley & Sons. 2007.
- Ang, G.L., Ang, A. H.-S. and Tang, W.H. Optimal importance-sampling density estimator, *Journal of Engineering Mechanics*, 118(6): 1146-1163. 1992.
- Ariff, N. M., Jemain, A.A., Ibrahim, K. and Wan Zin, W.Z. IDF relationships using bivariate copula for storm events in Peninsular Malaysia, *Journal of Hydrology*, 470: 158-171. 2012.

- Ashkar, F. and Tatsambon, C.N. Revisiting some estimation methods for the generalized Pareto distribution, *Journal of Hydrology*, 346: 136-143. 2007.
- Au, S.K. and Beck, J.L. A new adaptive importance sampling scheme for reliability calculations, *Structural Safety*. 21:135-158. 1999.
- Au, S.K. and Beck, J.L. Estimation of small failure probabilities in high dimensions by subset simulation, *Probabilistic Engineering Mechanics*, 16(4): 263-277. 2001a.
- Au, S.K. and Beck, J.L. First excursion probabilities for linear systems by very efficient importance sampling, *Probabilistic Engineering Mechanics*, 16:193-207. 2001b.
- Augenbaugh, J.M. and Paredis, C.J.J. The Value of Using Imprecise Probabilities in Engineering Design. *Journal of Mechanical Design* 128(4), 969-979. 2006.
- Baarhold, G.S., Haver, S. and Økland, O.D. Combining contours of significant wave height and peak period with platform response distributions for predicting design response, *Marine Structures*, 23: 147-163. 2011.
- Baarholm, G.S. and Moan, T. Efficient estimation of extreme long-term stresses by considering a combination of longitudinal bending stresses, *Journal of Marine Science and Technology*, 6(3): 122-134. 2002.
- Battjes, J.A. Long term wave height distributions at seven stations around the British Isles, *Deutsche Hydrographische Zeitschrift*, 25: 179–189. 1972.
- Baker, J.F. Plastic design of frames: fundamentals. Cambridge University Press, London. 1969.
- Beck, J.L. and Katafygiostis, L.S. Updating models and their uncertainties – Bayesian statistical framework, *Journal of Engineering Mechanics*, 124, 455-461. 1998.
- Beer, M. Engineering quantification of inconsistent information, *International Journal of Reliability and Safety*, 3, 174-200. 2009.
- Beer, M., Ferson, S. and Kreinovich, V. Imprecise probabilities in engineering analyses, *Mechanical Systems and Signal Processing*, 37 (1-2): 4-29. 2013.
- Beer, M. Fuzzy Probability Theory. In: Meyers, R. (ed.), *Encyclopedia of Complexity and Systems Science*, 6: 4047–4059. Springer, New York. 2009.
- Beguer á, S. Uncertainties in partial duration series modeling of extremes related to the choice of the threshold value. *Journal of Hydrology* 303, 215-230. 2005.

- Berens, P. CircStat: A Matlab Toolbox for Circular Statistics, *Journal of Statistical Software*, 31: 10. 2009.
- Berleant, D. Automatically verified reasoning with both intervals and probability density functions, *Interval Computations*, 2, 48-70. 1993.
- Bendat, J.S. and Piersol, A.C. *Random data: analysis and measurement procedures*. New York, Wiley. 1971.
- Bermudez, P.Z. and Kotz, S. Parameter estimation of the generalized Pareto distribution-Part I, II, *Journal of Statistical Planning and Inference*, 140(6): 1353-1397. 2010.
- Bernardini, A. and Tonon, F. *Bounding Uncertainty in Civil Engineering*. Springer-Verlag, Berlin Heidelberg. 2010
- Bernaola-Galvan, P., Oliver, J.L. and Carpena, P. Roman-Roldan R: Isochore chromosome maps of eukaryotic genomes, *Gene*, 276: 47-56. 2001.
- Biondini, F., Bontempi, F., Frangopol, D. and Malerba, P. Probabilistic service life assessment and maintenance planning of concrete structures, *Journal of Structural Engineering*, 132(5): 810-825. 2006.
- Bitner-Gregersen, E.M. and Hagen,  . Uncertainties in data for the offshore environment, *Structural Safety*, 7(1): 11-34. 1990.
- Bitner-Gregersen, E.M. and Haver, S. Joint long term description of environmental parameters for structural response calculation, *Proceedings of the 2nd International Workshop on Wave Hindcasting and Forecasting*. 1989.
- Bitner-Gregersen, E.M., Guedes Soares, C., Machado, U. and Cavaco, P. Comparison of different approaches to joint environmental modeling, *Proceedings of the 17th International Conference on Offshore Mechanics and Arctic Engineering*, ASME, New York, Vol. II. 1998.
- Bitner-Gregersen, E.M. Joint probabilistic description for combined seas, In *Proceedings of 24th International Conference on Offshore Mechanics and Arctic Engineering (OMAE)*, June 12-17, Halkidiki, Greece. 2005.
- BMT. Fluid Mechanics Limited. Review of issues associated with the stability of semi-submersibles, Health and Safety Executive. 2006.
- Boccotti, P. *Wave mechanics for ocean engineering*, Elsevier Science B.V., Amsterdam, Netherlands. 2000.
- Box, G.E.P. and Cox, D.R. An analysis of transformations, *Journal of the Royal Statistical Society Series B*, 26: 211-252. 1964.

- Bracewell, R.N. *The Fourier Transform and its Applications*. New York, McGraw-Hill. 1986.
- Bradley, J.V. *Distribution-Free Statistical Tests*. New York, Academic Press. 1968.
- Burton, T., Sharpe, D., Jenkins, N. and Bossanyi, E. *Wind Energy Handbook*. John Wiley, Chichester, England. 2001.
- Caires, S. and Sterl, A. 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data, *Journal of Climate*, 18(7): 1032-48. 2005.
- Carter, D.J.T. and Challenor, P.G. Estimating return values of environmental parameters, *Q. J. R. Meteorology Society*, 107: 259–266. 1981.
- Castillo, E. and Hadi, A.S. Fitting the generalized Pareto distribution to data, *Journal of American Statistics Association*, 92(440): 1609-1620. 1997.
- Castillo, E., Hadi, A.S., Balakrishnan, N. and Sarabia, J.M. *Extreme Value and Related Models with Applications in Engineering and Science*, Wiley, Chichester. 2004.
- Celik, A.N. A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey, *Renew Energy*, 29 (4): 593–604. 2004.
- Chakrabarti, S.K. *Hydrodynamics of offshore Structures*. Computational Mechanics Publications, Southampton, UK, 1987.
- Chen, L., Singh, V., Shenglian, G., Hao, Z. and Li, T. Flood coincidence risk analysis using multivariate copula functions, *Journal of Hydrologic Engineering*, 17(6): 742-755. 2012.
- Chen, T.C. and Yoon, J. Interdecadal variation of the North Pacific wintertime blocking, *Monthly Weather Review*, 130: 3136-3143. 2002.
- Cheng, P.W., van Bussel, G.J.W., van Kuik, G.A.M. and Vugts, J.H. Reliability-based design methods to determine the extreme response distribution of offshore wind turbines, *Wind Energy*, 6(1): 1-22. 2003.
- Cherubini, U., Luciano, E. and Vecchiato, W. *Copula Methods in Finance*. John Wiley & Sons, New York. 2004.
- Ching, J.Y. and Chen, J.R. Predicting displacement of augered cast-in-place piles based on load test database, *Structural Safety*, 32(6): 372–383. 2010.

- Choulakian, V. and Stephens, M.A. Goodness-of-Fit Tests for the Generalized Pareto Distribution, *Technometrics*, 43(4): 478-484. 2001.
- Coles, S.G. An introduction to statistical modeling of extreme values. Springer, London. 2001.
- Coles, S.G. and Walshaw, D. Directional modeling of extreme wind speeds, *Applied Statistics-Journal Of The Royal Statistical Society Series C*, 43(1): 139-157. 1994.
- Corti, S., Giannini, A., Tiabldi, S. and Molteni, F. Patterns of low-frequency variability in a three-level quasi-geostrophic model, *Climate Dynamics*, 13: 883-904. 1997.
- Corbella, S. and Stretch, D.D. Simulating a multivariate sea storm using Archimedean copulas, *Coastal Engineering*, 76: 68-78. 2013.
- Cunnane, C. A note on the Poisson assumption in partial duration series models, *Water Resources Research*, 15 (2): 489-494. 1979.
- Danielsson, J., de Haan, L., L., P. and de Vries, C.G. Using a bootstrap method to choose the sample fraction in tail index estimation, *Journal of Multivariate Analysis*, 76: 226-248. 2001.
- Davison, A.C., Padoan, S.A. and Ribatet, M. Statistical modelling of spatial extremes, *Statistical Science*, 27: 161-186. 2012.
- Davison, A.C. and Smith, R.L. Models for exceedances over high thresholds (with discussion), *Journal of the Royal Statistical Society*, B52, 393-442. 1990.
- De Michele, C., Salvadorib, G., Passoni, G. and Vezzoli, R. A multivariate model of sea storms using copulas, *Coastal Engineering*, 54(10): 734-751. 2007.
- De Zea Bermudez, P. and Kotz, S. Parameter estimation of the generalized Pareto distribution – Parts I & II, *Journal of Statistical Planning and Inference*, 140: 1374-1388. 2010.
- Degrauwe, D., Lombaert, G. and De Roeck, G. Improving interval analysis in finite element calculations by means of affine arithmetic. *Computers and Structures* 88 (3-4), 247-254. 2010.
- Deidda, R. and Puliga, M. Performances of some parameter estimators of the generalized Pareto distribution over rounded-off samples, *Physics and Chemistry of the Earth*, 34(10-12): 626-634. 2009.
- Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping, *Ann Math Statist*, 38: 325-39. 1967.

- Deodatis, G. and Shinozuka, M. Simulation of seismic ground motion using stochastic waves, *Journal of Engineering Mechanics-ASCE*, 115(2): 2723-2737. 1989.
- Der Kiureghian, A. and Dakessian, T. Multiple design points in first and second-order reliability, *Structural Safety*, 20:37-49. 1998.
- Dijkstra, T.A. and Dixon, N. Climate change and slope stability in the UK: Challenges and approaches, *Journal of Engineering Geology and Hydrogeology*, 43(4):371-385. 2010.
- Ditlevsen, O. Stochastic model for joint wave and wind loads on offshore structures, *Structural Safety*, 24(2-4): 139-163. 2002.
- DNV. Technical Report 95-3203, Det Norske Veritas. 1995.
- DNV. Environmental conditions and environmental loads, Recommended Practice, DNV-RP-C205. 2007.
- DNV. Environmental conditions and environmental loads, Recommended Practice, DNV-RP-C205. 2012.
- Dong, S., Wang, N. and Liu, W. Bivariate maximum entropy distribution of significant wave height and peak period, *Ocean Engineering*, 59: 86-99. 2013.
- Dourte, D., Shukla, S. and Singh, P. Rainfall Intensity-Duration-Frequency Relationships for Andhra Pradesh, India: Changing Rainfall Patterns and Implications for Runoff and Groundwater Recharge, *Journal of Hydrologic Engineering*, 18(3): 324-330. 2013.
- Dukes, M.D.G. and Palutikof, J.P. Estimation of extreme wind speeds with very long return periods, *Journal of Applied Meteorology*, 34(9): 1950-1961. 1995.
- Efron, B. and Tibshirani, R.J. *An Introduction to the Bootstrap*. Chapman & Hall, New York. 1993.
- Elsinghorst, C., Jonathan, P., Smulders, L., Taylor, P.H. and Groeneboom, P. Extreme value analysis of North Sea storm severity, *J. Offshore Mechanics Arctic Engineering*, 120:177-183. 1998.
- Embrechts, E. Copulas: a personal view, *Journal of Risk and Insurance*, 76(3): 639-650. 2009.
- Enders, W. *Applied Econometric Time Series*. John Wiley & Sons, United States. 2004.

- Enszer, J.A., Lin, Y., Ferson, S., Corliss, G.F. and Stadtherr, M.A. Probability bounds analysis for nonlinear dynamic process models, *AIChE Journal*, 57(2): 404-422. 2011.
- Ernst, B. and Seume, J.R. Investigation of site-specific wind field parameters and their effect on loads of offshore wind turbines, *Energies*, 5 (10): 3835-3855. 2012.
- Farnes, K. and Moan, T. Extreme dynamic, nonlinear response of fixed platforms using a complete long-term approach, *Applied Ocean Research*, 15(6): 317-326. 1994.
- Fellin, W., Lessmann, H., Oberguggenberger, M. and Vieider, R. Analyzing uncertainty in civil engineering. Springer-Verlag, Berlin. 2005.
- Fellin, W., King, J., Kirsch, A. and Oberguggenberger, M. Uncertainty modelling and sensitivity analysis of tunnel face stability, *Structural Safety*, 32(6): 402–410. 2010.
- Ferreira, A. and Guedes Soares, C. Modelling the Long-Term Distribution of Significant Wave Height with the Beta and Gamma Models, *Ocean Engineering*, 26(8): 713-725. 1999.
- Ferreira, J.A. and Guedes Soares, C. Modelling bivariate distributions of significant wave height and mean wave period, *Applied Ocean Research*, 24(1): 31-45. 2002.
- Ferreira, J.A., and Guedes Soares, C. An application of the peaks over threshold method to predict extremes of significant wave height, *Journal of Offshore Mechanics and Arctic Engineering*, 120: 165-176. 1998.
- Ferson, S. and Donald, S. Probability Bounds Analysis, Probabilistic Safety Assessment and Management, Mosleh, A., and Bari, R.A. (eds). Springer-Verlag, New York. 1998.
- Ferson, S. and Hajagos, J.G. Arithmetic with uncertain numbers: rigorous and (often) best possible answers, *Reliability Engineering & System Safety*, 85(1–3): 135–152. 2004.
- Ferson, S. RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers. Lewis Publishers, Boca Raton, Florida. 2002.
- Fitzwater, L.M. and Winterstein, S.R. Predicting Design Wind Turbine Loads from Limited Data: Comparing Random Process and Random Peak Models. In *ASME Wind Energy Symposium*, 355-364, AIAA-2001-0046, Reno, NV. 2001.

- Forristall, G.Z. On the use of directional wave criteria, *Journal of Waterway Port Coastal and Ocean Engineering*, 130: 272–275. 2004.
- Franchin, P., Ditlevsen, O. and Kiireghian, A.D. Model correction factor method for reliability problem involving integrals of non-Gaussian random fields, *Probabilistic Engineering Mechanics*, 17: 109-122. 2002.
- Frangopol, D.M., Kallen, M-J. and Noortwijk, M. Probabilistic models for life-cycle performance of deteriorating structures: review and future directions, *Progress in Structural Engineering and Materials*, 6: 197–212. 2004.
- Frees, E. and Valdez, E. Understanding Relationships Using Copulas, *North American Actuarial Journal*, 3(1): 137-142. 1999.
- Furrer, R., Nychka, D. and Sain, S. fields: Tools for spatial data, R package version 6.3. 2010.
- Galambos, J., Leigh, S. and Simiu, E. *Extreme Value Theory and Applications*. Kluwer, Amsterdam. 1994.
- Galeano, P. The use of cumulative sums for detection of change points in the rate parameter of a Poisson Process, *Computational Statistics and Data Analysis*, 51(12): 6151-6165. 2007.
- Galiatsatou, P. and Prinos, P. Modeling non-stationary extreme waves using a point process approach and wavelet, *Stochastic Environmental Research Risk Assessment*, 25: 165-183. 2011.
- Gao, Z. and Moan, T. Extreme Value Prediction of Inundation Drag Force with and without Current, *Ocean Engineering*, 36: 1244-1250. 2009.
- García-Ruiz, J.M., Arnáez, J., White, S.M., Lorente, A. and Beguería, S. Uncertainty assessment in the prediction of extreme rainfall events: an example from the Central Spanish Pyrenees, *Hydrological Processes*, 14: 887–898. 2000.
- Genest, C. and Favre, A. Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, 12(4): 347-368. 2007.
- Genest, C., Gendron, M. and Bourdeau-Brien, M. The advent of copulas in finance, *European Journal of Finance*, 15: 609-618. 2009.
- Genest, C. and MacKay, R.J. The joy of copulas: Bivariate distributions with uniform marginals, *The American Statistician*, 40: 280-283. 1986.
- Goda, Y. *Random seas and design of maritime structures*. University of Tokyo Press, Japan. 2000.

- Goda, K. Statistical modeling of joint probability distribution using copula: Application to peak and permanent displacement seismic demands, *Structural Safety* 32(2): 112-123. 2010.
- Gomes, M.L. On the estimation of parameters of rare events in environmental time series. *Statistics for the environment*, Barnett, V., and Turkman, K.F., (Eds), Wiley, New York. 1993.
- Goutsias, J., Mashler, R.P.S. and Nguyen, H.T. *Random sets*. Springer-Verlag, New York. 1997.
- Graham, C.G. Problems with the design statistical approach, In *Proceedings of Extreme Value Analysis of Directional Wind and Wave Data*, International Conference on Wave and Wind Directionality, Paris, 1981.
- Green, P.J. and Silverman, B.W. *Nonparametric regression and generalised linear models: a roughness penalty approach*. Chapman and Hall, London, UK. 1994.
- Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, *Water Resources Research*, 15 (5): 1049-1054. 1979.
- Guedes Soares, C., Lopes, L.C. and Costa, M.D.S. *Wave Climate Modelling for Engineering Purposes*, *Computer Modelling in Ocean Engineering*, Rotterdam: A.A. Balkema Pub. 1988.
- Guedes Soares, C. and Henriques, A.C. Statistical uncertainty in long-term distributions of significant wave height, *Journal of Offshore Mechanics and Arctic Engineering*, 11: 284–291. 1996.
- Guedes Soares, C. and Scotto, M.G. Modelling uncertainty in long-term predictions of significant wave height, *Ocean Engineering*, 28:329–342. 2001.
- Guedes Soares, C. and Scotto, M.G. Application of the r largest-order statistics for long-term predictions of significant wave height, *Coastal Engineering*, 51(5-6): 387-394. 2004.
- Guedes Soares, C. and Scotto, M.G. Application of the r -order statistics for long-term predictions of significant wave heights, *Coastal Engineering*, 54(5): 393-400. 2007.
- Guedes Soares, C. *Probabilistic models of waves in the coastal zone*, Elsevier *Oceanography Series*, 67: 159-187. 2003.
- Guedes Soares, C., Weisse, R., Alvarez, E. and Carretero, J.C. A 40 years hindcast of wind, sea level and waves in European waters. In *Proceedings of*

- the 21st International Conference on Offshore Mechanics and Arctic Engineering (OMAE 2002), 23–28 June 2002, Oslo, Norway.
- Gulev, S.K. and Grigorieva, V. Last century in ocean wind wave height from global visual wave data, *Geophysical Research Letters*, 31(24): 1-4. 2004.
- Gumbel, E.J. *Statistics of Extremes*. Columbia University Press, New York. 1958.
- Hailperin, T. *Boole's Logic and Probability*. North-Holland, Amsterdam. 1986.
- Hanss, M. and Turrin, S. A Fuzzy-Based Approach to Comprehensive Modeling and Analysis of Systems with Epistemic Uncertainties, *Structural Safety*, 32(6): 433–441. 2010.
- Harr, M.E. *Reliability-based design in civil engineering*. McGraw-Hill, New York. 1987.
- Harrison, G.P. and Wallace, R.A. Sensitivity of wave energy to climate change, *IEEE Transactions on Energy Conversion*, 20(4): 870-877. 2005.
- Hasofer, A.M. and Lind, N.C. Exact and invariant second-moment code format, *Journal Of The Engineering Mechanics Division-ASCE*, 100: 111-121. 1974.
- Haver, S. Wave climate off northern Norway, *Applied Ocean Research*, 7: 85–92. 1985.
- Hawkins, D.M. Fitting multiple change-point models to data, *Computational Statistics and Data Analysis*, 37: 323–341. 2001.
- Heideman, J. Parametric Response Model for Wave/Current Joint Probability, *API TAC 88-20*, Jan 1980.
- Helton, J.C. and Oberkampf, W.L. Special Issue on alternative representations of epistemic uncertainty, *Reliability Engineering & System Safety*, 85. 2004.
- Hodapp, D.P., Kim, D.H. and Troesch, A.W. On the finite approximation of a Gaussian process and its effect on extreme value theory, *Ocean Engineering*, 58: 135-143. 2013.
- Hokstad, P., Oien, K. and Reinertsen, R. Recommendations on the use of expert judgment in safety and reliability engineering studies: Two offshore case studies, *Reliability Engineering & System Safety*, 61(1-2): 65-76. 1998.
- Hosking, J.R.M., Wallis, J.R. and Wood, E.F. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, 27: 251-261. 1985.

- Hosking, J.R.M. L-moments: analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society, Series B* 52: 105–124. 1990.
- Hughes, P.N., Glendinning, S., Mendes, J., Parkin, G., Toll, D.G., Gallipoli, D. and Miller, P.E. Full-scale testing to assess climate effects on embankments. In: *Proceedings of the Institution of Civil Engineers: Engineering Sustainability*, 162(2): 67-79. 2009.
- Hyman, J.M. FORSIG: An Extension of FORTRAN with Significance Arithmetic. LA-9448-MS, Los Alamos National Laboratory, Los Alamos, New Mexico. 1982.
- Inclán, C. and Tiao, G.C. Use of cumulative sums of squares for retrospective detection of changes of variance, *Journal of the American Statistical Association*, 89: 913–923. 1994.
- IPCC. *Climate change 2007: synthesis report*, Intergovernmental Panel on Climate Change. 2007.
- Ivan, K. and Yan, J. Modeling Multivariate Distributions with Continuous Margins Using the copula R Package, *Journal of Statistical Software*, 34(9): 1-20. 2010.
- Jaspers, N.H. Statistical distribution patterns of ocean waves and of wave induced stresses and motions with engineering applications, *Transactions Society Naval Architects and Marine Engineers*, 64: 375–432. 1956.
- Jiang, W. and Xilia, L. Prediction of structural performance and life-cycle analysis based on Bayesian dynamic models. In: Augusti, G., Schuëler, G.I. and Ciampoli, M. (eds) *ICOSSAR'05*. Millpress, Rotterdam, 1679–86. 2005.
- Joe, H. *Multivariate Models and Dependence Concepts*, Chapman & Hall, London. 1997.
- Jonathan, P. and Ewans, K. Modeling the seasonality of extreme waves in the Gulf of Mexico, *Journal of Offshore Mechanics and Arctic Engineering*, 133: 0211041-49. 2011.
- Jonathan, P. and Ewans, K. The Effect of Directionality on Extreme Wave Design Criteria, *Ocean Engineering*, 34(14-15): 1977-1994. 2007.
- Jonathan, P. and Ewans, K.C. Uncertainties in extreme wave height estimates for hurricane dominated regions. In: *Proceedings of the 25th International Conference on Offshore Mechanics and Arctic Engineering*, June 4–8, 2006, Hamburg, Germany.

