ACKNOWLEDGEMENTS

First of all, I would like to thank dr.ir. Leo H. Holtuijsen, my thesis project supervisor at *Technische Universiteit Delft* (TU Delft), for giving me the chance to carry out this research which will be presented at the 2008 ICCE, and, above all, for his daily theoretical, practical and personal support. Thanks are also given to dr.ir Pieter H.A.J.M. Van Gelder for his valuable assistance.

I also want to express my gratitude to all the staff at the Department of Hydraulic Engineering at TU Delft, especially the Section of Environmental Fluid Mechanics, for their pleasant welcome and help throughout my work.

I acknowledge Dr. Joan Pau Sierra Pedrico from *Universitat Politècnica de Catalunya* (UPC), for accepting to be my supervisor in Barcelona.

Furthermore, I am very grateful to the people who have provided me with data, without which this study would not have been possible; firstly to all the staff at *Laboratori d'Enginyeria Marítima* of my home university (UPC), especially Jesús Gómez Aguar, who supplied the data for the four buoys off the Catalan coast, which are managed by XIOM (*Xarxa d'Instruments Oceanogràfics i Meteorològics de la Generalitat de Catalunya*). Secondly, I thank Ph. D. Stephen Barstow, senior ocean wave climatologist, who gave me permission to use some of the WADIC project storm data, in Norway.

Last but not least, I need to express my best deepest appreciation to the invaluable support of my partner, Iván, and my parents for their constant support and encouragement.

1. INTRODUCTION

1.1. Research motivation

In order to predict high waves, understanding ocean wave height statistics has become an important matter for civil engineering, especially for safety and prevention reasons. Since Longuet-Higgins (1952) introduced the Rayleigh distribution for describing the probability of wave height occurrence under certain conditions, based on results from Rice (1945), it has become the basis for describing wave height statistics. Many studies (Al-Humoud et al., 2001; Cartwright & Longuet-Higgins, 1956; Forristall, 1978; Longuet-Higgins, 1957, 1963, 1975, 1980, 1984; Mori & Yasuda, 2001; Tayfun, 1981a, 1981b, 1983, 1984, 1990, 1994, 2004, 2006, 2007; etc.) have been carried out in the last half century in order to find a better adjusted probability distribution, either theoretical or empirical, or simply to demonstrate agreement by using observations in the field. However, no clear conclusion has been derived as yet. The present study intends to give an overview of the most important existing theories and compare them with the observations from the Mediterranean and North Sea.

The motivation for the present research was the apparent “Maximum wave height paradox”. On the one hand, one found that the significant wave height was overpredicted by linear theory (Holthuijsen, 2007), having a discrepancy of approximately 7.5%. On the other hand, Cartwright (1958) found that the maximum wave crest was reasonably well predicted by a modified linear theory, including the spectral bandwidth of the spectrum. For his observations, such an inclusion caused a difference with the original linear theory lower than 2%, being far from 7.5%. This discrepancy led to investigate deeply the wave height statistics.

1.2. Objectives

The present study, which focuses on deep water, aims to gain an insight into some of the most important existing theories of short term statistics. It is based on the analysis of 40,000 time wave records (approx. 10 million waves) from 4 buoys off the Catalan coast (Mediterranean Sea) and 69 wave records (approx. 9,000 waves) from 2 laser altimeters in the North Sea (the WADIC project).

The Mediterranean data is the result of a meticulous quality control process performed on raw data directly obtained from the buoys during the period of 1991-2006. This large amount of data will hopefully serve to obtain more reliable results and diminish sampling errors. Although the length for each record is limited by the requirement of stationary, by normalizing the records with the standard deviation of the surface elevation, one can construct sets of large samples.

Although the amount of the Mediterranean data is clearly enough, a complementary analysis has been made with 69 wave records from the North Sea obtained with two laser altimeters, which were used in the WADIC project (Allender et. al, 1989). The reasons are mainly twofold: the inclusion of higher wave heights and a comparison between statistics of buoy and laser registrations. The buoy might tend to avoid the wave crests and therefore underestimate them whereas a laser altimeter does not suffer from this effect.

The starting point is the well-known Rayleigh distribution for wave heights for a narrow-band spectrum, assuming the surface elevation being Gaussian distributed. It will be shown that, although the Rayleigh distribution is derived for such limited conditions, its capacity of prediction is reasonably good. However, the Rayleigh distribution often overpredicts the observed wave height (Holthuijsen, 2007; Forristall, 1978 1984; Longuet-Higgins, 1980; Naess, 1985; Nolte et al., 1979; Rodríguez et al., 2002; Vinje 1989, etc.) but sometimes the effect is the opposite, observed higher wave heights are larger than those predicted (Mori & Janssen 2005). Therefore, other theories are studied, trying to describe better wave height statistics. Firstly, some versions of linear theory are made (Longuet-Higgins, 1980; Tayfun, 1981b, 1990; Naess, 1985) in order to try to solve the overprediction problem. That is partly related with assumption of narrow-band spectrum. Secondly, nonlinearities are invoked (Al-Humoud et al., 2001; Mori & Yasuda, 2001; Tayfun, 1983, 1984, 1994, 2006, 2007) with the aim of understanding the reason for encountering large wave heights in some records. In that case, the surface elevation is not considered to be Gaussian distributed since terms of higher order are included.

In general, as the theories include more general situations and are less limited by assumptions, the formulation becomes noticeably more complicated, in some cases leading to complex expressions for the probability density function and therefore requiring numerical integration. One important question is whether such more complicated probability distributions should be preferred, or, otherwise, the simpler Rayleigh distribution. One consideration is that it is preferable to have a formulation depending only on spectral parameters (which can be predicted from wind prediction) than on analysis from time records themselves.

In conclusion, the principal objectives are:

- Statistical and spectral analysis of the Mediterranean data, being the raw data previously filtered.
- Comparison of observations to linear theory predictions, and, if necessary, to more advanced theories.
- Complementary analysis, using the North Sea data (from 2 laser altimeters).