KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT BURGELIJKE BOUWKUNDE
LABORATORIUM voor HYDRAULICA
Kasteelpark Arenberg, 40, B-3001 Heverlee

BUOY DATA ASSIMILATION IN NEARSHORE WAVE MODELLING

Promotor:
Prof. J. Monbaliu

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen
door
Jesús PORTILLA

Mei 2009

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT BURGELIJKE BOUWKUNDE
LABORATORIUM voor HYDRAULICA
Kasteelpark Arenberg, 40, B-3001 Heverlee

BUOY DATA ASSIMILATION IN NEARSHORE WAVE MODELLING

Jury:
Prof. Dirk Vandermeulen, Voorzitter
Prof. Jaak Monbaliu, Promotor
Prof. Nicole Van Lipzig
Prof. Jean Berlamont
Prof. Geert Degrande
Prof. Gerbrandt van Vledder (TUDelft)
Prof. Peter Troch (UGent)

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen
door
Jesús PORTILLA

U.D.C. 519.2:551.466:627.52

Mei 2009

Wettelijk depot D/2009/7515/57

ISBN 978-94-6018-074-3

A la memoria de mi viejo: Segundo Jesús Portilla E.

A la Joyi, que es mi clepsidra, mi brújula, mi astrolabio

A esos dos soles: Laura y Mauro

"Como todos los hombres de Babilonia, he sido procónsul; como todos, esclavo; también he conocido la omnipotencia, el oprobio, las cárceles. Miren: a mi mano derecha le falta el índice. Miren: por este desgarrón de la capa se ve en mi estómago un tatuaje bermejo: es el segundo símbolo, Beth. Esta letra, en las noches de luna llena, me confiere poder sobre los hombres cuya marca es Ghimel, pero me subordina a los de Aleph, que en las noches sin luna deben obediencia a los Ghimel. En el crepúsculo del alba, en un sótano, he yugulado ante una piedra negra toros sagrados. Durante un año de la luna, he sido declarado invisible: gritaba y no me respondían, robaba el pan y no me decapitaban. He conocido lo que ignoran los griegos: la incertidumbre. En una cámara de bronce, ante el pañuelo silencioso del estrangulador, la esperanza me ha sido fiel; en el río de los deleites, el pánico. Heráclides Póntico refiere con admiración que Pitágoras recordaba haber sido Pirro y antes Euforbo y antes algún otro mortal; para recordar vicisitudes análogas yo no preciso recurrir a la muerte ni aun a la impostura . . . "

Jorge Luis Borges
LA LOTERIA EN BABILONIA (*Ficciones*, 1944)

Abstract

The study of wind waves deserves major attention at present time as they play a very important role in many human activities. Thanks to numerical modelling, at present wave estimates are available at global scale and with an outstanding degree of accuracy. Previously, measurements have served mainly to verify and calibrate models, but more recently the abundance of in-situ and remote sensed data have motivated the incorporation of measurements into the models in a process known as data assimilation.

The focus of this study is the assimilation of offshore buoy data in the nearshore wave modelling scenario, looking for improvements of wave hindcasts at coastal locations. The model used is WAM, a state-of-the-art third generation wave model. The assimilation approach is the so called Optimal Interpolation (OI). This method has some advantages over other methods as it is rather straightforward to implement and it is computationally very efficient. However, two major limitations are inherent in this method. The first is the a-priori representation of the errors characteristics of the system, and the second the retrieval of analyzed magnitudes into the wave spectrum. The confrontation of these shortcomings has largely drawn the experimental line followed in this study.

For the former, it is recognized that the successful implementation of any assimilation system requires a proper knowledge of the spatial structure of the system's errors (model and observations). However, that knowledge is in general very poor and that is the case for the present study area. Therefore, different parameterizations of the gain matrix were tested. The OI scheme was implemented initially following a conventional approach where the errors of the system are considered homogeneous and isotropic. However, these assumptions were found to be restrictive in the nearshore. Therefore, experiments with other approximations introducing a number of anisotropic functions were carried out. Namely, a parametric 2D Gaussian function, a function based on long-term model output, and a function based on short-term model output. The advantage of the last two is that they are more transferable to other locations. Nevertheless, the results obtained using the three approximations are similar.

For the latter, from early data assimilation studies the distinction of different wave systems was believed to increase the robustness of the retrieval algorithm and palliate this deficiency of the OI approach. In this study, some steps forward have

been done in this regard with the development of more robust partitioning and identification algorithms. It is pointed out that their practical implementation in the assimilation system is not trivial. The development of a robust algorithm for cross-assigning model and buoy partitions needs further investigation.

From the results of the numerical experiments, the benefit of the assimilation is evidenced by the capability of the system to correct some of the model deficiencies. Especially, excessive energy dissipation could be compensated resulting not only in better estimates of mean wave parameters, but also in better spectral representations. The most evident enhancements were obtained in conditions where the model performance was low, namely, in conditions of moderated wind with the presence of swell.