- Jonathan, P., Flynn, J. and Ewans, K. Joint modeling of wave spectral parameters for extreme sea states, *Ocean Engineering*, 37(11-12): 1070-1080. 2010.
- Jonathan, P., Ewans, K. and Forristall, G. Statistical estimation of extreme ocean environments: the requirement for modeling directionality and other covariate effects, *Ocean Engineering*, 35: 1211–1225. 2008.
- Jonathan, P. and Ewans, K. Statistical modeling of extreme ocean environments for marine design: A review, *Ocean Engineering*, 62: 91–109. 2013.
- Kao, S.C. and Govindaraju, R.S. A copula-based joint deficit index for droughts, *Journal of Hydrology*, 380(1-2): 121-134. 2010.
- Katz, R.W., Parlange, M.B., and Naveau, P. Statistics of extremes in hydrology, *Advances in Water Resources*, 25: 1287-1304. 2002.
- Kawano, K. and Venkataramana, K. Dynamic response and reliability analysis of large offshore structures, *Computer Methods in Applied Mechanics and Engineering*, 168(1-4): 255-272. 1999.
- Kharin, V.V. and Zwiers, F.W. Estimating extremes in transient climate change simulations, *Journal of Climate*, 18: 1156–1173. 2005.
- Kiureghian, A.D. and Liu, P.L. Structural reliability under incomplete probability information, *Journal of Engineering Mechanics-ASCE*, 112(1): 85-104. 1986.
- Klir, G. *Uncertainty and information: foundations of generalized information theory*. Hoboken: Wiley-Interscience. 2006.
- Klir, G. and Yuan, B. *QS-CALC: An Interpreter for Uncertainty Propagation*. Quaternary Software, Fort Collins, Colorado. 1995.
- Kollat, J. B., Kasprzyk, J. R. and Thomas, W. O. Estimating the Impacts of Climate Change and Population Growth on Flood Discharges in the United States, *Journal of Water Resources Planning and Management-ASCE*, 138(5): 442-452. 2012.
- Kotz, S. and Nadarajah, S. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, London, UK. 2000.
- Kotz, S. Parameter estimation of the generalized Pareto distribution-Part I, II, *Journal of Statistical Planning and Inference*, 140(6): 1353-1397. 2010.
- Kurowicka, D. and Joe, H. *Dependence modeling: handbook on Vine Copula*. World Scientific Publishing, Singapore. 2010.

- Kyselý, J., Pícek, J. and Beranová R. Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold, *Global and Planetary Change* 72: 55–68. 2010.
- Leadbetter, M.R., Lindgren, G. and Rootzén, H. *Extremes and Related Properties of Random Sequences and Series*. Springer Verlag, New York. 1983.
- Lebrun, R. and Dutfoy, A. A generalization of the Nataf transformation to distributions with elliptical copula, *Probabilistic Engineering Mechanics* 24(2): 172-178. 2009.
- Lecacheux, S., Pedreros, R., Le Cozannet, G., Thiébot, J., De La Torre, Y. and Bulteau, T. A method to characterize the different extreme waves for islands exposed to various wave regimes: a case study devoted to Reunion Island, *Natural Hazards and Earth System Sciences*, 12(7): 2425-2437. 2012.
- Lee, Y.J. and Song, J. Risk Analysis of Fatigue-Induced Sequential Failures by Branch-and-Bound Method Employing System Reliability Bounds, *Journal of Engineering Mechanics-ASCE*, 137(12): 807-821. 2011.
- Leviandier, T. A model capable of changing the heaviness of the tail of a probability distribution function, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 55(6): 992-1001, 2010.
- Li, D.Q., Tang, X.S. and Phoon, K.K. Bivariate simulation using copula and its application to probabilistic pile settlement analysis, *International Journal for Numerical and Analytical Methods in Geomechanics* 37(6): 597-617. 2013.
- Li, D.Q., Tang, X.S. and Zhou, C.B. Uncertainty analysis of correlated non-normal geotechnical parameters using Gaussian copula, *Science China-Technological Sciences* 55(11): 3081-3089. 2012.
- Li, F., Bicknell, C. and Lowry, R. A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset, *Coastal Engineering*, 69: 1-11. 2012.
- Low, Y.M. An algorithm for accurate evaluation of the fatigue damage due to multimodal and broadband processes, *Probabilistic Engineering Mechanics*, 26(3), 435-446. 2011.
- Low, Y.M. and Cheung, S.H. On the long-term fatigue assessment of mooring and riser systems, *Ocean Engineering*, 53, 60-71. 2012.
- Lu, Y.J., Chen, Y.N., Tan, P.L. and Bai, Y. Prediction of most probable extreme values for jackup dynamic analysis, *Marine Structures*, 15(1): 15-34. 2002.
- Luceño, A. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Stat. Data Anal*, 51: 904:917. 2006.

- Luceño, A., Menéndez, M. and Menéndez, F.J. The effect of temporal dependence on the estimation of the frequency of extreme ocean climate events, *Proceedings of the Royal Society*, 462: 1683-1697. 2006.
- Mackay, E.B.L., Challenor, P.G. and Bahaj, A.B.S. A comparison of estimators for the generalised Pareto distribution, *Ocean Engineering*, 38(11-12): 1338-1346. 2011.
- Mackay, E.B.L., Challenor, P.G. and Bahaj, A.B.S. On the use of discrete seasonal and directional models for the estimation of extreme wave conditions, *Ocean Engineering*, 37(5-6): 425-442. 2010.
- Mahdia, S. and Ashkarb, F. Exploring generalized probability weighted moments, generalized moments and maximum likelihood estimating methods in two-parameter Weibull model, *Journal of Hydrology*, 285(1-4): 62-75. 2004.
- Mahjoobi, J. and Mosabbeq, E.A. Prediction of significant wave height using regressive support vector machines, *Ocean Engineering*, 36(5): 339-347. 2009.
- Mano, J.F. Keynote: Biomimetic surfaces exhibiting extreme wettability properties for tissue engineering applications, *Journal of Tissue Engineering and Regenerative Medicine*, 6(1): 229-229. 2012.
- Manwell, J.F., McGowan, J.G. and Rogers, A.L. *Wind energy explained: theory, design and application*. 2002.
- Marcos, M., Jordà G., Gomisa, D. and Begoña Pérez, B. Changes in storm surges in southern Europe from a regional model under climate change scenarios, *Global and Planetary Change*, 77(3-4): 116-128. 2011.
- Markovic, M., Krauberger, N. and Saje, M. Non-linear analysis of pre-tensioned concrete planar beams, *Engineering Structures*, 46: 279-293. 2013.
- Marx, B.D. and Eilers, P.H.C. Direct generalized additive modeling with penalized likelihood, *Computational Statistics and Data Analysis*, 28: 193-209. 1998.
- Massel, S.R. *Ocean surface waves: their physics and prediction*. World Scientific, Singapore. 1996.
- Materon, G. *Random sets and integral geometry*, Wiley. 1975.
- Mazas, F. and Hamm, L. A multi-distribution approach to POT methods for determining extreme wave heights, *Coastal Engineering*, 58(5): 385-394. 2011.
- Mazumder, R. and Mazumber, B.S. Statistical characterization of circulation patterns and direction of turbulent flow over a waveform structure, *Environmetrics* 17(5): 417-434. 2006.

- Melchers, R.E. Recent progress in the modeling of corrosion of structural steel immersed in seawaters, *Journal of Infrastructure Systems*, 12(3): 154–62. 2006.
- Melchers, R.E. *Structural Reliability Analysis and Prediction*. John Wiley & Sons, New York. 1999.
- Melchers, R.E. The effect of corrosion on the structural reliability of steel offshore structures, *Corrosion Science*, 47(10): 2391-2410. 2004.
- Mendenhall, W. and Reinmuth, J. *Statistics for Management and Economics*, Duxbury Press. 1982.
- Méndez, F.J., Menéndez, M., Luceño, A. and Losada, I.J. Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model, *Journal of Geophysical Research*, 111(C7): 8-13. 2006.
- Menéndez, M., Méndez, F.J., Izaguirre, C., Luceño, A. and Losada, I.J. The influence of seasonality on estimating return values of significant wave height, *Coastal Engineering*, 56(3): 211-219. 2009.
- Menéndez, M., Méndez, F.J., Losada, I.J. and Graham, N.E. Variability of extreme wave heights in northeast Pacific Ocean based on buoy measurements, *Geophysical Research Letters*, 35 L22607, doi: 10.1029/2008GL035394. 2008.
- Michele, C.D., Salvadori, G., Passoni, G. and Vezzoli, R. A multivariate model of sea storms using copulas, *Coastal Engineering*, 54: 734–751. 2007.
- Moan, T. Safety of offshore structures. Report No. 4, National University of Singapore. 2005.
- Moharram, S.H., Gosain, A.K. and Kapoor, P.N. A comparative study for the estimators of the generalized Pareto distribution, *Journal of Hydrology*, 150: 169-185. 1993.
- Möller, B. and Beer, M. Engineering Computation Under Uncertainty – Capabilities of Non-traditional Models, Special Issue on Uncertainties in Structural Analysis, *Computers and Structures*, 86(10): 1024–1041. 2008.
- Möller, B., Graf, W. and Beer, M. Fuzzy structural analysis using α -level optimization, *Computational Mechanics*, 26(6), 547–565. 2000.
- Moore, R. E. *Interval Analysis*. Prentice Hall, Inc, Englewood Cliffs, N.J. 1966.
- Morgan, E.C., Lackner, M. and Vogel, R. M. Probability distributions for offshore wind speeds, *Energy Conversion and Management*, 52(1): 15-26. 2011.

- Mori, Y. and Ellingwood, B. Maintaining reliability of concrete structures I: role of inspection repair, *ASCE Journal Structural Engineering*, 120(3): 824–5. 1994.
- Moriarty, P. J., Holley, W. E. and Butterfield., S. P. Extrapolation of Extreme and Fatigue Loads using Probabilistic Methods, Technical Report NREL/TP-500-34421, National Renewable Energy Laboratory, Golden, CO. 2006.
- Morton, I.D., Bowers, J., and Mould, G. Estimating return period wave heights and wind speeds using a seasonal point process model, *Coastal Engineering*, 31: 305-326. 1997.
- Muir, J.R., and El Shaarawi, A.H. On the calculation of extreme wave heights, a review, *Ocean Engineering*, 13: 93–118. 1986.
- Muraleedharan, G., Rao, A.D., Kurup, P.G., Nair, N.U. and Sinha, M. Modified Weibull distribution for maximum and significant wave height simulation and prediction, *Coastal Engineering*, 54(8): 630-638. 2007.
- Muraleedharan, G., Lucas, C. and Guedes Soares, C. Modelling significant wave height distributions with quantile functions for estimation of extreme wave heights, *Ocean Engineering*, 54: 119-131. 2012.
- Myers, R.H. Response surface methodology-current status and future directions, *Journal of Quality Technology*, 31: 30-44. 1999.
- Na, S.J., Do, K.D. and Suh, K.D. Forecast of wave run-up on coastal structure using offshore wave forecast data, *Coastal Engineering*, doi:10.1016/j.coastaleng.2011.03.012. 2011.
- Naess, A. and Clausen, P.H. Combination of the peaks-over-threshold and bootstrapping methods for extreme value prediction, *Structural Safety*, 23(4): 315-330. 2001.
- Naess, A. and Gaidai, O. Monte Carlo methods for estimating the extreme response of dynamical systems. *Journal of Engineering Mechanics-ASCE*, 134(8): 628-636. 2008.
- Naess, A. and Haug, E. Extreme Value Statistics of Wind Speed Data by the POT and ACER Methods, *Journal of Offshore Mechanics and Arctic Engineering*, 132(4): 041604. 2010.
- Naess, A. Estimation of long return period design values for wind speeds. *Journal of Engineering Mechanics-ASCE*, 124(3): 252-259. 1998.
- Naess, A., Gaidai, O. and Haver, S. Efficient Estimation of Extreme Response of Drag-dominated Offshore Structures by Monte Carlo Simulation, *Ocean Engineering*, 34(16): 2188-2197. 2007.

- Naess, A., Gaidai, O. and Batsevych, O. Extreme value statistics of combined load effect processes, *Structural Safety*, 31(4): 298-305. 2009.
- Naess, A., Gaidai, O. and Teigen, P. S. Extreme response prediction for nonlinear floating offshore structures by Monte Carlo simulation, *Applied Ocean Research*, 29(4): 221-230. 2007.
- Nasekhian, A. and Schweiger, H.F. Random Set Finite Element Method Application to Tunnelling, In: Beer, M., Muhanna, R.L. and Mullen, R.L. (eds) *Proceedings of the 4th International Workshop on Reliable Engineering Computing (REC2010), Robust Design – Coping with Hazards, Risk and Uncertainty*, Research Publishing, Singapore, 369–385. 2010.
- Nataf, A. Determination des distributions de probabilités dont les marges sont données, *Comptes Rendus de l'Académie des Sciences* 225: 42-43. 1962.
- Nelson, R.B. *An introduction to Copulas*. Springer, New York. 2006.
- Nguyen, H.T. *An Introduction to Random Sets*. Chapman and Hall/CRC Press, Boca Raton, Florida. 2006.
- NIST/SEMATECH, e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>. 2012.
- Nolte, K.G. Statistical methods for determining extreme sea states, proceedings of the second International conference on port and ocean engineering under arctic conditions, University of Iceland: 705–742. 1971.
- Northrop, P. and Jonathan, P. Threshold modeling of spatially-dependent non-stationary extremes with application to hurricane-induced wave heights, *Environmetrics*, 22: 799–809. 2011.
- Norton, E. Investigation into IEC Offshore Draft Standard Design Load Case 1.1. RECOFF Document 59, Recommendations for Design of Offshore Wind Turbines, Denmark. 2004.
- NPD. Acts, regulations and provisions for the petroleum activities. The Norwegian Petroleum Directorate. 1996.
- Oberguggenberger, M. and Fellin, W. Reliability bounds through random sets: nonparametric methods and geotechnical applications, *Computers & Structures*, 86: 1093–1101. 2008.
- Oberkampf, W.L., Helton, J.C. and Sentz, K. Mathematical representation of uncertainty. In: *AIAA Non-deterministic approaches forum*, no.AIAA 2001–1645, Seattle, WA. 2001.

- Obeyskera, J., Park, J. and Irizarry-Ortiz, M. Probabilistic Projection of Mean Sea Level and Coastal Extremes, *Journal of Waterway Port Coastal and Ocean Engineering-ASCE*, 139(2): 135-141. 2013.
- Ochi, K. New approach for estimating the severest sea state from statistical data, *Coastal Engineering Conference*, 512–525. 1992.
- Ochi, M. New approach for estimating the severest sea state from statistical data, *Coastal Engineering Proceedings*, 1(23): doi:10.9753/icce.v23. 2011.
- Oehlert, G.W. A Note on the Delta Method, *The American Statistician*, 46(1): 27-29. 1992.
- OGP. Risk assessment data directory, Reports No. 434-17. 2010.
- Oo, K.M. The design of semi-submersibles for minimum vertical motion, Ph.D thesis. 1974.
- Oil Rig Disasters (ORD). Usumacinta Jack-up Fire, May 2013 Retrieved from: <http://www.oilrigdisasters.co.uk/>.
- Palutikof, J.P., Brabson, B.B. and Lister, D.H. A review of methods to calculate extreme wind speeds, *Meteorological Applications*, 6(2): 119-132. 1999.
- Papadimitriou, C., Beck, J.L. and Katafygiotis, L.S. Asymptotic expansions for reliabilities and moments of uncertain dynamic systems, *Journal of Engineering Mechanics, ASCE*, 123(12): 1219-1229. 1997.
- Parey, S., Malek, F., Laurent, C. and Dacunha-Castelle, D. Trends and climate evolution: statistical approach for very high temperatures in France, *Climate Change*, 81: 331–352. 2007.
- Parrott, L.J. Factors influencing relative humidity in concrete, *Mag Concr Res*, 43(154): 45-52. 1991.
- Patton, A. Applications of copula theory in financial econometrics, *J. Appl. Econometrics*, 21: 147–173. 2002.
- Peng, L. and Welsh, A. Robust Estimation of the Generalized Pareto Distribution, *Extremes*, 4:53–65. 2001.
- Petrov, V., Guedes Soares, C. and Gotovac, H. Prediction of extreme significant wave heights using maximum entropy, *Coastal Engineering*, 74: 1-10. 2013.
- Phoon, K.K. and Kulhawy, F.H. Characterization of geotechnical variability, *Canadian Geotechnical Journal*, 36: 612–624. 1999.

- Pickands, J. Statistical inference using extreme order statistics, *Annals of Statistics*, 3(1): 119-131. 1975.
- Pradlwarter, H.J. and Schueller, G.I. Uncertain linear structural systems in dynamics: Efficient stochastic reliability assessment, *Computers and Structures*, 88(1-2): 74-86. 2010.
- Prince-Wright, R. Maximum likelihood models of joint environmental data for TLP design, *Proceedings of the 14th International Conference on Offshore Mechanics and Arctic Engineering*. 1995.
- Pugh, D.T. Estimating extreme currents by combining tidal and surge probabilities, *Ocean Engineering*, 9(4): 361-372. 1982.
- Quek, S.T. and Cheong, H.F. Prediction of extreme 3-sec gusts accounting for seasonal effects, *Structural Safety*, 11(2): 121-129. 1992.
- Rahman, M. Application of Fourier transforms to generalized functions. WIT press, Southampton, UK. 2011.
- Ramirez, P. and Carta, J.A. Influence of the data sampling interval in the estimation of the parameters of the Weibull wind speed probability density distribution: a case study, *Energy Convers Manage*, 46 (15–16): 2419–2438. 2005.
- Ramsay, T. Spline smoothing over difficult regions. *Journal of Royal Statistics Society Series B*, 64: 307–319. 2002.
- Randell, D., Wu, Y., Jonathan, P. and Ewans, K. Modeling covariate effects in extremes of storm severity on the Australian north west shelf. In *Proceedings of the ASME 32nd International Conference on Ocean, Offshore and Arctic Engineering*. Nantes, France, 9-14 June, 2013.
- Reiss, R.D. and Thomas, M. *Statistical Analysis of Extreme Values*. Birkhauser Verlag, Basel, Switzerland. 2001.
- Renard, B. and Lang, M. Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology, *Advances in Water Resources*, 30: 897 – 912. 2007.
- Repko, A., Van Gelder, P.H.A.J.M., Voortman, H.G. and Vrijling, J.K. Bivariate description of offshore wave conditions with physics-based extreme value statistics, *Applied Ocean Research*, 26(3-4): 162-170. 2004.
- Ribatet, M. POT: Generalized Pareto Distribution and Peaks Over Threshold. R package version 1.1-0. 2009.

- Ribatet, M. SpatialExtremes: Modelling Spatial Extremes. R package version 1.7-0. 2010.
- Ribereau, P., Naveau, P. and Guillou, A. A note of caution when interpreting parameters of the distribution of excesses, *Advances in Water Resources*, 34(10): 1215-1221. 2011.
- Robinson, M.E. and Tawn, J.A. Statistics for extreme sea currents, *Applied Statistics-Journal of the Royal Statistical Society Series C*, 46(2): 183-205. 1997.
- Roy, S. Foundation design philosophy for process equipment, May 2013 Retrieved from: <http://www.civildesignhelp.info/offshore.html>.
- Rubio, E., Hall, J.W. and Anderson, M.G. Uncertainty analysis in a slope hydrology and stability model using probabilistic and imprecise information, *Computers and Geotechnics*, 31: 529–536. 2004.
- Rue, H. and Held, L. *Gaussian Markov Random Fields*. Chapman and Hall, London. 2005.
- Ruggiero, P., Komar, P.D. and Allan, J.C. Increasing wave heights and extreme value projections: The wave climate of the U.S. Pacific Northwest, *Coastal Engineering*, 57(5): 539-552. 2010.
- Ruppert, D., Wand, M.P. and Carroll, R.J. *Semiparametric Regression*. Cambridge University Press, UK. 2003.
- Sagrilo, L.V.S., Naess, A. and Doria, A.S. On the long-term response of marine structures, *Applied Ocean Research*, 33(3): 208-214. 2011.
- Salvadori, G. and De Michele, C. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resource Research*, 40: W12511, doi:10.1029/2004WR003133. 2004.
- Salvadori, G., De Michele, C., Kottegoda, N.T. and Rosso, R. An approach using copulas, In *Extremes in nature*, Dordrecht: Springer. 2007.
- Sanchez-Silva, M., Klutke, G.A. and Rosowsky, D.V. Life-cycle performance of structures subject to multiple deterioration mechanisms, *Structural Safety*, doi:10.1016/j.strusafe.2011.03.003. 2011
- Santos, R.S. and Feijo, L.P. Safety Challenges Associated With Deepwater Concepts Utilized in the Offshore Industry, *Mine Safety*. 2010.
- Saranyasontorn, K. and Manuel, L. Design Loads for Wind Turbines using the Environmental Contour Method, *Journal of Solar Energy Engineering, Transactions of the ASME*, 128(4):554-561. 2006.

- Saranyasoontorn, K. and Manuel, L. On Assessing the Accuracy of Offshore Wind Turbine Reliability-based Design Loads from the Environmental Contour Method, *International Journal of Offshore and Polar Engineering*, 15(2):132-140. 2005.
- Sarpkaya, T. and Isaacson, M. *Mechanics of Wave Forces on Offshore Structures*. Van Nostrand Reinhold, New York. 1981.
- Sauvaget, P., David, E. and Guedes Soares, C. Modelling Tidal Currents on the Coast of Portugal, *Coastal Engineering*, 40(4): 393-409. 2000.
- Schuëller, G.I. Computational Stochastic Mechanics - Recent Advances, *Computers and Structures*, 79: 2225-2234. 2001.
- Segers, J. Generalized Pickands estimators for the extreme value index, *Journal of Statistical Planning and Inference*, 128(2): 381-396. 2005.
- Shafer, G. *A mathematical theory of evidence*. Princeton University Press, Princeton. 1976.
- Shahid, M. and Das, P.K. Structural system reliability using FEM and reliability processors, In Guedes Soares, C. and Das, P.K. (eds): *Proceedings of the 1st conference on marine structures*, Taylor & Francis, Glasgow, United Kingdom, 547-556. 2007.
- Shariff, A., Hafezi, A. and Hadi, M. Modelling significant wave height data of north sea: Rayleigh vs Weibull distribution, In *Proceedings of International Conference on Mechatronics and Applied Mechanics (ICMAM 2011)*. Hong Kong, China, Dec 27-28, 2011.
- Shinozuka, M. *Stochastic Fields and their digital Simulation*, *Stochastic methods in structural dynamics*: 93-133. 1987.
- Silva-González, F., Heredia-Zavoni, E. and Montes-Iturrizaga, R. Development of environmental contours using Nataf distribution model, *Ocean Engineering*, 58: 27-34. 2013.
- Sklar, A. Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 8: 229–231. 1959.
- Smith, R.L. *Environmental statistics*, technical report. Dep. of Stat., Univ. of North Carolina, Chapel Hill. 2001.
- Spanos, P.D. and Deodatis, G. *Computational Stochastic Mechanics*. Millpress, Rotterdam. 2007.
- Stephenson, A.G. evd: Extreme Value Distributions. *R News*, 2(2):31-32. 2002.

- Stephenson, A.G. and Ribatet, M. evdbayes: Bayesian Analysis in Extreme Value Theory. R package version 1.0-8. 2010a.
- Stephenson, A.G. ismev: An Introduction to Statistical Modeling of Extreme Values. R package version 1.35. 2010b.
- Straub, D. and Faber, M.H. Risk based inspection planning for structural systems, *Structural Safety*, 27(4): 335-355. 2005.
- Sudati Sagrilo, L.V., Papaleo, A. and de Lima, E.C.P. A joint probability model for environmental parameters, *Journal of Offshore Mechanics and Arctic Engineering*, 133(3): 031605. 2011.
- Tao, Z., Li, H. and Liu, D. Risk assessment method for offshore structure based on global sensitivity analysis, *Modelling and Simulation in Engineering art. no. 671934*. 2012.
- Teena, N.V., Sanil Kumar, V., Sudheesh, K. and Sajeev, R. Statistical analysis on extreme wave height, *Natural Hazards*, 64(1): 223-236. 2012.
- Toth B., Lillo, F. and Farmer, J.D. Segmentation algorithm for non-stationary compound Poisson processes, *European Physical Journal B*, 78(2): 235-243. 2010.
- Tucker, T. and Ferson, S. *Probability Bounds Analysis in Environmental Risk Assessments*, Applied Biomathematics. Setauket, New York. 2003.
- US Army Corps of Engineers, *Wave Information Studies (WIS) Project Documentation*, Coastal and Hydraulics Laboratory, Dec, 2010.
- USFOS Course Manual. Workshop II, USFOS jacket pushover. MARINTEK SINTEF Group. 2001.
- USFOS. *Ultimate Strength of Frame Offshore Structures, User's Manual*, SINTEF, 2007.
- Vandenbergh, S., Verhoest, N.E.C., Onof, C. and De Baets, B. A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall, *Water Resources Research*, 47: W07529 DOI: 10.1029/2009WR008388. 2011.
- Vanem, E. and Bitner-Gregersen, E. M. Stochastic modelling of long-term trends in the wave climate and its potential impact on ship structural loads. *Applied Ocean Research* 37: 235-248. 2012.
- Vanem, E. Long-term time-dependent stochastic modelling of extreme waves, *Stochastic Environmental Research on Risk Assessment*, 25:185–209. 2011.

- Vikebó, F., Furevik, T., Furnes, G., Kvamstø, N.G. and Reistad, M. Wave height variations in the North Sea and on the Norwegian Continental Shelf, 1881-1999, *Continental Shelf Research*, 23: 251-263. 2003.
- Walley, P. *Statistical reasoning with imprecise probabilities*. Chapman & Hall, London. 1991.
- Wang, X.L., Zwiers, F.W. and Swail, V.R. North Atlantic ocean wave climate change scenarios for the twenty-first century, *Journal of Climate*, 17: 2368-2383. 2003.
- Webb, C.S. and Zank, H. Accounting for optimism and pessimism in expected utility, *Journal of Mathematical Economics*, 47(6): 706-717. 2011.
- Williamson, R.C. and Downs, T. Probabilistic Arithmetic I: numerical methods for calculating convolutions and dependency bounds, *International Journal of Approximate Reasoning*, 4: 89-158. 1990.
- Winterstein, S.R., Jha, A.K. and Kumar, S. Reliability of floating structures: Extreme response and load factor design, *Journal of Waterway Port Coastal and Ocean Engineering-ASCE*, 125(4): 163-169. 1999.
- Winterstein, S.R., Kleiven, G. and Hagen, Ø. Comparing extreme wave estimates from hourly and annual data. *Proceedings ISOPE. Conference*. 2001.
- Wist, H.T., Myrhaug, D. and Rue, H. Statistical properties of successive wave heights and successive wave periods, *Applied Ocean Research*, 26(3-4): 114-136. 2004.
- Yager, R.R., Fedrizzi, M. and Kacprzyk, J. *Advances in the Dempster-Shafer Theory of Evidence*. John Wiley & Sons. 1994.
- Yamazaki, F. and Shinozuka, M. Digital generation of non-Gaussian stochastic fields, *Journal of Engineering Mechanics-ASCE*, 114(7): 1183-1197. 1988.
- Yan, J. Enjoy the Joy of Copulas: With a Package copula, *Journal of Statistical Software*, 21(4): 1-21. 2007.
- Yang, T.Y. and Kuo, L. Bayesian binary segmentation procedure for a Poisson process with multiple change points, *Journal of Computational and Graphical Statistics*, 10: 772-785. 2011.
- Zachary, S., Feld, G., Ward, G. and Wolfram, J. Multivariate extrapolation in the offshore environment, *Applied Ocean Research*, 20(5): 273-295. 1998.
- Zadeh, L.A. Fuzzy sets, *Information and Control*, 8, 338-353. 1965.

- Zayed, A., Garbatov, Y. and Guedes Soares, C. Time variant reliability assessment of ship structures with fast integration techniques, *Probabilistic Engineering Mechanics*, 32: 93-102. 2013.
- Zhang, J. Improving on Estimation for the Generalized Pareto Distribution, *Technometrics*, 52(3): 335-339. 2010.
- Zhang, J. Likelihood Moment Estimation for the Generalized Pareto Distribution, *Australian and New Zealand Journal of Statistics*, 49(1): 69–77. 2007.
- Zhang, J. and Stephens, M.A. A new and efficient estimation method for the generalized Pareto distribution, *Technometrics*, 51(3): 316-325. 2007.
- Zhang, L. and Singh, V.P. Frequency analysis of flood damage, *Journal of Hydrologic Engineering*, ASCE, 10(2): 100-109. 2005.
- Zhang, M.Q., Beer, M., Quek, S.T. and Choo, Y.S. Comparison of uncertainty models in reliability analysis of offshore structures under marine corrosion, *Structural Safety*, 32: 425-432. 2010.
- Zhang, M.Q., Beer, M. and Koh, C. Interval Analysis for System Identification of Linear MDOF Structures in the Presence of Modeling Errors, *Journal of Engineering Mechanics*, 138(11): 1326-38. 2012.
- Zhao, M. and Gu, M. Extreme wind pressure estimation based on the r largest order statistics model, *Chinese Journal of Theoretical and Applied Mechanics*, 42 (6): 1074-1082. 2010.
- Zhou, L., Riska, K., Polach, R.V.B.U., Moan, T. and Su, B. Experiments on level ice loading on an icebreaking tanker with different ice drift angles, *Cold Regions Science and Technology*, 85: 79-93. 2013a.
- Zhou, L., Riska, K., Moan, T. and Su, B. Numerical modeling of ice loads on an icebreaking tanker: Comparing simulations with model tests, *Cold Regions Science and Technology*, 87: 33-46. 2013b.
- Zimmermann, H-J. Fuzzy set theory and its applications. Kluwer Academic Publishers, Boston/London. 1992.

Appendix A. Detailed Information of Four Discretization Steps

The detailed information of the discretization steps for the bivariate copula (H_S, T_P) are given herein. Figures A.1-A.4 show the subdomains of the copula which are discretized in each step. The structural analysis results for each of the discretized subcopulas are recorded in Table A.1.

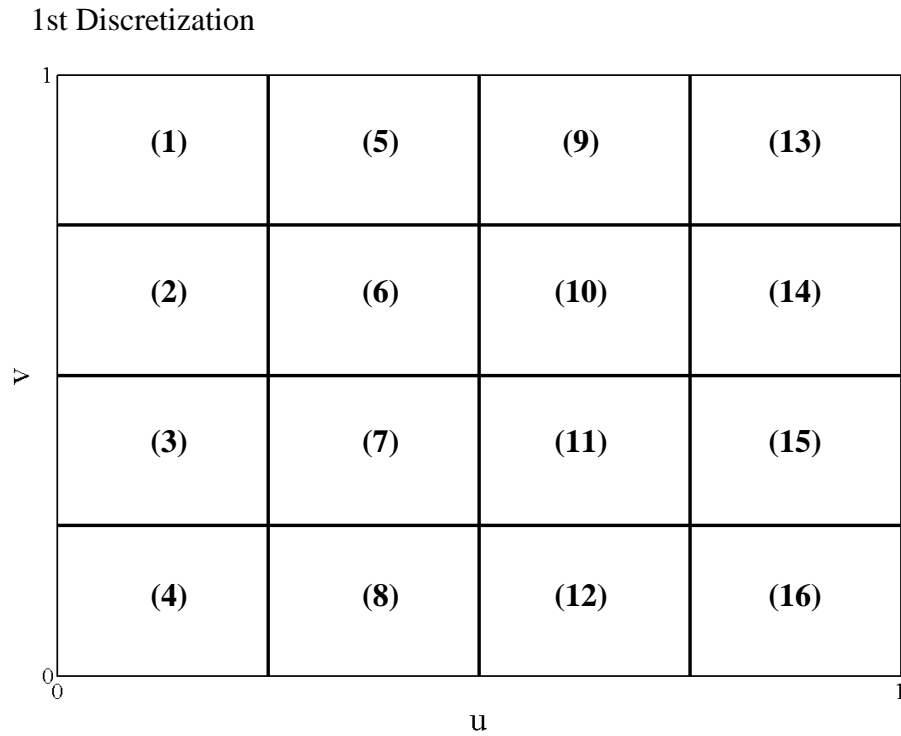


Figure A.1 Numbering of subcopulas in first discretization ($u=F_1(H_S)$, $v=F_2(T_P)$).

2nd Discretization

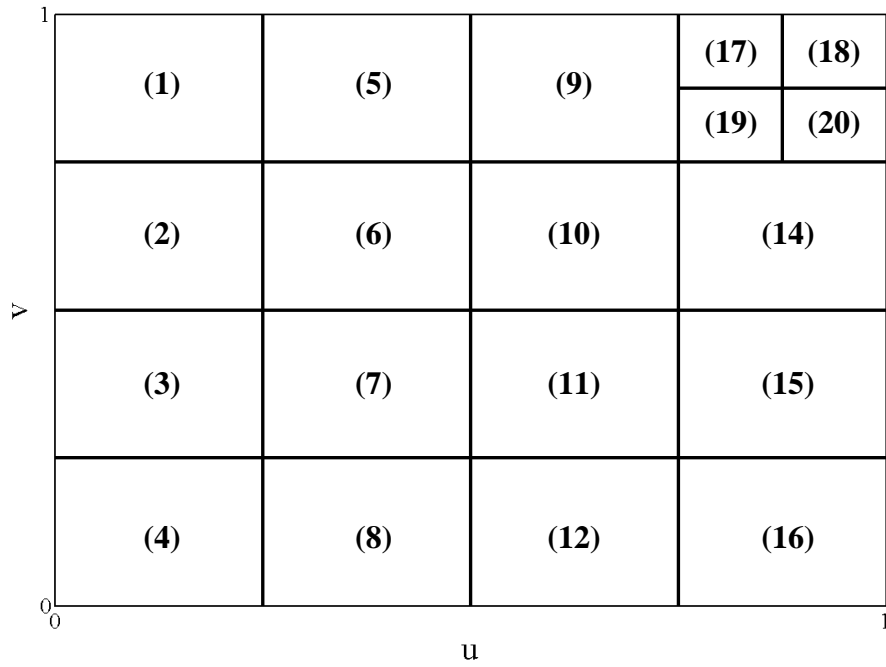


Figure A.2 Numbering of subcopulas in second discretization ($u=F_1(H_S)$, $v=F_2(T_P)$).

3rd Discretization

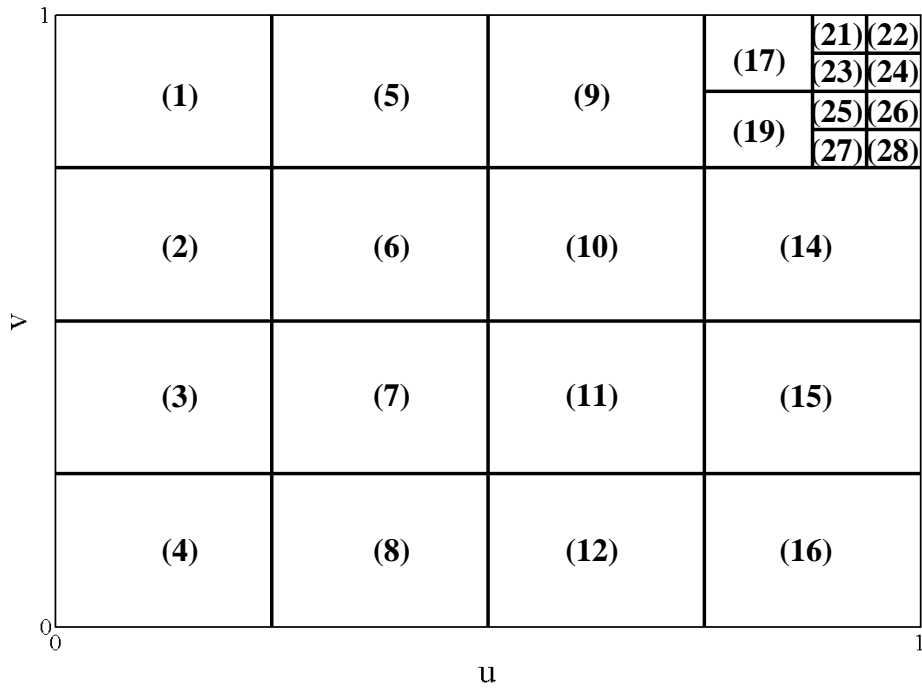


Figure A.3 Numbering of subcopulas in third discretization ($u=F_1(H_S)$, $v=F_2(T_P)$).

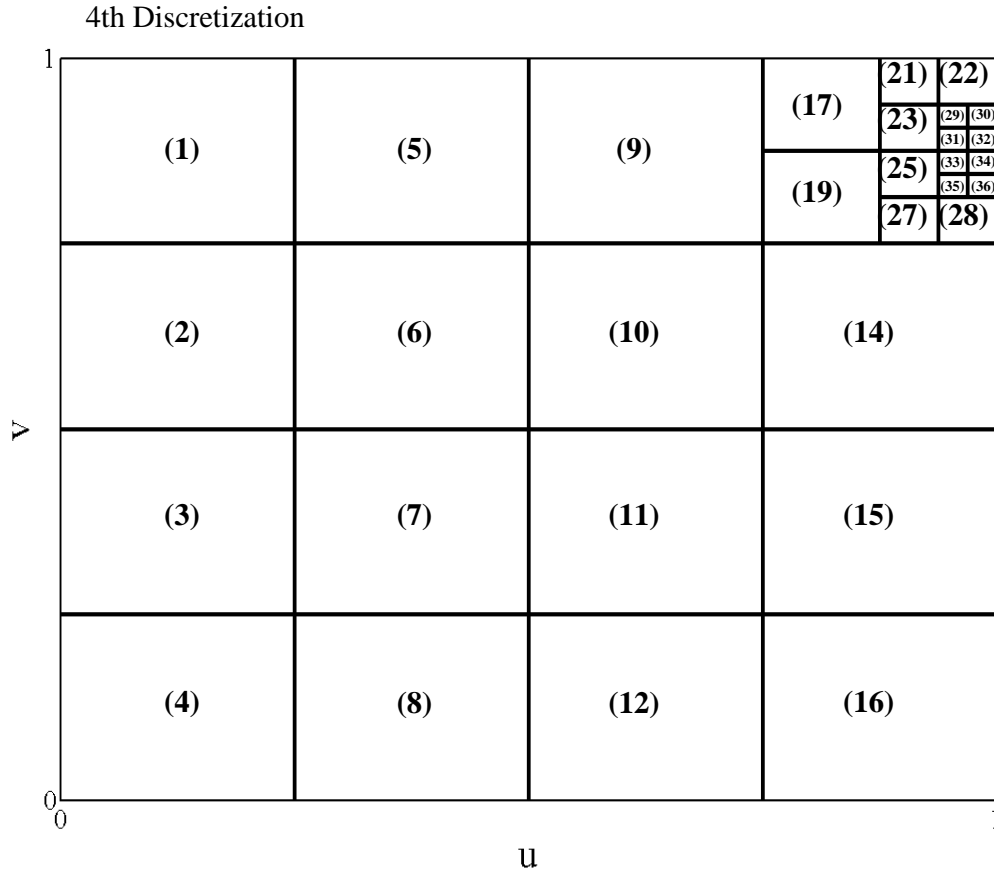


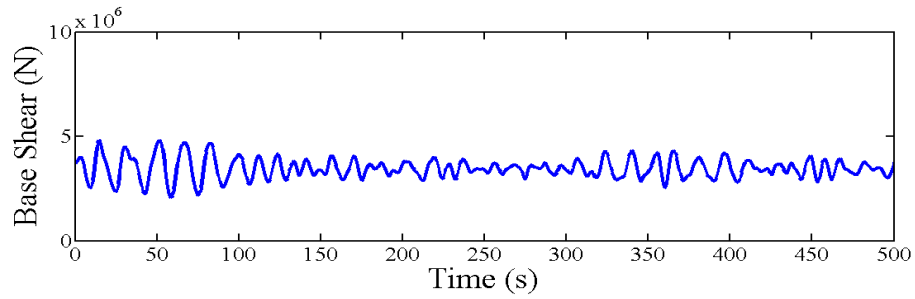
Figure A.4 Numbering of subcopulas in fourth discretization ($u=F_1(H_S)$, $v=F_2(T_P)$).

Comparison of base shear response under different sea states is shown in Fig. A.5 which demonstrates the importance of critical sea states in the multivariate analysis. From the results in Table A.1, it can be seen that the significance of H_S is relatively larger than T_p since large magnitude in the response is associated with large wave height for this particular selected example. Through the discretizations, the minimum ΔC of the blocks in each step is decreased from 0.0205 to 0.0017 and maximum ΔP_E of the blocks in each step is decreased from 0.0081 to 0.0014. This implies that the computation of Eq. (2.12)

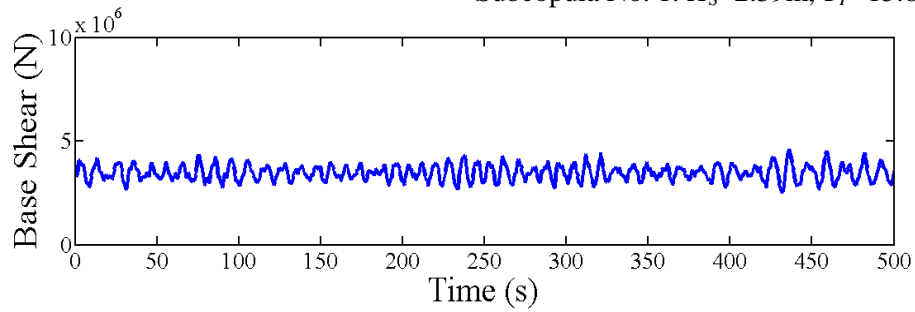
now relies on finer and finer subdomains of the copula through the discretization steps.

Table A.1 Results of structural analysis in each subcopula.

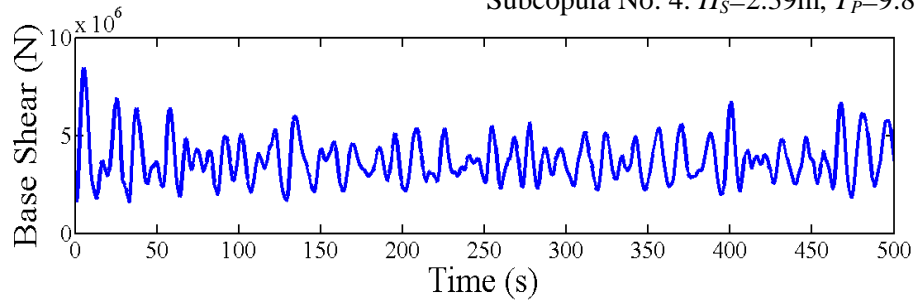
No	H_S (m) T_P (s)		1 st Discretization		2 nd Discretization		3 rd Discretization		4 th Discretization	
			Long Term Load 8937629 N		Long Term Load 9886673 N		Long Term Load 10806707 N		Long Term Load 10976878 N	
			ΔP_E	ΔC	ΔP_E	ΔC	ΔP_E	ΔC	ΔP_E	ΔC
1	2.39	15.62	0.0000	0.0205	0.0000	0.0205	0.0000	0.0205	0.0000	0.0205
2	2.39	13.32	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441
3	2.39	11.72	0.0000	0.0695	0.0000	0.0695	0.0000	0.0695	0.0000	0.0695
4	2.39	9.83	0.0000	0.1158	0.0000	0.1158	0.0000	0.1158	0.0000	0.1158
5	3.34	15.62	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441
6	3.34	13.32	0.0000	0.0640	0.0000	0.0640	0.0000	0.0640	0.0000	0.0640
7	3.34	11.72	0.0000	0.0724	0.0000	0.0724	0.0000	0.0724	0.0000	0.0724
8	3.34	9.83	0.0000	0.0695	0.0000	0.0695	0.0000	0.0695	0.0000	0.0695
9	4.33	15.62	0.0003	0.0695	0.0001	0.0695	0.0001	0.0695	0.0000	0.0695
10	4.33	13.32	0.0000	0.0724	0.0000	0.0724	0.0000	0.0724	0.0000	0.0724
11	4.33	11.72	0.0009	0.0640	0.0005	0.0640	0.0005	0.0640	0.0003	0.0640
12	4.33	9.83	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441
13	6.06	15.62	0.0081	0.1158	-	-	-	-	-	-
14	6.06	13.32	0.0004	0.0695	0.0000	0.0695	0.0000	0.0695	0.0000	0.0695
15	6.06	11.72	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441	0.0000	0.0441
16	6.06	9.83	0.0002	0.0205	0.0000	0.0205	0.0000	0.0205	0.0000	0.0205
17	5.45	16.76			0.0000	0.0255	0.0000	0.0255	0.0000	0.0255
18	7.08	16.76			0.0054	0.0420	-	-	-	-
19	5.48	14.86			0.0000	0.0227	0.0000	0.0227	0.0000	0.0227
20	7.08	14.86			0.0039	0.0255	-	-	-	-
21	6.49	17.78					0.0004	0.0094	0.0002	0.0094
22	8.09	17.78					0.0004	0.0154	0.0004	0.0154
23	6.49	16.11					0.0000	0.0077	0.0000	0.0077
24	8.09	16.11					0.0044	0.0094	-	-
25	6.49	15.21					0.0000	0.0067	0.0000	0.0067
26	8.09	15.21					0.0031	0.0072	-	-
27	6.49	14.55					0.0005	0.0059	0.0004	0.0059
28	8.09	14.55					0.0000	0.0058	0.0000	0.0058
29	7.50	16.41							0.0009	0.0023
30	9.11	16.41							0.0012	0.0027
31	7.50	15.85							0.0000	0.0021
32	9.11	15.85							0.0014	0.0023
33	7.50	15.41							0.0000	0.0019
34	9.11	15.41							0.0014	0.0019
35	7.50	15.03							0.0000	0.0017
36	9.11	15.03							0.0013	0.0017



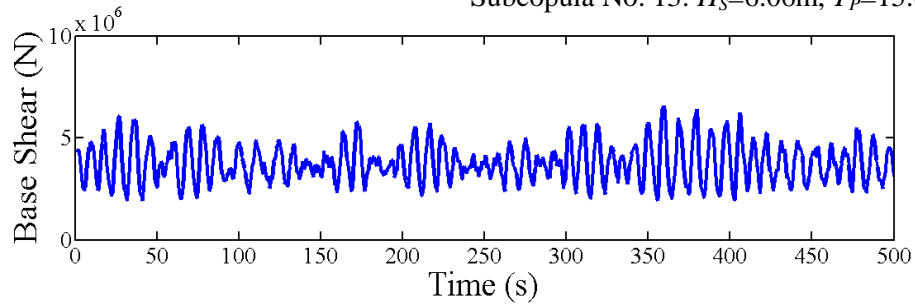
Subcopula No. 1: $H_S=2.39\text{m}$, $T_P=15.62\text{s}$



Subcopula No. 4: $H_S=2.39\text{m}$, $T_P=9.83\text{s}$



Subcopula No. 13: $H_S=6.06\text{m}$, $T_P=15.62\text{s}$



Subcopula No. 16: $H_S=6.06\text{m}$, $T_P=9.83\text{s}$

Figure A.5 Comparison of structural base shear for different sea states.

Appendix B. Uncertainty Assessment in POT Method

To study the effects of tail behavior, noise and dependency in the time series on the POT method, numerical simulations are performed. The overall results were plotted in Section 3.3. First, the performance of each estimator is tested via Monte Carlo simulation with different data sample sizes, namely, $n = 10, 20, 30, 50, 80, 100, 150$ and 200 . In studying the effects of tail behavior, the shape parameter are examined for values of $\zeta = -0.5, -0.25, 0, 0.25$ and 0.5 , with the threshold and scale parameter fixed at $u = 1$ and $\sigma = 2$, respectively. For each value of ζ and n , the results are calculated based on the mean of 10,000 simulations to ensure the conclusions made are meaningful. The parameters studied include the estimated shape parameter, scale parameter and the 99-percentile values. Detailed results in terms of the bias percentage are presented in Tables B.1 to B.5. The abbreviations in the table stand for different estimators. MOM: method of moments; MLE: maximum likelihood method; PWMU: unbiased probability weighted moments method ($p_{i:n}=(i-0.5)/n$ in Eq. 3.21); PWMB: biased probability weighted moments method ($p_{i:n}=(i-0.35)/n$ in Eq. 3.21); AD: A-D test based goodness-of-fit method; KS: K-S test based goodness-of-fit method.