A remarkable aspect is the long duration effect of the assimilation. Corrections prevail in the system for a time lag of the order of days while the travelling time of waves in the domain is of the order of hours. This effect is attributed to the spectral shape corrections, which act as improved initial conditions for the proceeding wind events. Low frequency bands (swell) are better represented in the assimilation run than in the normal run. Through the quadruplet non-linear interaction it is then easier to transfer energy input from wind to lower frequencies

The simultaneous assimilation of wave height and period did not show an advantage over the assimilation of wave height only. One of the reasons is that by the assimilation of wave height only, the spectrum is subject to variation and wave period changes accordingly. However, other considerations should be taken into account for further research as the physical assumptions used for updating the spectrum and the simplifications imposed in the applied algorithm.

The assimilation of wave height during wind-sea dominated periods did not show a large impact. On the one hand, this is because in purely wind-sea conditions the model better reproduces waves. On the other hand, the performance of the assimilation scheme in those conditions is affected by the fact that a typical model underestimation of wave energy is lower at the coastal locations than at offshore locations. This produces conflicting assimilation requirements and lowers the impact of assimilation in wind-sea conditions.

Samenvatting

Windgolven spelen een belangrijke rol bij heel wat menselijke activiteiten. Met behulp van numerieke modellering kan men nu over schattingen voor golfparameters beschikken op globale schaal en dit met een verbazingwekkende nauwkeurigheid. In het verleden werden metingen vooral gebruikt voor kalibratie en validatie van modellen. Maar dankzij de steeds grotere beschikbaarheid worden metingen, zowel bekomen in situ als met aardobservatietechnieken, meer en meer geïntegreerd in modellen via zogenaamde data-assimilatie.

De nadruk in deze studie ligt op de assimilatie van meetgegevens afkomstig van offshore boeien in golfmodelleringstoepassingen voor kustzones met als doel het verbeteren van de kwaliteit van de output van het golfmodel dicht bij de kust. Het gebruikte golfmodel is het WAM-model, een state-of-the-art spectraal golfmodel van de derde generatie. De gebruikte assimilatietechniek is de zogenaamde Optimale Interpolatie (OI) techniek. Deze methode heeft een aantal voordelen. Ze is vrij eenvoudig te implementeren en heel efficiënt naar rekenkracht toe. Ze heeft echter ook twee belangrijke nadelen. Men moet a priori de foutkarakteristieken van het systeem kennen en men moet weten hoe men karakteristieke grootheden van golven gaat verwerken in het spectrum van de golven. De confrontatie met deze beperkingen heeft in grote mate de onderzoekslijn bepaald.

Wat het eerste nadeel betreft, is het algemeen gekend dat om met succes aan data-assimilatie te kunnen doen, men heel goed de ruimtelijke structuur van de foutkarakteristieken moet kennen en dit zowel voor fouten afkomstig van het model als voor fouten verbonden aan de metingen. Jammer genoeg is die kennis meestal heel beperkt en dat is ook bij deze studie het geval. Verschillende parameterisaties voor de gain-matrix werden uitgetest. Het OI-algoritme werd eerst op conventionele wijze geïmplementeerd waarbij de fouten in het systeem homogeen en isotroop verondersteld worden. Deze veronderstellingen bleken hier echter te restrictief. Daarom werden er een aantal experimenten opgezet met anisotrope functies, met name met een parametrische 2D Gauss-functie, met een functie gebaseerd op lange termijn model output, en met een functie gebaseerd op korte termijn model output. Het voordeel van de laatste twee benaderingen is dat ze makkelijker overdraagbaar zijn naar andere locaties. De resultaten die werden bekomen voor dit studiegebied waren evenwel vergelijkbaar voor deze drie benaderingen.

Wat het tweede nadeel betreft, is uit vroegere assimilatiestudies gekend dat men erin gelooft dat het onderscheiden van verschillende golfsystemen zou bijdragen tot de robuustheid van het assimilatieschema en dat op die manier deze beperking zou

kunnen gemilderd worden. In dit onderzoek zijn een aantal stappen gezet naar een meer robuust splitsings- en allocatie-algoritme. De praktische implementatie in een assimilatieprocedure is echter geen sinecure. De ontwikkeling van een robuust algoritme voor het aan elkaar toewijzen van de verschillende golfsystemen die in een gemeten en in een gemodelleerd golvenspectrum kunnen onderscheiden worden, vraagt bijkomend onderzoekswerk.

Uit de resultaten van de numerieke experimenten is gebleken dat het assimilatiesysteem in staat is om te compenseren voor sommige beperkingen in het golfmodel. In het bijzonder kwam tot uiting dat een te hoge dissipatie van energie gecompenseerd werd, waardoor niet alleen de karakteristieke golfparameters, maar ook de vorm van het berekende spectrum beter werden gemodelleerd. De grootste verbeteringen werden gevonden in condities met lage performantie van het golfmodel, in het bijzonder condities met matige wind in aanwezigheid van deining.