Table B.1 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=-0.5, \sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	16.39%	34.52%	-0.22%	MOM	7.78%	14.06%	0.02%
MLE	33.91%	76.42%	-11.43%	MLE	21.19%	46.09%	-7.28%
PWMU	9.25%	13.62%	9.05%	PWMU	4.84%	5.00%	4.41%
PWMB	12.32%	23.60%	3.40%	PWMB	6.66%	10.80%	1.98%
AD	13.70%	23.00%	18.28%	AD	4.92%	7.39%	2.87%
KS	4.60%	-14.30%	64.80%	KS	4.14%	-5.59%	19.14%
n=30				n=50			
MOM	4.40%	7.66%	-0.32%	MOM	2.53%	3.71%	-0.41%
MLE	15.51%	30.68%	-5.24%	MLE	8.62%	17.34%	-3.78%
PWMU	2.56%	2.24%	2.38%	PWMU	1.44%	0.21%	1.18%
PWMB	3.85%	6.23%	0.89%	PWMB	2.25%	2.66%	0.31%
AD	1.88%	1.38%	3.21%	AD	0.87%	-1.56%	1.17%
KS	3.05%	-5.57%	12.90%	KS	1.61%	-5.98%	7.89%
n=80				n=100			
MOM	1.79%	2.63%	-0.07%	MOM	1.55%	2.48%	0.01%
MLE	5.58%	12.08%	-2.53%	MLE	4.16%	9.65%	-2.05%
PWMU	1.06%	0.45%	0.90%	PWMU	1.11%	0.89%	0.88%
PWMB	1.57%	2.00%	0.37%	PWMB	1.51%	2.12%	0.46%
AD	0.40%	-1.29%	0.92%	AD	0.11%	-1.29%	0.86%
KS	0.90%	-1.94%	5.67%	KS	1.01%	-2.33%	4.11%
n=150				n=200			
MOM	1.99%	0.43%	0.34%	MOM	0.36%	0.63%	0.02%
MLE	3.76%	5.89%	-1.15%	MLE	2.13%	4.74%	-1.12%
PWMU	1.70%	-0.80%	0.94%	PWMU	0.10%	-0.42%	0.46%
PWMB	1.98%	0.03%	0.66%	PWMB	0.31%	0.20%	0.25%
AD	0.76%	-1.88%	1.06%	AD	-0.44%	-1.33%	0.53%
KS	1.73%	-2.25%	3.05%	KS	-0.02%	-2.22%	1.94%

Table B.2 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=-0.25, \sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	15.27%	66.85%	-4.41%	MOM	6.82%	30.08%	-1.95%
MLE	42.82%	171.40%	-12.47%	MLE	22.24%	81.30%	-8.07%
PWMU	6.85%	23.52%	8.17%	PWMU	3.01%	11.84%	4.53%
PWMB	12.00%	49.77%	0.57%	PWMB	5.81%	25.87%	0.90%
AD	14.14%	42.77%	35.59%	AD	5.41%	7.85%	9.00%
KS	15.85%	-13.89%	108.20%	KS	5.39%	-4.29%	25.76%
n=30				n=50			
MOM	5.11%	17.65%	-1.83%	MOM	2.78%	11.03%	-1.45%
MLE	13.06%	52.27%	-6.81%	MLE	7.57%	29.03%	-4.64%
PWMU	2.98%	4.69%	2.24%	PWMU	1.37%	4.04%	0.98%
PWMB	4.87%	14.24%	-0.08%	PWMB	2.52%	9.81%	-0.37%
AD	2.72%	1.28%	4.80%	AD	0.93%	-1.40%	2.68%
KS	4.27%	-3.22%	18.14%	KS	1.92%	-5.58%	12.06%
n=80				n=100			
MOM	1.68%	9.36%	-0.85%	MOM	1.64%	7.83%	-0.74%
MLE	4.68%	20.31%	-3.10%	MLE	4.08%	16.15%	-2.64%
PWMU	0.68%	5.53%	0.70%	PWMU	0.90%	5.17%	0.62%
PWMB	1.41%	9.14%	-0.14%	PWMB	1.48%	8.06%	-0.05%
AD	0.42%	-0.19%	2.00%	AD	0.47%	0.25%	1.77%
KS	0.79%	-3.17%	8.70%	KS	0.86%	0.19%	7.58%
n=150				n=200			
MOM	1.12%	4.26%	-0.21%	MOM	0.62%	1.52%	-0.16%
MLE	2.53%	9.61%	-1.65%	MLE	1.76%	7.36%	-1.34%
PWMU	0.62%	2.63%	0.63%	PWMU	0.40%	-0.28%	0.49%
PWMB	1.00%	4.56%	0.19%	PWMB	0.69%	1.18%	0.15%
AD	0.31%	-1.29%	1.33%	AD	-0.15%	-2.84%	0.88%
KS	0.27%	-0.82%	5.83%	KS	-0.02%	-5.45%	4.33%

Table B.3 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=0, \sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	19.65%	-	-11.06%	MOM	11.15%	-	-7.02%
MLE	47.72%	-	46.72%	MLE	21.25%	-	-5.75%
PWMU	6.32%	-	6.42%	PWMU	3.66%	-	2.90%
PWMB	13.93%	-	-2.72%	PWMB	7.59%	-	-1.72%
AD	15.44%	-	80.84%	AD	6.53%	-	20.86%
KS	9.55%	-	212.21%	KS	6.70%	-	30.62%
n=30				n=50			
MOM	7.61%	-	-5.44%	MOM	3.89%	-	-3.16%
MLE	11.97%	-	-5.89%	MLE	5.82%	-	-3.56%
PWMU	2.88%	-	1.46%	PWMU	0.76%	-	1.34%
PWMB	5.52%	-	-1.62%	PWMB	2.35%	-	-0.54%
AD	3.65%	-	10.89%	AD	1.00%	-	7.00%
KS	6.23%	-	21.92%	KS	3.10%	-	12.94%
n=80				n=100			
MOM	2.47%	-	-2.06%	MOM	2.05%	-	-1.49%
MLE	3.31%	-	-2.33%	MLE	2.75%	-	-1.89%
PWMU	0.12%	-	0.68%	PWMU	0.34%	-	0.83%
PWMB	1.12%	-	-0.49%	PWMB	1.14%	-	-0.11%
AD	0.25%	-	4.39%	AD	0.27%	-	3.42%
KS	1.20%	-	6.17%	KS	1.14%	-	5.18%
n=150				n=200			
MOM	1.85%	-	-1.06%	MOM	1.37%	-	-0.30%
MLE	2.32%	-	-1.06%	MLE	1.71%	-	-0.35%
PWMU	0.75%	-	0.76%	PWMU	0.46%	-	1.22%
PWMB	1.28%	-	0.12%	PWMB	0.86%	-	0.74%
AD	0.72%	-	2.97%	AD	0.36%	-	2.94%
KS	1.40%	-	3.49%	KS	0.90%	-	4.36%

Table B.4 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=0.25, \sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	34.09%	-117.85%	-22.10%	MOM	21.50%	-74.00%	-15.51%
MLE	52.55%	-145.17%	38.50%	MLE	18.42%	-58.31%	70.54%
PWMU	10.36%	-49.17%	-1.75%	PWMU	4.82%	-26.82%	0.18%
PWMB	21.11%	-81.45%	-9.53%	PWMB	10.27%	-43.33%	-4.42%
AD	19.56%	-25.33%	171.32%	AD	6.35%	-1.86%	56.52%
KS	22.08%	-77.51%	242.08%	KS	14.77%	-58.19%	55.23%
n=30				n=50			
MOM	17.43%	-58.71%	-13.20%	MOM	12.37%	-38.29%	-10.28%
MLE	10.37%	-37.89%	1.99%	MLE	5.83%	-15.05%	1.35%
PWMU	3.68%	-21.51%	-0.42%	PWMU	2.07%	-8.11%	0.23%
PWMB	7.37%	-32.51%	-3.59%	PWMB	4.30%	-14.74%	-1.82%
AD	3.48%	-2.06%	26.45%	AD	2.14%	6.31%	16.75%
KS	10.84%	-49.11%	24.06%	KS	6.68%	-36.11%	14.84%
n=80				n=100			
MOM	9.40%	-29.87%	-7.30%	MOM	8.14%	-29.22%	-8.64%
MLE	3.43%	-11.27%	0.02%	MLE	2.97%	-11.72%	-2.51%
PWMU	1.40%	-6.66%	0.23%	PWMU	1.43%	-8.06%	-2.28%
PWMB	2.78%	-10.81%	-1.08%	PWMB	2.54%	-11.38%	-3.32%
AD	1.50%	1.51%	7.67%	AD	1.34%	0.68%	4.25%
KS	4.68%	-28.53%	5.40%	KS	4.24%	-25.03%	-0.71%
n=150				n=200			
MOM	6.52%	-21.54%	-6.22%	MOM	5.33%	-18.35%	-3.87%
MLE	2.07%	-6.97%	-0.81%	MLE	1.28%	-4.87%	0.90%
PWMU	1.06%	-4.62%	-0.69%	PWMU	0.49%	-3.75%	0.84%
PWMB	1.80%	-6.83%	-1.40%	PWMB	1.04%	-5.42%	0.29%
AD	1.59%	0.03%	3.57%	AD	1.12%	0.48%	3.89%
KS	3.43%	-19.20%	-0.55%	KS	2.34%	-14.83%	0.27%

Table B.5 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=0.5, \sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	66.32%	-87.64%	-31.29%	MOM	49.75%	-64.03%	-27.54%
MLE	52.77%	-71.08%	821.61%	MLE	16.95%	-27.03%	24.76%
PWMU	16.35%	-44.05%	-18.23%	PWMU	7.39%	-27.53%	-12.29%
PWMB	32.41%	-60.57%	-18.44%	PWMB	15.44%	-35.86%	-14.12%
AD	23.93%	-14.51%	370.10%	AD	11.58%	-2.82%	79.08%
KS	62.00%	-111.96%	102.37%	KS	50.71%	-108.77%	14.3%
n=30				n=50			
MOM	43.19%	-53.49%	-25.70%	MOM	38.62%	-43.62%	-22.45%
MLE	9.00%	-2.70%	13.05%	MLE	5.75%	-7.37%	5.54%
PWMU	5.42%	-20.62%	-10.68%	PWMU	5.17%	-13.33%	-9.14%
PWMB	10.85%	-26.16%	-12.42%	PWMB	8.44%	-16.65%	-10.16%
AD	4.59%	0.31%	45.66%	AD	3.41%	1.71%	23.43%
KS	42.97%	-99.27%	-6.42%	KS	37.32%	-79.31%	-8.24%
n=80				n=100			
MOM	35.13%	-36.96%	-18.71%	MOM	32.58%	-35.47%	-16.79%
MLE	3.18%	-4.28%	6.82%	MLE	2.87%	-3.78%	6.69%
PWMU	3.17%	-8.69%	-4.73%	PWMU	3.02%	-8.68%	-2.90%
PWMB	5.25%	-10.76%	-5.53%	PWMB	4.67%	-10.34%	-3.58%
AD	3.30%	-1.98%	16.98%	AD	2.87%	-2.50%	14.64%
KS	31.93%	-66.87%	-7.89%	KS	28.04%	-63.81%	-11.07%
n=150				n=200			
MOM	28.73%	-31.51%	-16.63%	MOM	27.22%	-29.89%	-14.94%
MLE	1.81%	-2.07%	3.19%	MLE	0.96%	-1.27%	1.28%
PWMU	2.24%	-6.05%	-3.31%	PWMU	1.58%	-5.39%	-2.89%
PWMB	3.33%	-7.15%	-3.81%	PWMB	2.40%	-6.21%	-3.25%
AD	1.08%	-7.18%	4.81%	AD	0.81%	-9.25%	-0.08%
KS	20.12%	-51.73%	-12.89%	KS	19.08%	-45.70%	-12.92%

The effect of noise is studied in the same manner with respect to the threshold u , shape parameter ζ and scale parameter σ . The noise, as described in Session 3.3, is assumed to be Gaussian with zero mean and coefficient of

variation of 0.1, 0.3 and 0.5. The procedures in assessing the relative bias follow the same way as described above and the results are summarized in Tables B.6 to B.14.

Table B.6 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1+N(0,0.1)$, $\zeta=-0.5$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	33.27%	59.41%	-1.88%	MOM	23.55%	36.38%	-2.09%
MLE	41.45%	83.50%	-11.82%	MLE	31.41%	53.43%	-5.97%
PWMU	26.33%	40.68%	4.68%	PWMU	21.74%	31.00%	0.35%
PWMB	28.72%	48.88%	0.17%	PWMB	23.36%	35.98%	-1.49%
AD	22.76%	35.95%	8.43%	AD	14.61%	15.04%	2.58%
KS	17.98%	10.29%	41.87%	KS	17.92%	20.91%	14.15%
n=30				n=50			
MOM	21.45%	32.92%	-2.06%	MOM	17.94%	28.33%	-2.07%
MLE	25.54%	37.58%	-3.68%	MLE	17.01%	24.89%	-1.92%
PWMU	21.10%	30.86%	-1.10%	PWMU	18.54%	29.39%	-1.94%
PWMB	22.23%	35.26%	-2.21%	PWMB	19.26%	31.50%	-2.58%
AD	11.67%	11.03%	1.84%	AD	9.18%	8.30%	1.80%
KS	17.11%	22.46%	9.97%	KS	15.60%	22.63%	3.78%
n=80				n=100			
MOM	17.16%	27.97%	-2.38%	MOM	16.90%	26.84%	-2.56%
MLE	14.04%	16.91%	-0.70%	MLE	12.45%	16.08%	-0.53%
PWMU	18.08%	30.86%	-2.67%	PWMU	17.99%	29.87%	-3.11%
PWMB	18.55%	32.18%	-3.06%	PWMB	18.36%	30.94%	-3.41%
AD	8.80%	7.24%	1.64%	AD	8.23%	7.10%	1.41%
KS	15.74%	23.07%	1.71%	KS	16.05%	23.96%	0.25%
n=150				n=200			
MOM	16.55%	26.41%	-2.83%	MOM	15.88%	24.70%	-2.54%
MLE	10.79%	12.45%	0.36%	MLE	9.67%	10.69%	0.53%
PWMU	17.88%	30.12%	-3.58%	PWMU	17.35%	28.55%	-3.39%
PWMB	18.13%	30.83%	-3.78%	PWMB	17.54%	29.09%	-3.54%
AD	8.19%	6.94%	1.33%	AD	7.64%	6.39%	1.39%
KS	16.79%	24.89%	-1.04%	KS	15.88%	25.21%	-0.95%

Table B.7 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1+N(0,0.3)$, $\zeta=-0.5$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	76.99%	133.14%	-5.26%	MOM	64.66%	101.35%	-5.30%
MLE	57.75%	97.24%	-8.17%	MLE	49.01%	67.34%	-3.22%
PWMU	66.45%	109.78%	-1.79%	PWMU	62.39%	97.81%	-4.63%
PWMB	66.68%	111.72%	-4.40%	PWMB	63.06%	100.08%	-5.63%
AD	45.01%	65.34%	12.20%	AD	33.41%	39.21%	2.55%
KS	45.02%	58.33%	33.61%	KS	49.50%	66.87%	11.40%
n=30				n=50			
MOM	62.70%	91.51%	-4.80%	MOM	57.47%	85.74%	-4.88%
MLE	46.75%	58.32%	-0.54%	MLE	34.25%	39.35%	1.38%
PWMU	62.13%	91.94%	-4.70%	PWMU	58.65%	88.97%	-5.44%
PWMB	62.65%	93.78%	-5.37%	PWMB	59.02%	90.13%	-5.82%
AD	30.62%	31.58%	3.55%	AD	27.02%	25.45%	3.90%
KS	49.78%	70.95%	7.12%	KS	48.93%	72.19%	1.93%
n=80				n=100			
MOM	56.35%	87.21%	-5.48%	MOM	53.21%	85.41%	-5.52%
MLE	28.30%	31.53%	2.64%	MLE	26.76%	28.59%	3.43%
PWMU	58.24%	91.91%	-6.42%	PWMU	55.53%	90.95%	-6.40%
PWMB	58.48%	92.63%	-6.64%	PWMB	55.74%	91.54%	-6.57%
AD	24.36%	24.39%	3.98%	AD	23.55%	23.08%	4.38%
KS	49.22%	72.08%	1.80%	KS	48.55%	70.74%	4.71%
n=150				n=200			
MOM	54.39%	82.49%	-5.55%	MOM	55.29%	82.93%	-5.66%
MLE	25.91%	24.72%	3.93%	MLE	24.37%	21.47%	4.48%
PWMU	56.89%	89.09%	-6.72%	PWMU	57.96%	89.66%	-6.86%
PWMB	57.02%	89.50%	-6.83%	PWMB	58.06%	89.96%	-6.94%
AD	23.24%	21.41%	4.45%	AD	23.15%	19.40%	4.65%
KS	48.65%	71.41%	3.27%	KS	48.50%	69.93%	3.54%

Table B.8 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1+N(0,0.5)$, $\zeta=-0.5$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	134.38%	217.49%	-6.63%	MOM	114.80%	165.00%	-6.22%
MLE	71.64%	104.70%	-6.06%	MLE	61.98%	74.05%	0.19%
PWMU	113.95%	178.41%	-3.88%	PWMU	107.75%	159.61%	-5.55%
PWMB	110.32%	172.23%	-5.50%	PWMB	103.92%	156.81%	-6.13%
AD	64.53%	84.94%	3.67%	AD	50.69%	51.19%	5.14%
KS	71.94%	96.34%	34.85%	KS	74.36%	103.51%	18.56%
n=30				n=50			
MOM	112.19%	160.74%	-6.44%	MOM	111.10%	151.33%	-6.76%
MLE	62.99%	59.92%	2.71%	MLE	50.75%	48.94%	4.51%
PWMU	106.57%	153.41%	-6.31%	PWMU	106.04%	147.64%	-7.06%
PWMB	103.99%	152.99%	-6.67%	PWMB	103.68%	147.54%	-7.26%
AD	47.30%	45.51%	5.92%	AD	45.50%	39.33%	6.25%
KS	79.05%	100.66%	14.59%	KS	80.87%	99.55%	13.18%
n=80				n=100			
MOM	106.53%	149.66%	-6.87%	MOM	106.75%	147.90%	-6.92%
MLE	42.82%	41.53%	6.43%	MLE	41.89%	36.10%	7.03%
PWMU	104.08%	144.43%	-7.44%	PWMU	104.95%	145.49%	-7.51%
PWMB	103.90%	144.44%	-7.55%	PWMB	104.82%	145.42%	-7.59%
AD	40.55%	33.78%	7.32%	AD	40.19%	34.48%	7.68%
KS	78.11%	99.06%	10.80%	KS	78.44%	101.21%	11.40%
n=150				n=200			
MOM	106.33%	146.63%	-6.80%	MOM	106.82%	147.30%	-6.91%
MLE	40.50%	33.17%	7.86%	MLE	37.84%	30.05%	8.02%
PWMU	102.69%	146.02%	-7.55%	PWMU	102.42%	151.96%	-7.68%
PWMB	102.62%	146.01%	-7.60%	PWMB	102.35%	151.93%	-7.72%
AD	38.62%	30.46%	8.08%	AD	37.71%	28.83%	8.10%
KS	80.34%	99.27%	13.85%	KS	79.34%	104.86%	12.23%

Table B.9 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1, \zeta=-0.5, \sigma=2+N(0,0.2)$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	16.81%	31.82%	-0.15%	MOM	4.77%	13.47%	-0.22%
MLE	32.64%	78.13%	-10.42%	MLE	20.16%	45.03%	-7.81%
PWMU	9.76%	10.51%	9.26%	PWMU	1.72%	4.43%	4.43%
PWMB	12.72%	20.75%	3.62%	PWMB	3.64%	10.22%	1.99%
AD	13.70%	21.97%	25.45%	AD	3.01%	7.02%	3.74%
KS	4.37%	-16.65%	70.08%	KS	1.20%	-5.33%	19.07%
n=30				n=50			
MOM	3.81%	9.75%	0.40%	MOM	2.65%	4.16%	-0.11%
MLE	15.91%	34.65%	-5.15%	MLE	8.85%	18.03%	-3.84%
PWMU	1.79%	4.00%	3.37%	PWMU	1.52%	0.82%	1.70%
PWMB	3.11%	7.96%	1.81%	PWMB	2.32%	3.27%	0.83%
AD	2.18%	4.21%	2.19%	AD	0.89%	-0.59%	1.27%
KS	1.37%	-2.82%	15.26%	KS	1.24%	-4.46%	9.60%
n=80				n=100			
MOM	1.95%	1.74%	0.20%	MOM	1.85%	2.30%	-0.38%
MLE	5.87%	11.67%	-2.49%	MLE	4.43%	8.79%	-2.21%
PWMU	1.25%	-0.77%	1.36%	PWMU	1.44%	0.71%	0.40%
PWMB	1.76%	0.78%	0.83%	PWMB	1.85%	1.94%	-0.01%
AD	0.63%	-1.08%	0.99%	AD	0.24%	-1.71%	0.62%
KS	1.00%	-4.97%	5.32%	KS	1.30%	-3.99%	3.09%
n=150				n=200			
MOM	0.49%	-0.32%	0.21%	MOM	0.31%	1.38%	0.17%
MLE	2.98%	5.99%	-1.25%	MLE	1.93%	4.83%	-1.21%
PWMU	0.13%	-1.72%	0.79%	PWMU	0.07%	0.75%	0.65%
PWMB	0.40%	-0.89%	0.51%	PWMB	0.27%	1.37%	0.44%
AD	-0.34%	-2.06%	0.82%	AD	-0.49%	-1.44%	0.64%
KS	-0.31%	-4.46%	3.00%	KS	-0.12%	-1.03%	1.93%

Table B.10 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1$, $\zeta =-0.5$, $\sigma=2+N(0,0.6)$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	16.56%	37.06%	-13.39%	MOM	6.11%	9.37%	-0.31%
MLE	33.08%	75.38%	-22.49%	MLE	20.49%	41.96%	-7.65%
PWMU	9.52%	15.95%	-5.08%	PWMU	3.22%	0.19%	3.92%
PWMB	12.52%	25.72%	-9.97%	PWMB	5.08%	6.01%	1.54%
ADR	13.82%	22.11%	2.18%	AD	3.84%	2.40%	2.60%
KS	2.02%	-18.39%	57.28%	KS	1.35%	-14.24%	26.83%
n=30				n=50			
MOM	4.47%	8.73%	-0.35%	MOM	2.57%	5.56%	0.66%
MLE	16.05%	31.05%	-5.66%	MLE	8.65%	19.00%	-3.15%
PWMU	2.63%	3.10%	2.56%	PWMU	1.51%	2.77%	2.49%
PWMB	3.93%	7.06%	1.04%	PWMB	2.31%	5.19%	1.59%
ADR	2.54%	0.03%	2.24%	ADR	0.53%	0.06%	1.98%
KS	1.20%	-10.34%	18.10%	KS	0.57%	-6.21%	11.62%
n=80				n=100			
MOM	1.21%	3.88%	2.00%	MOM	1.37%	1.64%	-0.20%
MLE	5.04%	12.62%	-0.44%	MLE	4.51%	8.60%	-2.37%
PWMU	0.57%	1.85%	3.09%	PWMU	0.84%	-0.01%	0.69%
PWMB	1.07%	3.38%	2.54%	PWMB	1.25%	1.23%	0.27%
ADR	-0.23%	-0.13%	2.98%	ADR	0.17%	-2.13%	0.64%
KS	-0.35%	-5.97%	9.11%	KS	0.04%	-5.43%	5.48%
n=150				n=200			
MOM	1.09%	2.70%	1.30%	MOM	-0.87%	0.39%	0.69%
MLE	3.47%	7.09%	-0.06%	MLE	0.75%	4.29%	-0.40%
PWMU	0.74%	1.83%	1.84%	PWMU	-1.08%	-0.46%	1.07%
PWMB	1.01%	2.65%	1.56%	PWMB	-0.87%	0.16%	0.87%
ADR	0.15%	-0.65%	1.96%	ADR	-1.82%	-2.13%	1.37%
KS	0.32%	-1.55%	5.21%	KS	-1.57%	-3.76%	3.11%

Table B.11 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1$, $\zeta =-0.5$, $\sigma=2+N(0,1.0)$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	16.74%	31.67%	2.71%	MOM	10.82%	11.00%	5.70%
MLE	34.36%	72.70%	-5.21%	MLE	25.79%	45.26%	-2.28%
PWMU	9.41%	10.17%	12.44%	PWMU	7.77%	1.44%	10.47%
PWMB	12.56%	20.33%	6.51%	PWMB	9.68%	7.31%	7.85%
ADR	14.20%	20.62%	16.75%	ADR	8.26%	4.06%	9.89%
KS	6.03%	-34.09%	78.84%	KS	3.84%	-20.36%	36.80%
n=30				n=50			
MOM	6.74%	5.18%	4.15%	MOM	5.41%	4.78%	4.06%
MLE	18.40%	31.01%	-1.78%	MLE	11.63%	17.53%	0.41%
PWMU	4.85%	-0.99%	7.20%	PWMU	4.35%	1.32%	5.91%
PWMB	6.17%	3.07%	5.56%	PWMB	5.18%	3.76%	5.00%
ADR	4.60%	0.66%	6.45%	ADR	3.52%	-0.93%	5.50%
KS	1.74%	-19.61%	25.09%	KS	2.14%	-9.56%	19.16%
n=80				n=100			
MOM	4.99%	2.11%	1.19%	MOM	5.64%	2.44%	0.50%
MLE	8.07%	11.19%	-1.63%	MLE	9.15%	9.44%	-1.57%
PWMU	4.44%	-0.19%	2.37%	PWMU	5.04%	0.70%	1.35%
PWMB	4.96%	1.36%	1.82%	PWMB	5.47%	1.93%	0.93%
ADR	3.02%	-1.60%	1.98%	ADR	4.65%	-0.92%	1.38%
KS	2.29%	-9.00%	10.16%	KS	3.11%	-7.86%	7.68%
n=150				n=200			
MOM	5.69%	2.03%	2.15%	MOM	4.38%	1.14%	1.40%
MLE	8.38%	6.94%	0.72%	MLE	5.97%	5.05%	0.13%
PWMU	5.33%	1.05%	2.68%	PWMU	4.13%	0.30%	1.88%
PWMB	5.62%	1.88%	2.40%	PWMB	4.34%	0.92%	1.67%
ADR	4.80%	-1.07%	2.86%	ADR	3.53%	-1.05%	1.89%
KS	3.56%	-7.47%	6.89%	KS	2.96%	-6.25%	5.51%

Table B.12 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1$, $\zeta = -0.5 + N(0, 0.05)$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	15.12%	37.78%	-0.02%	MOM	6.62%	12.27%	0.48%
MLE	32.61%	78.36%	-11.79%	MLE	20.04%	42.25%	-7.06%
PWMU	7.40%	16.16%	9.42%	PWMU	3.97%	3.33%	5.15%
PWMB	10.47%	25.87%	3.72%	PWMB	5.83%	9.12%	2.64%
ADR	12.40%	24.22%	13.04%	ADR	3.54%	3.31%	4.34%
KS	3.27%	-11.52%	60.91%	KS	3.85%	-7.96%	23.13%
n=30				n=50			
MOM	3.06%	5.75%	0.12%	MOM	1.32%	2.24%	-0.21%
MLE	14.39%	30.52%	-4.70%	MLE	7.71%	17.22%	-3.32%
PWMU	1.20%	-0.22%	2.83%	PWMU	0.14%	-1.34%	1.35%
PWMB	2.51%	3.81%	1.31%	PWMB	0.95%	1.13%	0.49%
ADR	1.16%	1.27%	2.80%	ADR	0.09%	-2.17%	1.57%
KS	0.24%	-7.23%	14.71%	KS	-0.26%	-6.91%	8.88%
n=80				n=100			
MOM	1.81%	4.47%	0.73%	MOM	2.44%	1.69%	0.53%
MLE	5.95%	13.22%	-2.17%	MLE	5.32%	9.63%	-1.46%
PWMU	1.14%	2.33%	1.96%	PWMU	1.97%	-0.06%	1.36%
PWMB	1.65%	3.86%	1.41%	PWMB	2.38%	1.18%	0.93%
ADR	0.42%	-0.20%	1.48%	ADR	0.94%	-1.41%	1.44%
KS	1.20%	-1.53%	7.57%	KS	1.53%	-3.49%	4.16%
n=150				n=200			
MOM	0.64%	1.21%	0.17%	MOM	0.57%	1.15%	0.70%
MLE	3.07%	6.99%	-1.17%	MLE	2.37%	5.25%	-0.63%
PWMU	0.29%	-0.01%	0.66%	PWMU	0.34%	0.24%	1.13%
PWMB	0.56%	0.82%	0.39%	PWMB	0.55%	0.86%	0.91%
ADR	-0.30%	-0.90%	0.83%	ADR	-0.28%	-0.90%	1.23%
KS	0.06%	-1.76%	2.97%	KS	0.04%	-1.22%	2.51%

Table B.13 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1$, $\zeta = -0.5 + N(0, 0.15)$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	15.69%	32.98%	2.73%	MOM	9.49%	10.48%	4.06%
MLE	32.24%	73.18%	-7.98%	MLE	20.97%	39.42%	-3.35%
PWMU	8.21%	11.56%	13.02%	PWMU	6.67%	1.02%	9.38%
PWMB	11.32%	21.56%	6.85%	PWMB	8.53%	6.83%	6.57%
ADR	12.80%	21.89%	19.81%	ADR	5.94%	3.28%	10.41%
KS	7.02%	-18.85%	67.65%	KS	4.67%	-11.1%	32.01%
n=30				n=50			
MOM	4.84%	7.70%	4.47%	MOM	2.32%	3.68%	3.46%
MLE	14.75%	29.85%	-1.23%	MLE	7.56%	16.41%	-0.18%
PWMU	2.84%	2.06%	7.82%	PWMU	1.29%	0.35%	5.39%
PWMB	4.13%	6.02%	6.06%	PWMB	2.09%	2.78%	4.41%
ADR	2.17%	0.82%	7.80%	ADR	0.45%	-1.74%	5.41%
KS	2.23%	-4.84%	19.27%	KS	1.03%	-4.61%	14.62%
n=80				n=100			
MOM	1.54%	3.07%	3.37%	MOM	0.64%	2.39%	2.78%
MLE	4.91%	10.61%	1.10%	MLE	3.81%	10.06%	0.77%
PWMU	0.86%	1.02%	4.50%	PWMU	0.09%	0.58%	3.68%
PWMB	1.38%	2.54%	3.91%	PWMB	0.51%	1.80%	3.22%
ADR	0.28%	-1.19%	4.83%	ADR	-0.38%	-0.71%	3.92%
KS	0.61%	-1.88%	8.84%	KS	-0.18%	-2.64%	7.34%
n=150				n=200			
MOM	0.33%	1.52%	3.09%	MOM	0.22%	0.69%	2.86%
MLE	3.19%	7.20%	1.94%	MLE	1.85%	6.55%	2.01%
PWMU	-0.07%	0.44%	3.59%	PWMU	-0.06%	2.90%	3.20%
PWMB	0.21%	1.26%	3.29%	PWMB	0.15%	3.50%	2.98%
ADR	-0.60%	-1.21%	3.98%	ADR	-0.68%	1.36%	3.61%
KS	-0.47%	-2.10%	5.60%	KS	-0.21%	-1.58%	4.75%

Table B.14 Bias of estimated results (shape parameter ζ , scale parameter σ and 99th percentile) based on simulations from GPD model ($u=1$, $\zeta = -0.5 + N(0, 0.25)$, $\sigma=2$).

	σ	ζ	99th percentile		σ	ζ	99th percentile
n=10				n=20			
MOM	18.88%	38.32%	8.66%	MOM	6.06%	13.26%	8.22%
MLE	31.49%	69.17%	14.81%	MLE	16.81%	33.45%	2.21%
PWMU	10.85%	15.16%	21.27%	PWMU	2.85%	3.39%	14.13%
PWMB	13.94%	24.50%	14.14%	PWMB	4.77%	8.96%	11.03%
ADR	14.49%	22.94%	43.13%	ADR	3.42%	1.38%	16.85%
KS	5.36%	-14.28%	124.32%	KS	2.33%	-5.67%	35.43%
n=30				n=50			
MOM	3.73%	6.32%	9.18%	MOM	1.83%	2.25%	9.52%
MLE	12.76%	26.12%	4.87%	MLE	7.07%	12.80%	6.20%
PWMU	1.72%	-0.15%	13.49%	PWMU	0.42%	-1.23%	11.51%
PWMB	3.06%	3.74%	11.44%	PWMB	1.25%	1.17%	10.37%
ADR	1.82%	1.45%	16.05%	ADR	0.49%	-4.01%	12.09%
KS	1.36%	-2.68%	30.21%	KS	-0.11%	-5.41%	17.56%
n=80				n=100			
MOM	2.52%	3.04%	9.53%	MOM	1.09%	0.52%	10.14%
MLE	4.80%	8.60%	7.51%	MLE	3.23%	5.67%	8.56%
PWMU	1.63%	0.57%	11.24%	PWMU	0.42%	-1.54%	11.17%
PWMB	2.16%	2.06%	10.52%	PWMB	0.85%	-0.33%	10.60%
ADR	0.70%	-1.56%	12.06%	ADR	-0.23%	-3.53%	11.92%
KS	1.58%	-2.22%	16.59%	KS	0.37%	-4.26%	14.50%
n=150				n=200			
MOM	0.67%	0.44%	8.46%	MOM	0.22%	-0.05%	8.93%
MLE	2.42%	4.28%	7.11%	MLE	0.92%	1.64%	7.73%
PWMU	0.20%	-0.86%	9.28%	PWMU	-0.16%	-1.21%	9.59%
PWMB	0.49%	-0.05%	8.92%	PWMB	0.06%	-0.60%	9.32%
ADR	-0.37%	-3.02%	9.62%	ADR	-0.62%	-2.45%	9.95%
KS	0.01%	-2.87%	12.08%	KS	-0.37%	-3.43%	12.45%

Appendix C. Selection of Threshold and Time Span in POT

Method

As discussed in Chapter 3, the applicability of the POT method implies that the exceedances fit well to a Pareto distribution and the occurrence of the exceedances follow the Poisson process. Thus, any pair of threshold u and time interval Δt which leads to both conditions being satisfied is a feasible solution.

For the set of data used in Section 3.4.2, the K-S test is performed for both the Pareto distribution and the Poisson model. The computed p-values are shown in Tables C.1 and C.2. The estimated 0.95 confidence interval for the 100 year return value are also presented in Tables C.3 and C.4.

Table C.1 P-value of K-S test for exceedances following GPD with different values of $(u, \Delta t)$.

		u(m)										
		2	2.4	2.8	3.2	3.6	4	4.4	4.8	5.2	5.6	6
Δt (hrs)	12	0.300	0.830	0.911	0.675	0.971	0.633	0.798	0.948	0.861	0.576	0.189
	24	0.322	0.824	0.916	0.891	0.742	0.549	0.825	0.971	0.881	0.617	0.245
	36	0.206	0.554	0.721	0.654	0.514	0.665	0.834	0.967	0.898	0.683	0.219
	48	0.028	0.275	0.600	0.593	0.607	0.809	0.840	0.964	0.907	0.761	0.367
	60	0.004	0.118	0.481	0.501	0.352	0.872	0.962	0.986	0.947	0.761	0.439
	72	0.001	0.023	0.468	0.391	0.375	0.922	0.811	0.986	0.973	0.796	0.502
	84	0.001	0.034	0.194	0.181	0.256	0.930	0.697	0.970	0.936	0.772	0.515
	96	0.001	0.022	0.116	0.111	0.178	0.963	0.541	0.945	0.927	0.830	0.503
	108	0.001	0.018	0.136	0.141	0.152	0.973	0.687	0.965	0.943	0.767	0.365
	120	0.000	0.015	0.192	0.289	0.236	0.902	0.771	0.957	0.931	0.831	0.365
	132	0.000	0.023	0.176	0.277	0.252	0.900	0.745	0.900	0.877	0.808	0.387
	144	0.000	0.010	0.042	0.089	0.138	0.901	0.850	0.978	0.922	0.849	0.543

Notes: For p-value smaller than 0.05, hypothesis of exceedances following Pareto distribution will be rejected at significance level $\alpha=0.05$.

Table C.2 P-value of K-S test for occurrences of exceedances following Poisson process with different values of $(u, \Delta t)$.

		u(m)										
		2	2.4	2.8	3.2	3.6	4	4.4	4.8	5.2	5.6	6
Δt (hrs)	12	0.000	0.000	0.000	0.000	0.022	0.102	0.174	0.120	0.430	0.462	0.598
	24	0.000	0.000	0.000	0.000	0.001	0.009	0.082	0.074	0.426	0.405	0.508
	36	0.000	0.000	0.000	0.000	0.000	0.001	0.012	0.046	0.357	0.505	0.449
	48	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.024	0.225	0.411	0.459
	60	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.039	0.391	0.492	0.712
	72	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.018	0.240	0.402	0.612
	84	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.009	0.260	0.290	0.486
	96	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.011	0.177	0.213	0.398
	108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.139	0.163	0.322
	120	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.066	0.114	0.268
	132	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.026	0.067	0.213
	144	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.014	0.043	0.150

Notes: For p-value smaller than 0.05, hypothesis of exceedances following Poisson process will be rejected at significance level $\alpha=0.05$.

As described in Chapter 3, the feasible region for significance level of 0.01 is larger than that for 0.05. This can be observed from the following figure.

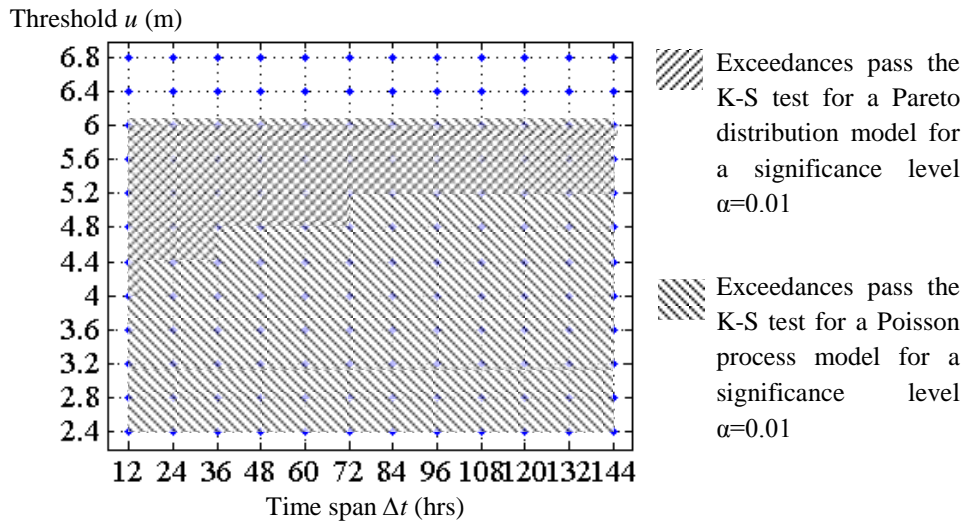


Figure C.1 Appropriate region for u and Δt for significance level equals to 0.01.

To show the significance of the selected values of $(u, \Delta t)$, the estimated scale parameter, shape parameter and 100-year return value for different values of u and Δt are shown in Fig. C.2. The plotted region is for $u \in [2.0, 6.0]$ and $\Delta t \in [12, 240]$. As shown in the figure, the estimated parameter values are very sensitive to the selected values of $(u, \Delta t)$. One could observe this from Fig. C.2 (c) where the estimated 100 year return value deviates quite a lot from each other. A high threshold leads to a higher estimate in the return value, whereas a larger value of time span leads to a lower value in the estimate.

Table C.3 Estimated lower bound of 0.95 confidence interval for 100 year design value with different values of $(u, \Delta t)$.

		u(m)										
		2	2.4	2.8	3.2	3.6	4	4.4	4.8	5.2	5.6	6
Δt (hrs)	12	11.652	11.707	11.676	11.802	11.370	11.305	11.283	11.303	11.265	11.220	11.285
	24	11.059	11.119	11.318	11.364	11.166	11.193	11.203	11.252	11.205	11.196	11.301
	36	10.950	10.969	11.077	11.176	11.109	11.173	11.176	11.201	11.182	11.186	11.316
	48	10.936	10.938	11.023	11.100	11.090	11.156	11.148	11.169	11.168	11.171	11.259
	60	10.986	10.951	11.002	11.067	11.070	11.161	11.169	11.182	11.186	11.179	11.224
	72	11.064	10.992	11.003	11.037	11.051	11.126	11.135	11.168	11.190	11.177	11.210
	84	11.120	11.053	11.021	11.031	11.049	11.132	11.133	11.176	11.196	11.192	11.212
	96	11.172	11.095	11.050	11.053	11.058	11.130	11.109	11.159	11.195	11.186	11.224
	108	11.215	11.135	11.076	11.079	11.077	11.146	11.122	11.160	11.195	11.182	11.240
	120	11.283	11.164	11.098	11.097	11.096	11.132	11.127	11.160	11.169	11.168	11.240
	132	11.190	11.193	11.128	11.117	11.114	11.141	11.140	11.160	11.171	11.163	11.228
	144	11.402	11.035	11.155	11.124	11.115	11.149	11.152	11.173	11.176	11.166	11.221

Table C.4 Estimated upper bound of 0.95 confidence interval for 100 year design value with different values of (u , Δt).

		u(m)										
		2	2.4	2.8	3.2	3.6	4	4.4	4.8	5.2	5.6	6
Δt (hrs)	12	13.723	13.819	13.910	14.277	13.877	14.039	14.246	14.643	14.870	15.054	16.088
	24	12.871	13.059	13.468	13.716	13.515	13.766	14.026	14.487	14.623	14.919	16.223
	36	12.599	12.716	13.074	13.423	13.376	13.740	13.961	14.314	14.511	14.864	16.381
	48	12.445	12.582	12.948	13.264	13.342	13.710	13.870	14.180	14.429	14.741	15.931
	60	12.300	12.518	12.872	13.168	13.284	13.761	13.978	14.259	14.542	14.802	15.643
	72	12.260	12.391	12.831	13.046	13.184	13.626	13.827	14.192	14.572	14.784	15.499
	84	12.252	12.379	12.767	12.938	13.149	13.666	13.832	14.253	14.626	14.905	15.527
	96	12.265	12.403	12.772	12.939	13.133	13.655	13.683	14.150	14.620	14.854	15.656
	108	12.299	12.419	12.795	13.006	13.148	13.731	13.728	14.139	14.629	14.800	15.849
	120	12.307	12.433	12.833	13.085	13.151	13.610	13.695	14.107	14.395	14.630	15.849
	132	12.342	12.471	12.784	13.076	13.191	13.630	13.721	14.053	14.356	14.455	15.711
	144	12.325	12.432	12.726	13.047	13.190	13.715	13.864	14.192	14.395	14.480	15.660

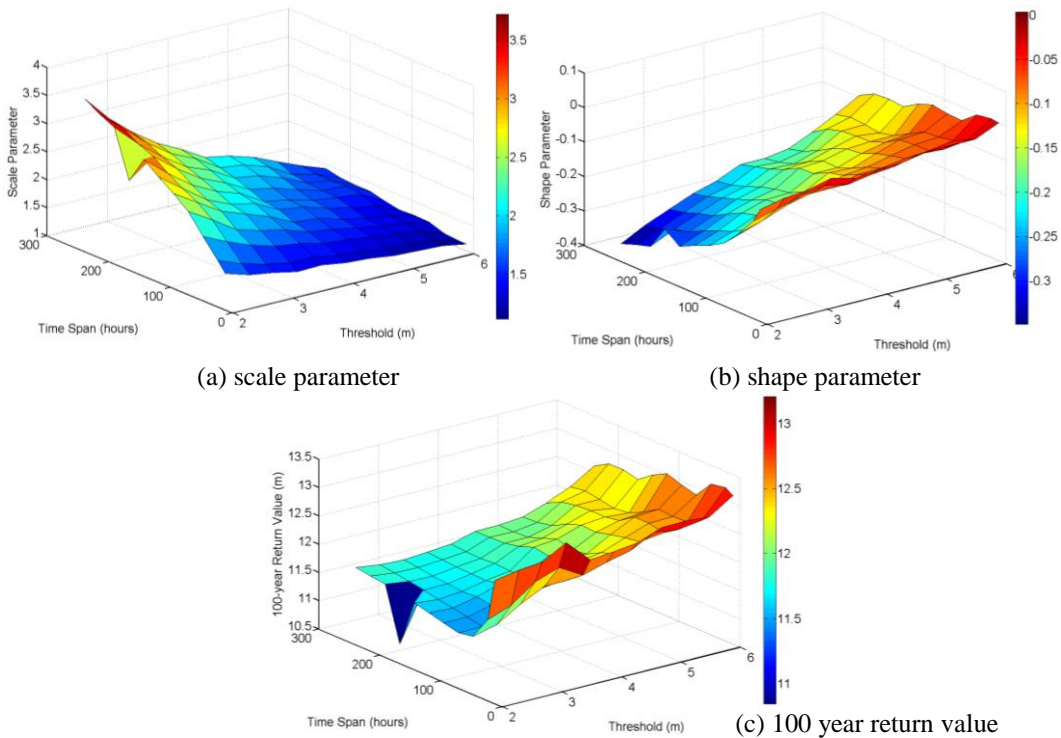


Figure C.2 Influence of u and Δt to (a) scale parameter, (b) shape parameter and (c) 100 year return value.

Table C.5 Number of exceedances above the threshold after applying the POT method with different values of $(u, \Delta t)$.

		u(m)										
		2.4	2.8	3.2	3.6	4	4.4	4.8	5.2	5.6	6	6.4
Δt (hrs)	12	1606	1344	1104	857	666	497	371	271	197	145	99
	24	1485	1259	1028	802	626	470	354	258	192	144	97
	36	1327	1151	950	751	587	447	336	249	187	142	96
	48	1202	1033	860	696	557	424	324	244	184	139	96
	60	1077	953	798	644	522	402	307	234	177	133	95
	72	972	875	743	612	501	391	303	232	176	132	94
	84	878	794	682	574	479	373	294	226	173	129	92
	96	813	737	636	543	461	362	286	222	169	128	91
	108	759	698	604	510	441	348	276	217	165	127	89
	120	718	661	581	483	413	330	265	210	162	127	89
	132	674	621	549	463	397	316	255	203	156	125	89
	144	635	590	525	447	386	310	249	194	151	122	87

Appendix D. Testing Values of Threshold and Time Span in POT Approach for Each Identified Time Sectors

To check the adequacy of the Poisson-GPD model for H_S in different time sectors as prescribed in Section 4.3, the exceedances are here tested in both the Pareto model and Poisson process model with reference to the appropriate threshold and minimum time difference to reduce serial correlation.

As discussed in Chapter 3, this model assumes the occurrence rate to follow a Poisson distribution. An easy test can be performed by transforming the random variable t_i ($i = 1, \dots, n$), which is the exact occurrence time of the extremes, to a uniformly distributed random variable y_i ($i = 1, \dots, n$) based on the property of Poisson process (Luceño 2006):

$$y_i = 1 - e^{-\int_{t_{i-1}}^{t_i} \lambda(t) dt} \quad (D.1)$$

where $\lambda(t)$ is the intensity function mentioned in Eq. (3.6). In this study, this refers to the average occurrence rate for the extremes within a reference time period.

Therefore, by checking the quantile-quantile plot (which compares the empirical value y_i and the theoretical value $r_{i=(i-1/2)/n}$), a suitable Δt can be selected in the POT method. Here, Δt is tested at a value of 12 hours for the applied data. The threshold selected here is taken as 1.5 times the mean value for

each season which is a reasonable selection from a practical point of view (Boccotti 2000). However, in order to justify the chosen threshold and time span are good enough, the suitability of several other values of threshold and time span are compared in the following. Figure D.1 illustrates the comparison between y_i and r_i in each of the identified four time sectors for using the time span at 12, 24 and 36 hours. The figure shows that the Poisson process model for adopting time span at 12 hours is most adequate as the maximum difference between the empirical CDF and the theoretical CDF is the smallest one among the tested three. This supports the adequate use of Δt .

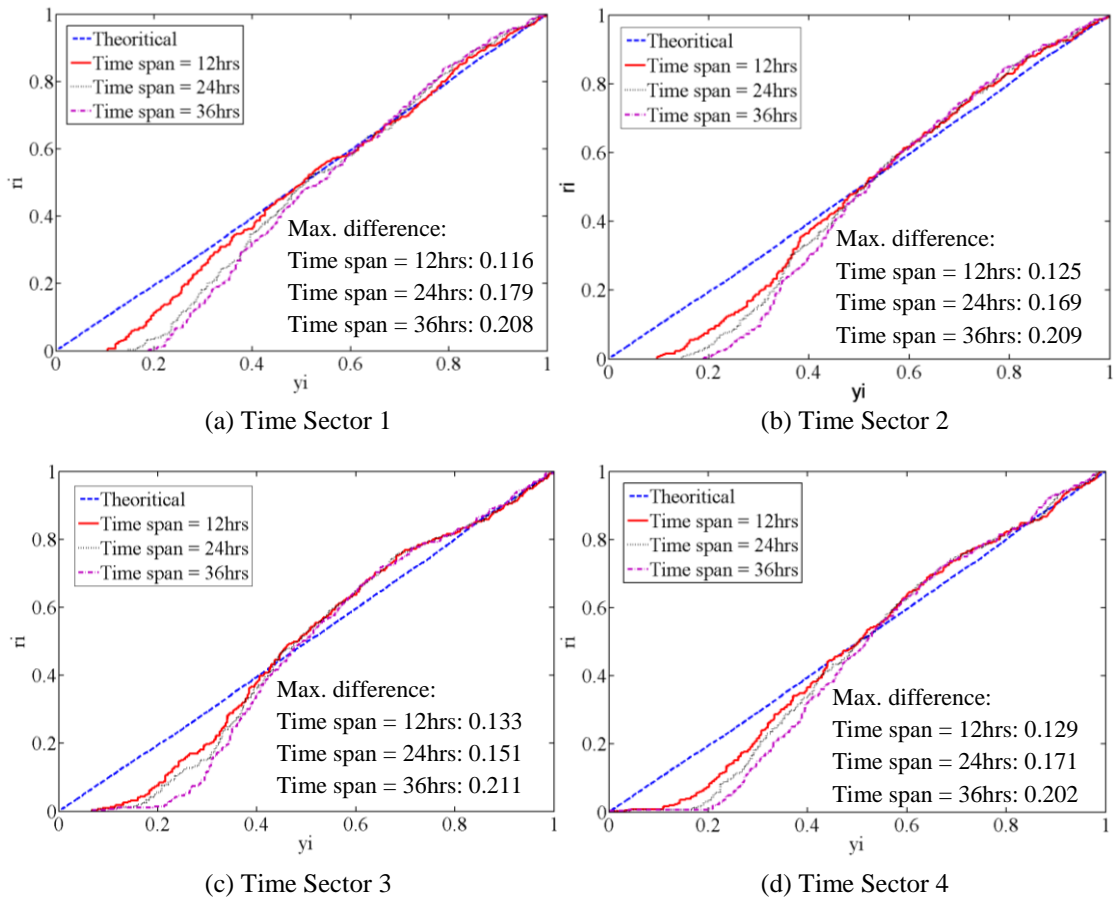
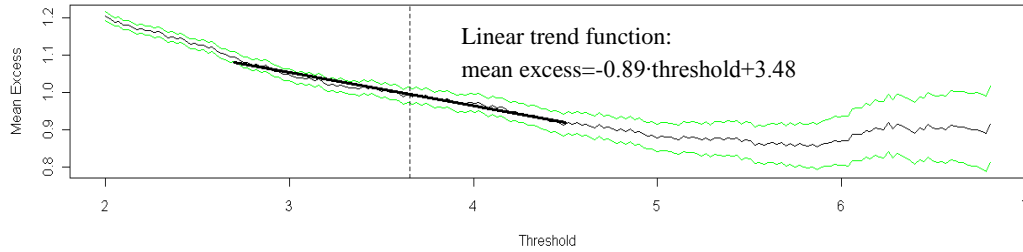
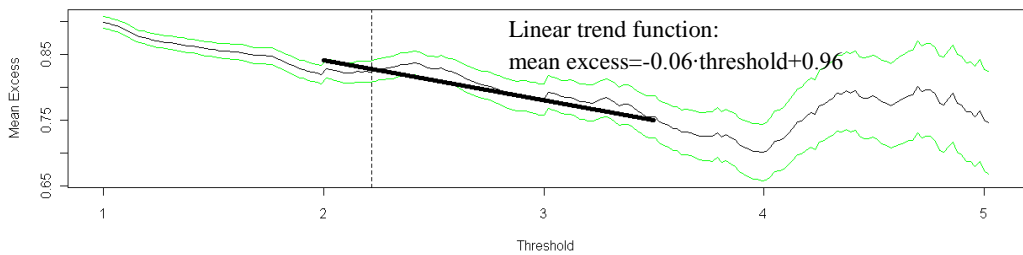


Figure D.1 Test of Poisson process in each identified time sector.

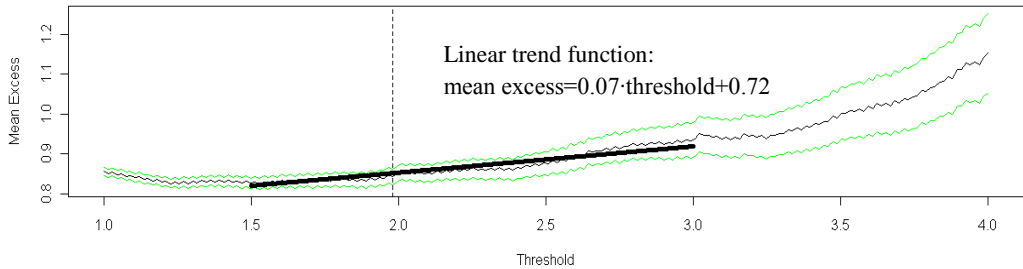
The appropriateness of the threshold is also tested by using the mean residual plot and the L-moment plot which have been discussed in Chapter 3. The mean residual plots for all the identified four time sectors are shown in Fig. D.2. It is seen that the selected threshold is within the region that has a linear trend



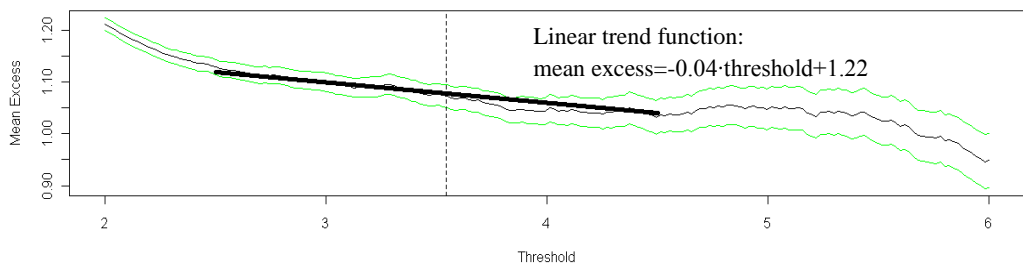
(a) Time Sector 1



(b) Time Sector 2



(c) Time Sector 3



(d) Time Sector 4

Figure D.2 Test of threshold (vertical dotted line) by mean residual plot with 95% confident intervals (green line) for all four time sectors.

drawn on the figures. This supports the hypothesis of GPD model for the data above the chosen threshold.

Figure D.3 shows the L-moment plot comparing the ratio of L-skewness and L-kurtosis between the empirical value and the theoretical value. In order to compare with other choices of thresholds, the figure also plots out the result for using threshold equals to 0.5, 1.0 and 2.0 times of sample mean. The selected threshold appears to be the most compatible with the GPD model since it

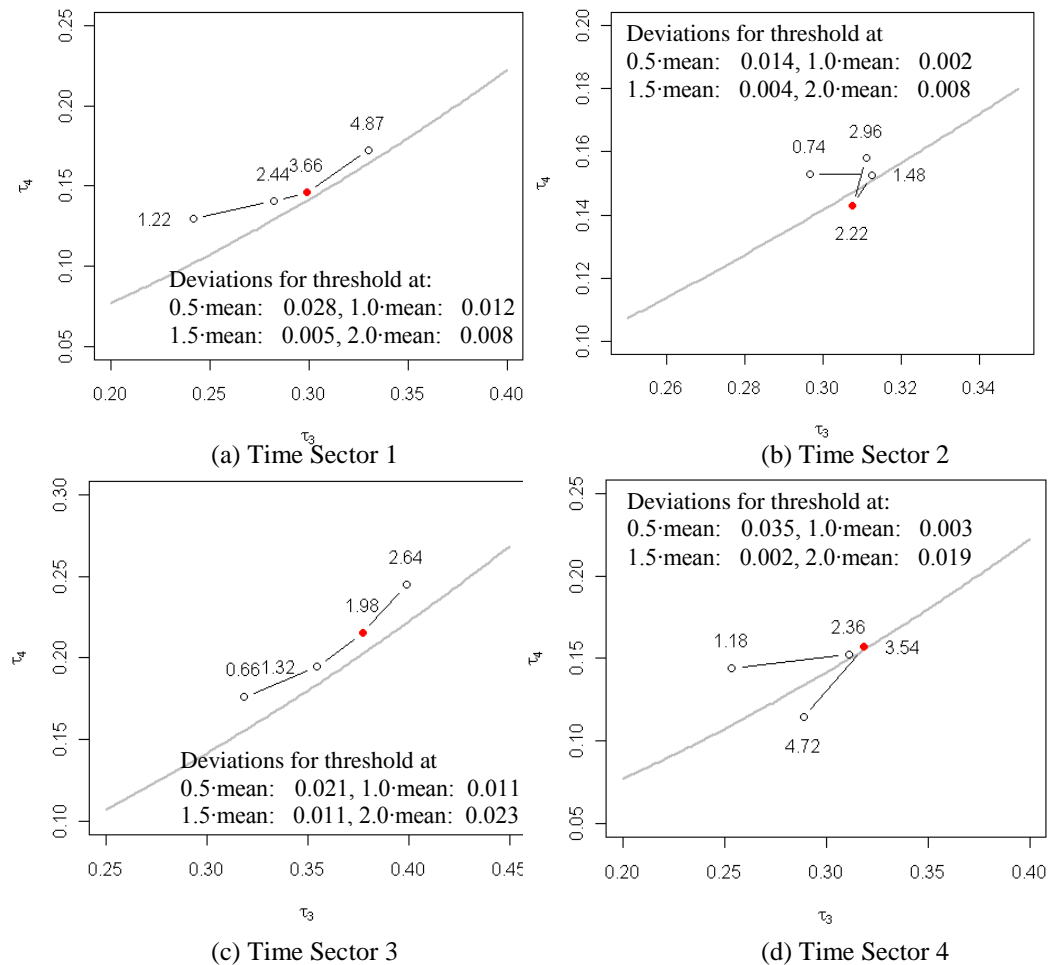


Figure D.3 L-moment plot for exceedances over the selected threshold (circle dot) with the theoretical GPD curve (grey line) for all four time sectors.

generally gives the overall least deviations from the theoretical line for all the four seasons. This confirms that a threshold of 1.5 times the mean is suitable and it agrees well with the previous inference.

Appendix E. USFOS Program Input for the Example

Structure

The following code is used to construct the structural model in USFOS program for the example analysis in Chapter 4.

```
HEAD          Dynamic Pushover
              Progressive Collapse Analysis / JACKET model
              SINTEF 1997
,
,
Node ID      X      Y      Z      Boundary code
NODE 51     -27.000  -13.500  25.000
NODE 52      27.000  -13.500  25.000
NODE 53      27.000   13.500  25.000
NODE 54     -27.000   13.500  25.000
NODE 57     -10.000  -13.500  25.000
NODE 58      10.000  -13.500  25.000
NODE 59     -10.000   13.500  25.000
NODE 60      10.000   13.500  25.000
NODE 61     -18.500   .000   25.000
NODE 62       .000   .000  25.000
NODE 63      18.500   .000  25.000
NODE 101     -27.000  -13.500  34.750
NODE 102      27.000  -13.500  34.750
NODE 103      27.000   13.500  34.750
NODE 104     -27.000   13.500  34.750
NODE 107     -10.000  -13.500  34.750
NODE 108      10.000  -13.500  34.750
NODE 109     -10.000   13.500  34.750
NODE 110      10.000   13.500  34.750
NODE 111     -27.000   .000  34.750
NODE 112     -10.000   .000  34.750
NODE 113      10.000   .000  34.750
NODE 114      27.000   .000  34.750
NODE 115     -18.500  -13.500  34.750
NODE 116       .000  -13.500  34.750
NODE 117      18.500  -13.500  34.750
```

NODE	118	-18.500	13.500	34.750
NODE	119	.000	13.500	34.750
NODE	120	18.500	13.500	34.750
NODE	121	-27.000	-6.750	34.750
NODE	122	-27.000	6.750	34.750
NODE	123	-10.000	-6.750	34.750
NODE	124	-10.000	6.750	34.750
NODE	125	10.000	-6.750	34.750
NODE	126	10.000	6.750	34.750
NODE	127	27.000	-6.750	34.750
NODE	128	27.000	6.750	34.750
NODE	130	-39.000	-13.500	34.750
NODE	131	-34.000	-13.500	34.750
NODE	132	-39.000	13.500	34.750
NODE	133	-34.000	13.500	34.750
NODE	134	34.000	-13.500	34.750
NODE	135	34.000	13.500	34.750
NODE	136	-34.000	-13.500	25.000
NODE	137	-34.000	13.500	25.000
NODE	138	34.000	-13.500	25.000
NODE	139	34.000	13.500	25.000
NODE	140	-27.000	.000	25.000
NODE	141	-10.000	.000	25.000
NODE	142	10.000	.000	25.000
NODE	143	27.000	.000	25.000
NODE	144	-18.500	.000	34.750
NODE	145	.000	.000	34.750
NODE	146	18.500	.000	34.750
NODE	201	-27.000	-13.500	12.000
NODE	202	27.000	-13.500	12.000
NODE	203	27.000	13.500	12.000
NODE	204	-27.000	13.500	12.000
NODE	205	.000	-13.500	12.000
NODE	206	.000	13.500	12.000
NODE	207	-10.000	-13.500	12.000
NODE	208	10.000	-13.500	12.000
NODE	209	-10.000	13.500	12.000
NODE	210	10.000	13.500	12.000
NODE	211	13.500	.000	12.000
NODE	212	-10.000	.000	12.000
NODE	213	-27.000	.000	12.000
NODE	214	.000	.000	12.000
NODE	215	10.000	3.500	12.000
NODE	216	10.000	-3.500	12.000
NODE	217	19.320	-8.290	12.000
NODE	218	19.320	8.290	12.000

NODE	219	20.904	-7.404	12.000
NODE	220	20.904	7.404	12.000
NODE	301	-28.650	-16.580	-14.000
NODE	302	28.650	-16.580	-14.000
NODE	303	28.650	16.580	-14.000
NODE	304	-28.650	16.580	-14.000
NODE	305	-10.000	-16.580	-14.000
NODE	306	10.000	-16.580	-14.000
NODE	307	10.000	16.580	-14.000
NODE	308	-10.000	16.580	-14.000
NODE	311	-28.650	.000	-14.000
NODE	312	-10.000	.000	-14.000
NODE	313	10.000	.000	-14.000
NODE	314	19.320	-8.290	-14.000
NODE	315	19.320	8.290	-14.000
NODE	316	28.650	.000	-14.000
NODE	351	-27.825	-15.040	-1.000
NODE	352	27.825	-15.040	-1.000
NODE	353	27.825	15.040	-1.000
NODE	354	-27.825	15.040	-1.000
NODE	355	-10.000	-15.040	-1.000
NODE	356	10.000	-15.040	-1.000
NODE	357	10.000	15.040	-1.000
NODE	358	-10.000	15.040	-1.000
NODE	361	-18.500	15.040	-1.000
NODE	362	.000	15.040	-1.000
NODE	363	18.500	15.040	-1.000
NODE	364	-18.500	-15.040	-1.000
NODE	365	.000	-15.040	-1.000
NODE	366	18.500	-15.040	-1.000
NODE	367	-27.825	-6.750	-1.000
NODE	368	-27.825	6.750	-1.000
NODE	369	-10.000	-6.750	-1.000
NODE	370	-10.000	6.750	-1.000
NODE	371	10.000	-6.750	-1.000
NODE	372	10.000	6.750	-1.000
NODE	373	27.825	-6.750	-1.000
NODE	374	27.825	6.750	-1.000
NODE	401	-30.480	-20.020	-43.000
NODE	402	30.480	-20.020	-43.000
NODE	403	30.480	20.020	-43.000
NODE	404	-30.480	20.020	-43.000
NODE	405	-10.000	-20.020	-43.000
NODE	406	10.000	-20.020	-43.000
NODE	407	10.000	20.020	-43.000
NODE	408	-10.000	20.020	-43.000

NODE	411	-30.480	.000	-43.000		
NODE	412	-10.000	.000	-43.000		
NODE	413	10.000	.000	-43.000		
NODE	414	20.240	-10.010	-43.000		
NODE	415	20.240	10.010	-43.000		
NODE	416	30.480	.000	-43.000		
NODE	501	-32.440	-23.700	-74.000		
NODE	502	32.440	-23.700	-74.000		
NODE	503	32.440	23.700	-74.000		
NODE	504	-32.440	23.700	-74.000		
NODE	505	-10.000	-23.700	-74.000		
NODE	506	10.000	-23.700	-74.000		
NODE	507	10.000	23.700	-74.000		
NODE	508	-10.000	23.700	-74.000		
NODE	511	-32.440	.000	-74.000		
NODE	512	-10.000	.000	-74.000		
NODE	513	10.000	.000	-74.000		
NODE	514	21.220	-11.850	-74.000		
NODE	515	21.220	11.850	-74.000		
NODE	516	32.440	.000	-74.000		
NODE	601	-34.340	-27.260	-104.000		
NODE	602	34.340	-27.260	-104.000		
NODE	603	34.340	27.260	-104.000		
NODE	604	-34.340	27.260	-104.000		
NODE	605	-10.000	-27.260	-104.000		
NODE	606	10.000	-27.260	-104.000		
NODE	607	10.000	27.260	-104.000		
NODE	608	-10.000	27.260	-104.000		
NODE	611	-34.340	.000	-104.000		
NODE	612	-10.000	.000	-104.000		
NODE	613	6.100	.000	-104.000		
NODE	614	34.340	.000	-104.000		
NODE	651	-21.676	25.408	-88.391		
NODE	652	21.676	25.408	-88.391		
NODE	653	-21.676	-25.408	-88.391		
NODE	654	21.676	-25.408	-88.391		
NODE	655	33.324	12.678	-87.952		
NODE	656	33.324	-12.678	-87.952		
NODE	701	-34.670	-27.890	-109.300	1	1 1
NODE	702	34.670	-27.890	-109.300	1	1 1
NODE	703	34.670	27.890	-109.300	1	1 1
NODE	704	-34.670	27.890	-109.300	1	1 1
NODE	1001	-37.823	-33.802	-159.300		
NODE	1002	37.823	-33.802	-159.300		
NODE	1003	37.823	33.802	-159.300		
NODE	1004	-37.823	33.802	-159.300		

NODE	2001	-40.976	-39.759	-209.300
NODE	2002	40.976	-39.759	-209.300
NODE	2003	40.976	39.759	-209.300
NODE	2004	-40.976	39.759	-209.300

	Elem ID	np1	np2	material	geom	lcoor	ecc1	ecc2
BEAM	201	201	207	2	48	21		
BEAM	202	205	208	2	48	10		
BEAM	203	201	213	2	48	23		
BEAM	206	205	214	2	55	23		
BEAM	207	205	216	2	59	27		
BEAM	208	202	219	2	59	28		
BEAM	209	206	215	2	59	29		
BEAM	210	203	220	2	59	30		
BEAM	211	202	203	2	48	23		
BEAM	212	204	209	2	48	21		
BEAM	213	206	210	2	48	10		
BEAM	214	207	214	2	58	1		
BEAM	215	207	213	2	54	1		
BEAM	216	209	213	2	54	1		
BEAM	217	209	214	2	58	1		
BEAM	218	207	212	2	62	1		
BEAM	219	209	212	2	62	1		
BEAM	220	212	213	2	61	1		
BEAM	221	212	214	2	61	1		
BEAM	222	208	216	2	62	1		
BEAM	223	215	216	2	62	1		
BEAM	224	210	215	2	62	1		
BEAM	225	211	216	2	63	1		
BEAM	226	211	215	2	63	1		
BEAM	227	205	207	2	48	1		
BEAM	228	206	209	2	48	1		
BEAM	229	202	208	2	48	1		
BEAM	230	203	210	2	48	1		
BEAM	231	211	219	2	63	1		
BEAM	232	211	220	2	63	1		
BEAM	233	208	217	2	119	1		
BEAM	234	202	217	2	120	1		
BEAM	235	216	217	2	120	1		
BEAM	236	217	219	2	119	1		
BEAM	237	203	218	2	120	1		
BEAM	238	210	218	2	119	1		
BEAM	239	215	218	2	120	1		
BEAM	240	218	220	2	119	1		
BEAM	241	219	220	2	121	1		
BEAM	242	204	213	2	48	1		

BEAM	243	206	214	2	55	1
BEAM	251	308	361	2	43	414
BEAM	252	361	204	2	43	414
BEAM	253	308	362	3	41	418
BEAM	254	362	210	3	41	418
BEAM	255	307	363	3	41	420
BEAM	256	363	203	3	41	420
BEAM	257	305	364	2	43	422
BEAM	258	364	201	2	43	422
BEAM	259	305	365	3	41	424
BEAM	260	365	208	3	41	424
BEAM	261	306	366	3	41	426
BEAM	262	366	202	3	41	426
BEAM	263	311	367	2	42	428
BEAM	264	367	201	2	42	428
BEAM	265	311	368	2	42	430
BEAM	266	368	204	2	42	430
BEAM	267	312	369	3	41	432
BEAM	268	369	207	3	41	432
BEAM	269	312	370	3	41	434
BEAM	270	370	209	3	41	434
BEAM	271	313	371	2	42	432
BEAM	272	371	208	2	42	432
BEAM	273	313	372	2	42	434
BEAM	274	372	210	2	42	434
BEAM	275	316	373	3	41	440
BEAM	276	373	202	3	41	440
BEAM	277	316	374	3	41	442
BEAM	278	374	203	3	41	442
BEAM	301	301	305	1	52	21
BEAM	302	305	306	2	50	21
BEAM	303	302	306	2	37	10
BEAM	304	301	311	2	40	23
BEAM	305	304	311	2	40	38
BEAM	306	305	311	2	39	39
BEAM	307	308	311	2	39	40
BEAM	308	305	312	2	39	23
BEAM	309	308	312	3	28	38
BEAM	310	306	312	2	45	43
BEAM	311	307	312	2	38	44
BEAM	312	306	313	2	39	23
BEAM	313	307	313	2	30	38
BEAM	314	306	314	11	58	47
BEAM	315	302	314	11	58	48
BEAM	316	313	314	11	58	49
BEAM	317	314	316	11	58	50

BEAM	318	313	316	2	53	21
BEAM	319	313	315	11	58	47
BEAM	320	315	316	11	58	53
BEAM	321	307	315	11	58	49
BEAM	322	303	315	11	58	55
BEAM	323	302	316	2	40	23
BEAM	324	303	316	2	40	38
BEAM	325	304	308	1	52	21
BEAM	326	307	308	2	50	10
BEAM	327	303	307	1	37	10
BEAM	351	305	401	2	25	161
BEAM	352	306	405	2	25	162
BEAM	353	306	402	3	21	163
BEAM	354	308	404	2	25	164
BEAM	355	307	408	2	25	165
BEAM	356	307	403	3	21	166
BEAM	357	304	411	2	39	167
BEAM	358	301	411	2	39	168
BEAM	359	308	412	2	33	169
BEAM	360	305	412	2	32	170
BEAM	361	307	413	2	33	169
BEAM	362	306	413	2	30	170
BEAM	363	303	416	2	30	173
BEAM	364	302	416	2	30	174
BEAM	401	401	405	1	47	21
BEAM	402	405	406	1	47	21
BEAM	403	402	406	2	39	10
BEAM	404	401	411	2	27	23
BEAM	405	404	411	2	27	38
BEAM	406	405	411	1	52	66
BEAM	407	408	411	2	51	67
BEAM	408	405	412	2	50	23
BEAM	409	408	412	2	39	38
BEAM	410	406	412	2	51	28
BEAM	411	407	412	2	45	30
BEAM	412	406	413	2	39	23
BEAM	413	407	413	2	36	38
BEAM	414	406	414	12	55	74
BEAM	415	402	414	12	60	66
BEAM	416	413	414	12	63	76
BEAM	417	414	416	12	55	74
BEAM	418	413	416	2	53	21
BEAM	419	413	415	12	63	74
BEAM	420	415	416	12	55	76
BEAM	421	407	415	12	55	76
BEAM	422	403	415	12	60	67

BEAM	423	402	416	2	38	23
BEAM	424	403	416	2	38	38
BEAM	425	404	408	1	47	21
BEAM	426	407	408	1	47	10
BEAM	427	403	407	2	40	10
BEAM	451	405	501	2	17	175
BEAM	452	406	505	2	25	176
BEAM	453	406	502	3	12	177
BEAM	454	408	504	2	17	178
BEAM	455	407	508	2	25	179
BEAM	456	407	503	3	12	180
BEAM	457	404	511	2	19	181
BEAM	458	401	511	2	19	182
BEAM	459	408	512	2	25	183
BEAM	460	405	512	2	38	184
BEAM	461	407	513	2	24	183
BEAM	462	406	513	2	32	184
BEAM	463	403	516	2	17	187
BEAM	464	402	516	2	17	188
BEAM	465	412	512	2	51	10
BEAM	466	413	513	2	51	10
BEAM	501	501	505	2	46	21
BEAM	502	505	506	2	51	21
BEAM	503	502	506	2	39	10
BEAM	504	501	511	2	50	23
BEAM	505	504	511	2	50	38
BEAM	506	505	511	2	51	93
BEAM	507	508	511	2	50	94
BEAM	508	505	512	2	51	23
BEAM	509	508	512	2	39	38
BEAM	510	506	512	2	51	97
BEAM	511	507	512	2	45	98
BEAM	512	506	513	2	50	23
BEAM	513	507	513	2	38	38
BEAM	514	506	514	13	54	101
BEAM	515	502	514	13	60	93
BEAM	516	513	514	13	60	103
BEAM	517	514	516	13	54	101
BEAM	518	513	516	2	55	21
BEAM	519	513	515	13	60	101
BEAM	520	515	516	13	54	103
BEAM	521	507	515	13	54	103
BEAM	522	503	515	13	60	94
BEAM	523	502	516	2	40	23
BEAM	524	503	516	2	40	38
BEAM	525	504	508	2	46	21

BEAM	526	507	508	2	51	10
BEAM	527	503	507	2	38	10
BEAM	551	501	653	2	20	191
BEAM	552	505	653	2	17	192
BEAM	553	506	605	2	27	193
BEAM	554	506	654	2	17	278
BEAM	555	502	654	2	20	194
BEAM	556	504	651	2	20	195
BEAM	557	508	651	2	17	196
BEAM	558	507	608	2	27	197
BEAM	559	507	652	2	17	198
BEAM	560	503	652	2	20	199
BEAM	561	511	604	2	18	200
BEAM	562	511	611	1	64	10
BEAM	563	511	601	2	18	202
BEAM	564	512	608	2	38	203
BEAM	565	512	612	1	64	10
BEAM	566	512	605	2	40	205
BEAM	567	513	607	2	38	203
BEAM	568	513	613	1	64	10
BEAM	569	513	606	2	40	205
BEAM	570	503	655	2	34	209
BEAM	571	516	655	3	15	210
BEAM	572	516	656	3	15	279
BEAM	573	502	656	2	34	211
BEAM	574	604	651	2	17	364
BEAM	575	608	651	2	20	365
BEAM	576	607	652	2	20	366
BEAM	577	603	652	2	17	367
BEAM	578	601	653	2	17	368
BEAM	579	605	653	2	20	369
BEAM	580	606	654	2	20	370
BEAM	581	602	654	2	17	371
BEAM	582	603	655	2	15	254
BEAM	583	614	655	2	34	253
BEAM	584	614	656	2	34	252
BEAM	585	602	656	2	15	251
BEAM	601	601	605	1	52	21
BEAM	602	605	606	1	35	21
BEAM	603	602	606	1	52	10
BEAM	604	601	611	1	56	23
BEAM	605	604	611	1	56	38
BEAM	606	605	611	2	51	119
BEAM	607	608	611	2	50	120
BEAM	608	605	612	2	51	23
BEAM	609	608	612	2	50	38

BEAM	610	606	612	1	52	123
BEAM	611	607	612	2	50	124
BEAM	612	606	613	13	42	125
BEAM	613	607	613	13	42	126
BEAM	614	606	614	13	50	127
BEAM	615	613	614	3	12	21
BEAM	616	607	614	13	45	129
BEAM	617	602	614	2	51	23
BEAM	618	603	614	2	51	38
BEAM	619	604	608	1	52	21
BEAM	620	607	608	1	35	10
BEAM	621	603	607	1	52	10
BEAM	622	612	613	1	57	21
BEAM	711	351	201	3	7	448
BEAM	712	301	351	3	7	448
BEAM	713	301	401	3	8	214
BEAM	714	401	501	3	69	214
BEAM	715	501	601	8	70	214
BEAM	716	601	701	2	70	217
BEAM	721	352	202	3	8	450
BEAM	722	302	352	3	8	450
BEAM	723	302	402	3	9	219
BEAM	724	402	502	3	68	219
BEAM	725	502	602	8	70	219
BEAM	726	602	702	2	70	281
BEAM	731	353	203	3	8	446
BEAM	732	303	353	3	8	446
BEAM	733	303	403	3	10	223
BEAM	734	403	503	3	69	223
BEAM	735	503	603	8	70	223
BEAM	736	603	703	2	70	226
BEAM	741	354	204	3	7	412
BEAM	742	304	354	3	7	412
BEAM	743	304	404	3	9	229
BEAM	744	404	504	3	69	229
BEAM	745	504	604	8	70	229
BEAM	746	604	704	2	70	232
BEAM	751	355	207	3	13	452
BEAM	752	305	355	3	13	452
BEAM	753	305	405	2	9	38
BEAM	754	405	505	2	9	38
BEAM	755	505	605	1	9	38
BEAM	761	356	208	3	13	452
BEAM	762	306	356	3	13	452
BEAM	763	306	406	2	9	38
BEAM	764	406	506	2	9	38

BEAM	765	506	606	1	9	38
BEAM	771	357	210	3	13	416
BEAM	772	307	357	3	13	416
BEAM	773	307	407	9	9	38
BEAM	774	407	507	9	9	38
BEAM	775	507	607	9	9	38
BEAM	781	358	209	3	13	416
BEAM	782	308	358	3	13	416
BEAM	783	308	408	9	9	38
BEAM	784	408	508	9	9	38
BEAM	785	508	608	9	9	38
BEAM	801	217	313	14	83	372
BEAM	802	217	314	14	83	10
BEAM	803	218	315	14	83	10
BEAM	804	218	316	14	83	233
BEAM	805	313	413	14	83	38
BEAM	806	314	414	14	83	256
BEAM	807	315	415	14	83	257
BEAM	808	316	416	14	83	258
BEAM	809	413	513	14	83	38
BEAM	810	414	514	14	83	260
BEAM	811	415	515	14	83	261
BEAM	812	416	516	14	83	258
BEAM	813	513	613	14	83	263
BEAM	814	514	613	14	83	264
BEAM	815	514	614	14	83	265
BEAM	816	515	613	14	83	266
BEAM	817	515	614	14	83	267
BEAM	818	516	614	14	83	258
BEAM	819	513	613	4	70	263
BEAM	820	516	614	3	71	258
BEAM	51057	51	57	16	55	1
BEAM	51061	51	61	16	55	1
BEAM	51101	51	101	16	15	10
BEAM	51115	51	115	16	43	140
BEAM	51121	51	121	16	43	14
BEAM	51131	51	131	16	43	24
BEAM	51136	51	136	16	55	1
BEAM	51140	51	140	16	55	1
BEAM	52058	52	58	16	55	1
BEAM	52063	52	63	16	55	1
BEAM	52102	52	102	16	15	10
BEAM	52117	52	117	16	43	25
BEAM	52127	52	127	16	43	14
BEAM	52134	52	134	16	43	159
BEAM	52138	52	138	16	55	1

BEAM	52143	52	143	16	55	1
BEAM	53060	53	60	16	55	1
BEAM	53063	53	63	16	55	1
BEAM	53103	53	103	16	15	10
BEAM	53120	53	120	16	43	25
BEAM	53128	53	128	16	43	12
BEAM	53135	53	135	16	43	159
BEAM	53139	53	139	16	55	1
BEAM	53143	53	143	16	55	1
BEAM	54059	54	59	16	55	1
BEAM	54061	54	61	16	55	1
BEAM	54104	54	104	16	15	10
BEAM	54118	54	118	16	43	140
BEAM	54122	54	122	16	43	12
BEAM	54133	54	133	16	43	24
BEAM	54137	54	137	16	55	1
BEAM	54140	54	140	16	55	1
BEAM	57058	57	58	16	55	1
BEAM	57061	57	61	16	55	1
BEAM	57062	57	62	16	55	1
BEAM	57107	57	107	16	15	10
BEAM	57115	57	115	16	43	25
BEAM	57116	57	116	16	43	11
BEAM	57123	57	123	16	43	14
BEAM	57141	57	141	16	55	1
BEAM	58062	58	62	16	55	1
BEAM	58063	58	63	16	55	1
BEAM	58108	58	108	16	15	10
BEAM	58116	58	116	16	43	136
BEAM	58117	58	117	16	43	140
BEAM	58125	58	125	16	43	14
BEAM	58142	58	142	16	55	1
BEAM	59060	59	60	16	55	1
BEAM	59061	59	61	16	55	1
BEAM	59062	59	62	16	55	1
BEAM	59109	59	109	16	15	10
BEAM	59118	59	118	16	43	25
BEAM	59119	59	119	16	43	11
BEAM	59124	59	124	16	43	12
BEAM	59141	59	141	16	55	1
BEAM	60062	60	62	16	55	1
BEAM	60063	60	63	16	55	1
BEAM	60110	60	110	16	15	10
BEAM	60119	60	119	16	43	136
BEAM	60120	60	120	16	43	140
BEAM	60126	60	126	16	43	12

BEAM	60142	60	142	16	55	1
BEAM	101115	101	115	16	55	1
BEAM	101121	101	121	16	55	1
BEAM	101131	101	131	16	55	1
BEAM	101144	101	144	16	55	1
BEAM	102117	102	117	16	55	1
BEAM	102127	102	127	16	55	1
BEAM	102134	102	134	16	55	1
BEAM	102146	102	146	16	55	1
BEAM	103120	103	120	16	55	1
BEAM	103128	103	128	16	55	1
BEAM	103135	103	135	16	55	1
BEAM	103146	103	146	16	55	1
BEAM	104118	104	118	16	55	1
BEAM	104122	104	122	16	55	1
BEAM	104133	104	133	16	55	1
BEAM	104144	104	144	16	55	1
BEAM	107115	107	115	16	55	1
BEAM	107116	107	116	16	55	1
BEAM	107123	107	123	16	55	1
BEAM	107144	107	144	16	55	1
BEAM	107145	107	145	16	55	1
BEAM	108116	108	116	16	55	1
BEAM	108117	108	117	16	55	1
BEAM	108125	108	125	16	55	1
BEAM	108145	108	145	16	55	1
BEAM	108146	108	146	16	55	1
BEAM	109118	109	118	16	55	1
BEAM	109119	109	119	16	55	1
BEAM	109124	109	124	16	55	1
BEAM	109144	109	144	16	55	1
BEAM	109145	109	145	16	55	1
BEAM	110119	110	109	16	55	1
BEAM	110120	110	120	16	55	1
BEAM	110126	110	126	16	55	1
BEAM	110145	110	145	16	55	1
BEAM	110146	110	146	16	55	1
BEAM	111121	111	121	16	55	1
BEAM	111122	111	122	16	55	1
BEAM	111130	111	130	16	55	1
BEAM	111132	111	132	16	55	1
BEAM	112123	112	123	16	55	1
BEAM	112124	112	124	16	55	1
BEAM	113125	113	125	16	55	1
BEAM	113126	113	126	16	55	1
BEAM	114127	114	127	16	55	1

BEAM	114128	114	128	16	55	1
BEAM	114134	114	134	16	55	1
BEAM	114135	114	135	16	55	1
BEAM	130131	130	131	16	55	1
BEAM	130132	130	132	16	55	1
BEAM	132133	132	133	16	55	1
BEAM	134135	134	135	16	55	1
BEAM	136130	136	130	16	43	143
BEAM	136131	136	131	16	43	10
BEAM	136137	136	137	16	55	1
BEAM	137132	137	132	16	43	143
BEAM	137133	137	133	16	43	10
BEAM	138134	138	134	16	43	10
BEAM	138139	138	139	16	55	1
BEAM	139135	139	135	16	43	10
BEAM	140111	140	111	16	43	10
BEAM	140121	140	121	16	43	12
BEAM	140122	140	122	16	43	14
BEAM	141112	141	112	16	43	10
BEAM	141123	141	123	16	43	12
BEAM	141124	141	124	16	43	14
BEAM	142113	142	113	16	43	10
BEAM	142125	142	125	16	43	12
BEAM	142126	142	126	16	43	14
BEAM	143114	143	114	16	43	10
BEAM	143127	143	127	16	43	12
BEAM	143128	143	128	16	43	14
BEAM	201051	201	51	17	15	10
BEAM	201140	201	140	16	43	154
BEAM	202052	202	52	17	15	10
BEAM	202143	202	143	16	43	154
BEAM	203053	203	53	17	15	10
BEAM	203143	203	143	16	43	153
BEAM	204054	204	54	17	15	10
BEAM	204140	204	140	16	43	153
BEAM	207057	207	57	17	15	10
BEAM	207141	207	141	16	43	154
BEAM	208058	208	58	17	15	10
BEAM	208142	208	142	16	43	154
BEAM	209059	209	59	17	15	10
BEAM	209141	209	141	16	43	153
BEAM	210060	210	60	17	15	10
BEAM	210142	210	142	16	43	153

'	Elem ID	np1	material	lcoor	ecc1
'	SPRNG2GR	901	701	19	

SPRNG2GR	902	702	19
SPRNG2GR	903	703	19
SPRNG2GR	904	704	19

	Geom ID	Do	Thick	Shear_y	Shear_z
PIPE	7	1.800	.075		
PIPE	8	1.800	.070		
PIPE	9	1.800	.067		
PIPE	10	1.800	.062		
PIPE	12	1.650	.060		
PIPE	13	1.600	.075		
PIPE	15	1.600	.060		
PIPE	17	1.600	.050		
PIPE	18	1.600	.045		
PIPE	19	1.500	.045		
PIPE	20	1.500	.040		
PIPE	21	1.400	.060		
PIPE	24	1.400	.045		
PIPE	25	1.400	.040		
PIPE	27	1.400	.030		
PIPE	28	1.300	.060		
PIPE	30	1.300	.050		
PIPE	32	1.300	.040		
PIPE	33	1.300	.035		
PIPE	34	1.300	.030		
PIPE	35	1.300	.025		
PIPE	36	1.200	.050		
PIPE	37	1.200	.025		
PIPE	38	1.200	.040		
PIPE	39	1.200	.035		
PIPE	40	1.200	.030		
PIPE	41	1.100	.055		
PIPE	42	1.100	.050		
PIPE	43	1.100	.045		
PIPE	45	1.100	.035		
PIPE	46	1.100	.030		
PIPE	47	1.100	.025		
PIPE	48	1.000	.045		
PIPE	50	1.000	.035		
PIPE	51	1.000	.030		
PIPE	52	1.000	.025		
PIPE	53	.900	.040		
PIPE	54	.900	.035		
PIPE	55	.900	.030		
PIPE	56	.900	.025		
PIPE	57	.900	.020		

PIPE	58	.800	.040
PIPE	59	.800	.035
PIPE	60	.800	.030
PIPE	61	.800	.025
PIPE	62	.700	.020
PIPE	63	.650	.030
PIPE	64	.600	.020
PIPE	68	3.000	.069
PIPE	69	3.000	.060
PIPE	70	3.250	.125
PIPE	71	2.000	.070
PIPE	83	.55000	.00100
PIPE	119	.700	.050
PIPE	120	.500	.055
PIPE	121	.700	.045

	Mat ID	E-mod	Poiss	Yield	Density	Thermal	
MISOIEP	1	2.100E+11	3.000E-01	3.550E+08	7.850E+03	.000E+00	
MISOIEP	2	2.100E+11	3.000E-01	3.400E+08	7.850E+03	.000E+00	
MISOIEP	3	2.100E+11	3.000E-01	3.200E+08	7.850E+03	.000E+00	
MISOIEP	4	2.100E+11	3.000E-01	3.100E+08	7.850E+03	.000E+00	
MISOIEP	8	2.100E+11	3.000E-01	3.100E+08	4.830E+03	.000E+00	
MISOIEP	9	2.100E+11	3.000E-01	3.400E+08	1.480E+04	.000E+00	
MISOIEP	11	2.100E+11	3.000E-01	3.400E+08	1.470E+04	.000E+00	
MISOIEP	12	2.100E+11	3.000E-01	3.400E+08	1.260E+04	.000E+00	
MISOIEP	13	2.100E+11	3.000E-01	3.400E+08	1.170E+04	.000E+00	
MISOIEP	14	2.100E+09	3.000E-01	1.000E+20	.000E+00	.000E+00	
MISOIEP	16	2.100E+11	3.000E-01	1.000E+15	.000E+00	.000E+00	
MISOIEP	17	2.100E+11	3.000E-01	3.100E+08	.000E+00	.000E+00	

	Loc-Coo	dx	dy	dz
UNITVEC	1	.000	.000	1.000
UNITVEC	10	.000	1.000	.000
UNITVEC	11	-.698	.000	.716
UNITVEC	12	.000	.822	.569
UNITVEC	14	.000	-.822	.569
UNITVEC	21	.000	-1.000	.000
UNITVEC	23	1.000	.000	.000
UNITVEC	24	.812	.000	.583
UNITVEC	25	.754	.000	.657
UNITVEC	27	.707	-.707	.000
UNITVEC	28	.707	.707	.000
UNITVEC	29	-.707	-.707	.000
UNITVEC	30	-.707	.707	.000

UNITVEC	38	-1.000	.000	.000
UNITVEC	39	.664	.747	.000
UNITVEC	40	-.664	.747	.000
UNITVEC	43	.638	.770	.000
UNITVEC	44	-.638	.770	.000
UNITVEC	47	.665	-.747	.000
UNITVEC	48	.664	.748	.000
UNITVEC	49	-.665	-.747	.000
UNITVEC	50	.664	-.748	.000
UNITVEC	53	-.664	-.748	.000
UNITVEC	55	-.664	.748	.000
UNITVEC	66	.699	.715	.000
UNITVEC	67	-.699	.715	.000
UNITVEC	74	.699	-.715	.000
UNITVEC	76	-.699	-.715	.000
UNITVEC	93	.726	.688	.000
UNITVEC	94	-.726	.688	.000
UNITVEC	97	.764	.645	.000
UNITVEC	98	-.764	.645	.000
UNITVEC	101	.726	-.688	.000
UNITVEC	103	-.726	-.688	.000
UNITVEC	119	.746	.666	.000
UNITVEC	120	-.746	.666	.000
UNITVEC	123	.806	.592	.000
UNITVEC	124	-.806	.592	.000
UNITVEC	125	.990	.142	.000
UNITVEC	126	-.990	.142	.000
UNITVEC	127	.746	-.666	.000
UNITVEC	129	-.746	-.666	.000
UNITVEC	136	.698	.000	.716
UNITVEC	140	-.754	.000	.657
UNITVEC	143	.890	.000	.456
UNITVEC	153	.000	.694	.720
UNITVEC	154	.000	-.694	.720
UNITVEC	159	-.812	.000	.583
UNITVEC	161	-.819	.068	.570
UNITVEC	162	-.825	.067	.561
UNITVEC	163	-.819	-.068	-.570
UNITVEC	164	-.819	-.068	.570
UNITVEC	165	-.825	-.067	.561
UNITVEC	166	-.819	.068	-.570
UNITVEC	167	-.031	.869	-.495
UNITVEC	168	.031	.869	.495
UNITVEC	169	.000	.868	-.496
UNITVEC	170	.000	.868	.496
UNITVEC	173	.031	.869	-.495

UNITVEC	174	-.031	.869	.495
UNITVEC	175	-.812	.069	.580
UNITVEC	176	-.842	.064	.536
UNITVEC	177	-.812	-.069	-.580
UNITVEC	178	-.812	-.069	.580
UNITVEC	179	-.842	-.064	.536
UNITVEC	180	-.812	.069	-.580
UNITVEC	181	-.034	.841	-.541
UNITVEC	182	.034	.841	.541
UNITVEC	183	.000	.840	-.543
UNITVEC	184	.000	.840	.543
UNITVEC	187	.034	.841	-.541
UNITVEC	188	-.034	.841	.541
UNITVEC	191	-.803	-.070	-.592
UNITVEC	192	-.779	.074	.623
UNITVEC	193	-.834	.065	.548
UNITVEC	194	-.803	.070	.592
UNITVEC	195	-.803	.070	-.592
UNITVEC	196	-.779	-.074	.623
UNITVEC	197	-.834	-.065	.548
UNITVEC	198	-.779	.074	-.623
UNITVEC	199	-.803	-.070	.592
UNITVEC	200	.042	.741	.670
UNITVEC	202	-.042	.741	-.670
UNITVEC	203	.000	.740	.673
UNITVEC	205	.000	.740	-.673
UNITVEC	209	.039	.785	-.618
UNITVEC	210	-.042	.741	.670
UNITVEC	211	-.039	.785	.618
UNITVEC	214	-.998	.007	.062
UNITVEC	217	-.998	.007	.061
UNITVEC	219	-.998	-.007	-.062
UNITVEC	223	-.998	.007	-.062
UNITVEC	226	-.998	.007	-.061
UNITVEC	229	-.998	-.007	.062
UNITVEC	232	-.998	-.007	.061
UNITVEC	233	.674	-.599	.433
UNITVEC	251	.051	-.737	.673
UNITVEC	252	.063	.781	.621
UNITVEC	253	.063	-.781	.621
UNITVEC	254	.051	.737	.673
UNITVEC	256	-.999	-.002	-.032
UNITVEC	257	-.999	.002	-.032
UNITVEC	258	-.998	.000	-.063
UNITVEC	260	-1.000	-.002	-.031
UNITVEC	261	-1.000	.002	-.031

UNITVEC	263	-.992	.000	.129
UNITVEC	264	-.905	-.156	.395
UNITVEC	265	-.926	.138	-.350
UNITVEC	266	-.905	.156	.395
UNITVEC	267	-.926	-.138	-.350
UNITVEC	278	-.779	-.074	-.623
UNITVEC	279	.042	.741	-.670
UNITVEC	281	-.998	-.007	-.061
UNITVEC	364	-.765	.112	.634
UNITVEC	365	.787	.125	.604
UNITVEC	366	-.787	.125	.604
UNITVEC	367	.765	.112	.634
UNITVEC	368	-.765	-.112	.634
UNITVEC	369	.787	-.125	.604
UNITVEC	370	-.787	-.125	.604
UNITVEC	371	.765	-.112	.634
UNITVEC	372	-.674	.599	.433
UNITVEC	412	-.468	.874	.133
UNITVEC	414	.820	.148	.553
UNITVEC	416	.000	.993	.118
UNITVEC	418	-.780	.120	.614
UNITVEC	420	-.820	.148	.553
UNITVEC	422	.820	-.148	.553
UNITVEC	424	-.780	-.120	.614
UNITVEC	426	-.820	-.148	.553
UNITVEC	428	-.108	.880	.464
UNITVEC	430	-.108	-.880	.464
UNITVEC	432	.000	.887	.461
UNITVEC	434	.000	-.887	.461
UNITVEC	440	.108	.880	.464
UNITVEC	442	.108	-.880	.464
UNITVEC	446	.468	.874	.133
UNITVEC	448	-.468	-.874	.133
UNITVEC	450	.468	-.874	.133
UNITVEC	452	.000	-.993	.118

' Ecc-ID Ex Ey Ez

' Mat ID SPRING CHAR.
 SPRIDIAG 19 6.72000E+09 6.72000E+09 2.07600E+10
 5.42000E+11 5.42000E+11 2.02000E+11

' Load Case Node ID LOAD INTENSITY
 NODELOAD 1 101 .00000E+00 .00000E+00 -1.94880E+07
 NODELOAD 1 102 .00000E+00 .00000E+00 -1.59120E+07

NODELOAD	1	103	.00000E+00	.00000E+00	-1.59120E+07
NODELOAD	1	104	.00000E+00	.00000E+00	-1.94880E+07
NODELOAD	1	107	.00000E+00	.00000E+00	-1.50000E+07
NODELOAD	1	108	.00000E+00	.00000E+00	-1.50000E+07
NODELOAD	1	109	.00000E+00	.00000E+00	-1.50000E+07
NODELOAD	1	110	.00000E+00	.00000E+00	-1.50000E+07
NODELOAD	1	111	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	112	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	113	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	114	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	115	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	116	.00000E+00	.00000E+00	-1.29000E+07
NODELOAD	1	117	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	118	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	119	.00000E+00	.00000E+00	-1.29000E+07
NODELOAD	1	120	.00000E+00	.00000E+00	-7.50000E+06
NODELOAD	1	130	.00000E+00	.00000E+00	-1.70000E+06
NODELOAD	1	132	.00000E+00	.00000E+00	-1.70000E+06
NODELOAD	1	134	.00000E+00	.00000E+00	-1.70000E+06
NODELOAD	1	135	.00000E+00	.00000E+00	-1.70000E+06
NODELOAD	1	201	.00000E+00	.00000E+00	-2.76000E+05
NODELOAD	1	202	.00000E+00	.00000E+00	-1.48000E+05
NODELOAD	1	203	.00000E+00	.00000E+00	-1.48000E+05
NODELOAD	1	204	.00000E+00	.00000E+00	-4.17000E+05
NODELOAD	1	205	.00000E+00	.00000E+00	-8.07000E+05
NODELOAD	1	206	.00000E+00	.00000E+00	-2.93000E+05
NODELOAD	1	301	.00000E+00	.00000E+00	-2.10000E+05
NODELOAD	1	302	.00000E+00	.00000E+00	-2.10000E+05
NODELOAD	1	303	.00000E+00	.00000E+00	-2.10000E+05
NODELOAD	1	304	.00000E+00	.00000E+00	-2.10000E+05
NODELOAD	1	305	.00000E+00	.00000E+00	-7.04000E+05
NODELOAD	1	306	.00000E+00	.00000E+00	-1.85000E+05
NODELOAD	1	307	.00000E+00	.00000E+00	-3.46000E+05
NODELOAD	1	308	.00000E+00	.00000E+00	-3.70000E+05
NODELOAD	1	311	.00000E+00	.00000E+00	-8.79000E+05
NODELOAD	1	312	.00000E+00	.00000E+00	-8.72000E+05
NODELOAD	1	401	.00000E+00	.00000E+00	-1.27500E+06
NODELOAD	1	402	.00000E+00	.00000E+00	-1.27500E+06
NODELOAD	1	403	.00000E+00	.00000E+00	-1.27500E+06
NODELOAD	1	404	.00000E+00	.00000E+00	-1.27500E+06
NODELOAD	1	405	.00000E+00	.00000E+00	-6.84000E+05
NODELOAD	1	406	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	407	.00000E+00	.00000E+00	-4.37000E+05
NODELOAD	1	408	.00000E+00	.00000E+00	-3.76000E+05
NODELOAD	1	411	.00000E+00	.00000E+00	-4.82000E+05
NODELOAD	1	412	.00000E+00	.00000E+00	-7.84000E+05

NODELOAD	1	501	.00000E+00	.00000E+00	-1.35600E+06
NODELOAD	1	502	.00000E+00	.00000E+00	-1.35600E+06
NODELOAD	1	503	.00000E+00	.00000E+00	-1.35600E+06
NODELOAD	1	504	.00000E+00	.00000E+00	-1.35600E+06
NODELOAD	1	505	.00000E+00	.00000E+00	-7.10000E+05
NODELOAD	1	506	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	507	.00000E+00	.00000E+00	-4.27000E+05
NODELOAD	1	508	.00000E+00	.00000E+00	-3.75000E+05
NODELOAD	1	511	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	512	.00000E+00	.00000E+00	-8.74000E+05
NODELOAD	1	513	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	516	.00000E+00	.00000E+00	-7.35000E+05
NODELOAD	1	601	.00000E+00	.00000E+00	-5.29200E+06
NODELOAD	1	602	.00000E+00	.00000E+00	-5.29200E+06
NODELOAD	1	603	.00000E+00	.00000E+00	-5.29200E+06
NODELOAD	1	604	.00000E+00	.00000E+00	-5.29200E+06
NODELOAD	1	605	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	606	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	607	.00000E+00	.00000E+00	-2.61000E+05
NODELOAD	1	608	.00000E+00	.00000E+00	-1.98000E+05
NODELOAD	1	611	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	612	.00000E+00	.00000E+00	-4.68000E+05
NODELOAD	1	613	.00000E+00	.00000E+00	-1.15000E+05
NODELOAD	1	614	.00000E+00	.00000E+00	-3.92000E+05
NODELOAD	1	701	.00000E+00	.00000E+00	-1.17620E+07
NODELOAD	1	702	.00000E+00	.00000E+00	-1.17620E+07
NODELOAD	1	703	.00000E+00	.00000E+00	-1.17620E+07
NODELOAD	1	704	.00000E+00	.00000E+00	-1.17620E+07

Load Case Elem ID L O A D I N T E N S I T Y

Load Case Elem ID Press1 Press2 Press3 Press4

Load Case Acc_X Acc_Y Acc_Z

GRAVITY 1 .0000E+00 .0000E+00 -9.8100E+00

Appendix F. Example of Constructing a Random Set Model

As described in Chapter 3, the random set which is also sometimes referred to as a Dempster-Shafer structure consists of a finite number of focal sets. For example, for discrete random variable x_i in the space X . Random sets are a collection of many imprecise observations $\mathfrak{F} = \{A_i : i = 1, \dots, n\}$ of the given fundamental set X and its probability weight mapping function:

$$m: \mathfrak{F} \rightarrow [0,1] \quad (\text{F.1})$$

where $m_i = m(A_i)$ $\sum m(A_i) = 1$. This gives a measure of the degree of confidence for the observations of X . For the occurrence event E in the space of X , the plausibility measure of $E \in X$ is defined by:

$$\text{Pls}(E) = \sum_{A_i \cap E \neq \emptyset} m(A_i) \quad (\text{F.2})$$

and the belief function is given by:

$$\text{Bel}(E) = \sum_{A_i \subseteq E} m(A_i) \quad (\text{F.3})$$

These plausibility and the belief functions can be interpreted as a prescription of a set of probability function. Or in other words, it gives the upper and lower probabilities of a certain set of probability distributions. For demonstration purpose, the following shows an example of how to construct the random set model.

Random Set Example

The uncertainty of X is represented by three random focal sets as:

$$A_1 = [2, 4], \quad A_2 = [3, 8], \quad A_3 = [1, 5]$$

and associated weights are given by:

$$m(A_1) = 0.2, \quad m(A_2) = 0.3, \quad m(A_3) = 0.5$$

These are presented in Fig. F.1 (a) and the probability weights which are represented by the three bins are also illustrated in this figure, see Fig. F.1 (b).

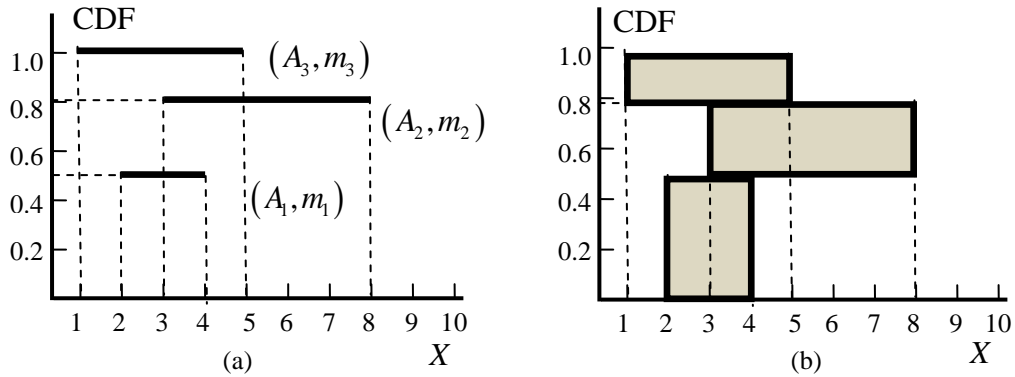


Figure F.1 Random set (a) and its assigned probability weight (b).

The imprecise probability (probability box) used to represent the random set is illustrated in Fig. F.2 (a) where the focal sets are plotted as a stack and the height of the lines is determined by the cumulative sum of weights (illustrated by the dotted line). The random set can also be visualized by its contour function which assigns each singleton x its plausibility $Pls(x)$. The value is obtained by

adding all the probability $m(A_i)$ of those focal sets A_i to which x belongs. This is illustrated in Fig. F.2 (b).

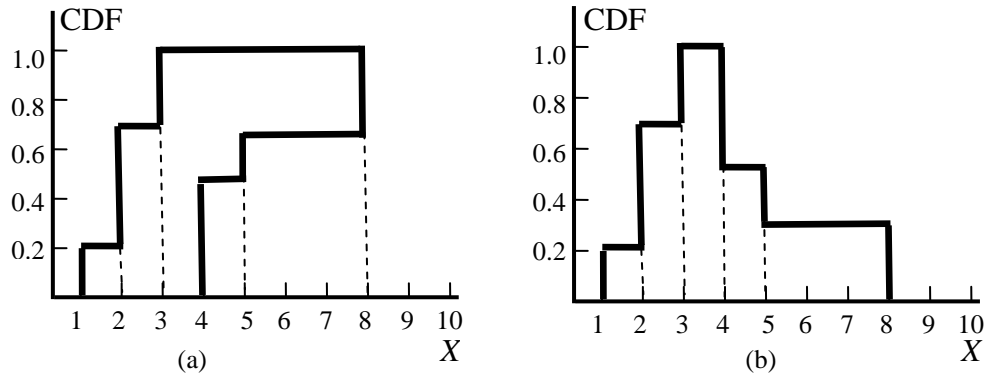


Figure F.2 Probability box (a) and contour function (b).

Appendix G. Information of Selected Wave Data and Basic Linear Wave Theory

G.1 Wave Information Studies

The wave data used in this thesis are taken from the Wave Information Studies (WIS) which is a US Army Corps of Engineers (USACE) sponsored project that generates consistent, hourly, long-term (20+ years) wave climatologies along all US coastlines, including the Great Lakes and US island territories. The WIS program originated in the Great Lakes in the mid 1970s and migrated to the Atlantic, Gulf of Mexico and Pacific Oceans. The currently available domains are depicted in Fig. G.1. The official website of WIS is at <http://wis.usace.army.mil/>.

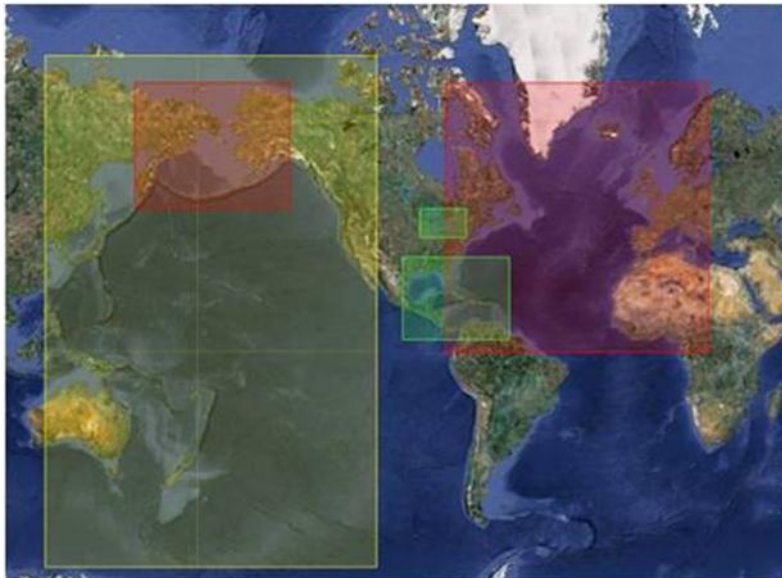


Figure G.1 WIS data domain.

This site provides access to the database of wave information for a densely-spaced series of wave gauges in water depths of 15-20 m and for a less-dense series in deeper water (100 m or more). Data available from each site include hourly wind speed, wind direction, and bulk wave parameters (significant wave height, period, and direction). Discrete directional wave spectra at 1 to 3-hour intervals are also available. A suite of tabular and graphic products for each location is also provided in the website.

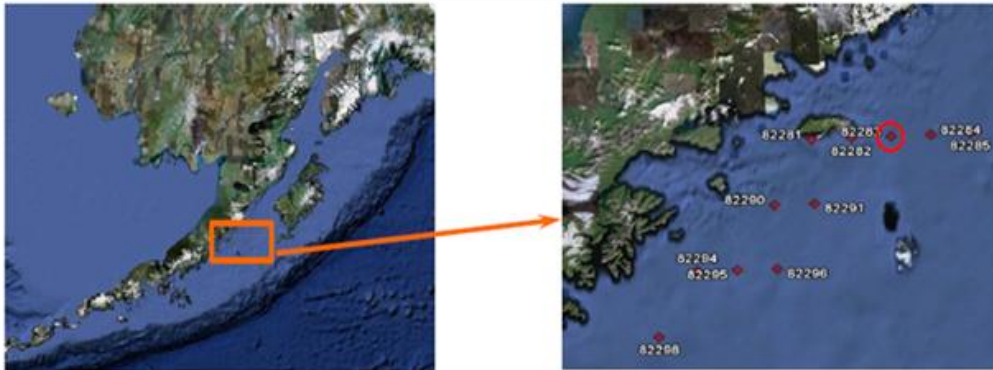


Figure G.2 Geological location of the selected buoy.

In this thesis, the set of data is taken from a buoy (No. 82283) locating in the south coast of Alaska (56.5°N 203.25°E), see Fig. G.2. The water depth at this location is 124m. The selected data set containing 25 years of hourly wave records (1985/1/1 01:00 to 2010/1/1 01:00) is filtered. In Chapter 3, the significant wave height H_S time series data are utilized. In Chapter 4, the record of significant wave height H_S , wave observation time t and wave directions θ_S are been used. In Chapter 5, a set of ocean parameters are been used, these include significant wave height H_S , peak wave period T_P , wave direction θ_S , wind speed V_W and wind direction θ_W .

G.2 Linear Wave Theory

The basic wave theory for analyzing an offshore structure relevant to the current study is presented. Detailed information could be found from the design codes (DNV 2007) and books (Sarpkaya and Isaacson 1981).

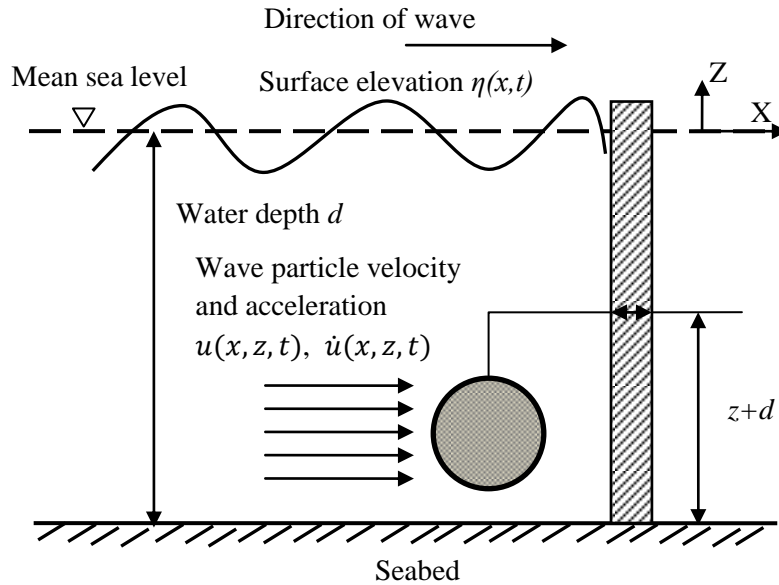


Figure G.3 Regular wave propagation properties (Sarpkaya and Isaacson 1981).

The linear wave theory is the most common way of representing the properties of a wave particle. The theory generally states that when the wave height H is much smaller compared to the wave length λ and still water depth d , the Airy wave theory may be adopted to give a linearised description of the propagation of gravity waves on the surface of a homogeneous fluid layer. The surface elevation of a wave, denoted as $\eta(x, t)$, can be expressed in terms of the

amplitude $a=H/2$, time t and spatial location x in the direction of wave propagation by:

$$\eta(x,t) = a \cos(\kappa x - \omega t) \quad (\text{G.1})$$

where $\kappa=2\pi/\lambda$ is the wave number and $\omega=2\pi/T$ is the angular frequency of the wave, in which T is the wave period. Based on the solution of the Laplace equation in terms of the velocity potential and the use of linearized boundary condition, the horizontal water particle velocity $u(x,z,t)$ and acceleration $\dot{u}(x,z,t)$ at depth z (for $z \leq 0$) measured from the mean sea level can be obtained by the derivative to the potential function which are expressed as:

$$u(x,z,t) = a\omega \frac{\cosh(\kappa(z+d))}{\sinh(\kappa d)} \cos(\kappa x - \omega t) \quad (\text{G.2})$$

$$\dot{u}(x,z,t) = a\omega^2 \frac{\cosh(\kappa(z+d))}{\sinh(\kappa d)} \sin(\kappa x - \omega t) \quad (\text{G.3})$$

In deep water, when $\kappa d > \pi$, the above formulae can be approximated by:

$$u(x,z,t) = a\omega e^{\kappa z} \cos(\kappa x - \omega t) \quad (\text{G.4})$$

$$\dot{u}(x,z,t) = a\omega^2 e^{\kappa z} \sin(\kappa x - \omega t) \quad (\text{G.5})$$

Usually, the ocean data are collected based on measurements at regular intervals (e.g. hourly record). The random behavior of ocean waves can be described by a statistical model. The variation of the physical variable within the time interval is assumed to be stationary and can be described by a stochastic

process. Based on the property of a zero-mean stationary process, the property of a sea state within such a short period can be conveniently represented in the frequency domain by the wave spectrum $S(\omega)$. Several well-established wave spectra are available in the literature. The most commonly applied wave spectrums are the *Pierson-Moskowitz* (P-M) spectrum and the *JONSWAP* spectrum (DNV 2007). The *Pierson-Moskowitz* (P-M) spectrum is given by:

$$S_{PM}(\omega) = \frac{5}{16} \cdot H_S^2 \omega_P^4 \cdot \omega^{-5} \exp \left[-\frac{5}{4} \left(\frac{\omega}{\omega_P} \right)^{-4} \right] \quad (\text{G.6})$$

where H_S is the significant wave height, T_P is the peak period, $\omega_P = 2\pi/T$ is the angular spectral peak frequency. *JONSWAP* spectrum $S_J(\omega)$ is formulated as a modification to P-M spectrum with the consideration of a developing sea state in a fetch limited situation:

$$S_J(\omega) = A_\gamma \cdot S_{PM}(\omega) \cdot \gamma^{\exp \left[-0.5 \left(\frac{\omega - \omega_P}{\sigma \omega_P} \right)^2 \right]} \quad (\text{G.7})$$

where σ is the spectral width parameter evaluated on the values of ω and ω_P :

$$\sigma = \begin{cases} 0.07 & \text{for } \omega \leq \omega_P \\ 0.09 & \text{for } \omega > \omega_P \end{cases} \quad (\text{G.8})$$

$A_\gamma = 1 - 0.287 \ln(\gamma)$ is the normalizing factor, γ is the non-dimensional peak shape parameter and the following value can be applied:

$$\gamma = \begin{cases} 5 & \text{for } \frac{T_p}{\sqrt{H_s}} \leq 3.6 \\ \exp\left(5.75 - 1.15 \frac{T_p}{\sqrt{H_s}}\right) & \text{for } 3.6 < \frac{T_p}{\sqrt{H_s}} < 5 \\ 1 & \text{for } 5 \leq \frac{T_p}{\sqrt{H_s}} \end{cases} \quad (\text{G.9})$$

Once the wave spectrum is established, the irregular random waves in the short term stationary sea state can be easily represented by a summation of sinusoidal wave components. The simulation of the surface elevation in time domain could be obtained by a superposition of large number of independent linear waves corresponding to different amplitudes, frequencies and arbitrary phase angles. The series representation of the elevation corresponding to a wave spectrum is given by:

$$\eta(x, t) = \sum_{i=1}^N a_i \cos(\kappa_i x - \omega_i t + \theta_i) \quad (\text{G.10})$$

where κ_i and ω_i are the wave number and discrete frequency which have the same meaning as in Eq. (5.26). θ_i are random phases, uniformly distributed between 0 and 2π , mutually independent of each other and of the amplitude a_i . The amplitude a_i of the i th component, which is Rayleigh distributed, is given by

$$a_i = \sqrt{2S(\omega_i)\Delta\omega_i} \quad (\text{G.11})$$

where $S(\omega_i)$ is the value of the i th component in the established wave spectrum and $\Delta\omega_i = \omega_i - \omega_{i-1}$ is the difference between successive frequencies. To simulate the random waves accurately, the simulation must contain enough

number of wave components. Or in other words, the value of $\Delta\omega_i$ needs to be small. The linear wave representation also allows the superposition of the water particle velocity and acceleration components. The water horizontal velocity and acceleration for this simulated sea state is given by:

$$u(x, z, t) = \sum_{i=1}^N \omega_i \sqrt{2S(\omega_i) \Delta\omega_i} \frac{\cosh(\kappa_i(z+d))}{\sinh(\kappa_i d)} \cos(\kappa_i x - \omega_i t + \theta_i) \quad (\text{G.12})$$

$$\dot{u}(x, z, t) = \sum_{i=1}^N \omega_i^2 \sqrt{2S(\omega_i) \Delta\omega_i} \frac{\cosh(\kappa_i(z+d))}{\sinh(\kappa_i d)} \sin(\kappa_i x - \omega_i t + \theta_i) \quad (\text{G.13})$$

which is employed as the basic simulation technique in the calculation of random wave loading to be used to compute the hydrodynamic response of offshore structures.