Een opmerkelijke bevinding was het langdurig effect van de assimilatie op de modelresultaten. De correcties bleven gedurende een grootteorde van dagen nawerken niettegenstaande de looptijd van de golven tot aan de kust slechts van de grootteorde een uur is. Dit effect wordt toegewezen aan de verbeteringen van de vorm van het spectrum. Dit zorgt voor verbeterde initiële condities voor de daaropvolgende energie-input door wind. De energie bij lagere frequenties (deining) is beter gemodelleerd in de run met assimilatie dan in de run zonder assimilatie. Energie-input door wind kan dan via de niet lineaire wisselwerking (quadruplets) makkelijker naar lagere frequenties gepompt worden.

Het gelijktijdig assimileren van golfhoogte en golfperiode leidde niet tot een verbetering t.o.v. assimilatie van enkel golfhoogte. Eén van de redenen is dat bij de assimilatie van enkel golfhoogte, het golvenspectrum ook al bijgestuurd wordt en dat de golfperiode daardoor ook al verandert. Andere beschouwingen dienen evenwel bij verder onderzoek meer in detail bekeken te worden, zoals de fysische veronderstellingen bij het updaten van het spectrum en de vereenvoudigingen die aan de grondslag liggen van het gebruikte assimilatie algoritme.

Het assimileren van golfhoogte in windzee gedomineerde perioden had geen grote impact op de resultaten. Dat is enerzijds te verklaren door het feit dat het golfmodel in staat is deze situatie beter te reproduceren en anderzijds door het feit dat het gebruikte golfmodel golfcondities dicht bij de kust minder onderschat dan meer offshore. Dit leidt tot tegenstrijdige eisen waardoor de impact van assimilatie in windzee gedomineerde situaties vermindert.

Acknowledgments

Funding from the Flemish Interuniversity Council (VLIR), the Research Foundation Flanders (FWO), contracts G.0477.04 and G.0477.07, and the Belgian Science Policy Projects Marebasse and Quest4D are gratefully acknowledged. For their financial support in the form of a page charge waiver for the publication of part of this work, I would also like to acknowledge the Bulletin of the American Meteorological Society.

To my promotor, Prof. Jaak Monbaliu, I express my gratitude for the confidence deposited upon me during these years, for his support and constructive criticism. The members of the evaluation commission, Prof. Nicole Van Lipzig, Prof. Gerbrandt van Vledder, Prof. Peter Troch, Prof. Jean Berlamont and Prof. Geert Degrande I thank for their time spent in revising this thesis and their constructive and dedicated remarks. I am also deeply thankful to Prof. Jan Feyen for his support and personal concern.

The Agency for Maritime and Coastal Services of the Ministry of Transport and Public Works (MD&K) and the *Vlaams Instituut voor de Zee* (VLIZ) are acknowledged for the wave data from Belgian locations. Data on wind fields was made available by the United Kingdom Meteorological Office (UKMO) and the Management Unit of the North Sea Mathematical Models (MUMM). The *Centro de Investigación Científica y de Educación Superior de Ensenada* (CICESE) provided the data from the Gulf of Tehuantepec collected during the intOA Experiment Field Campaign. Data from the European Centre for Medium-Range Weather Forecasts (ECMWF) was downloaded in the framework of the EU project SEAMOCS, and wave data from the Netherlands was downloaded from the website of the *Ministerie van Verkeer en Waterstaat*, Hydro Meteo Centrum Zeeland, Meetnet ZEGE. Data from the U.S.A buoys was downloaded from the website of the National Data Buoy Center (NDBC). The cooperation or contributions of all these institutions are gratefully acknowledged.

I would also like to acknowledge a number of colleagues from several institutions who contributed in different ways to the development of this study. A special thanks to Prof. Paco Ocampo and Héctor García from the CICESE for their contributions and the fruitful stay in Mexico. Guido Dumon from the MD&K I would like to thank for his cooperation and the data supply. For their constructive comments, numerical subroutines and bibliography I extend my thanks to Prof. Gerbrandt van

Vledder from Alkyon & Delft University of Technology and to Jean Bidlot from the ECMWF. I am very thankful to Erick Rogers from the Naval Research Laboratory of the Stennis Space Center (Mississippi) for his feedback and the interest shown in my work. To Luigi Cavaleri from the Institute of Marine Sciences CNR (Venice-Italy) I am grateful for his dedicated comments and support. My gratitude is also expressed to Brigitte Furevik and Anne Karin Magnuson from the Norwegian Meteorological Institute (DNMI) for their comments and collaboration. To Vera Van Lanker from the University of Gent and Dries van den Eynde from the MUMM I thank for their cooperation during the projects work. I am indebted to Patrick Willems from the Hydraulics Laboratory (KULeuven) for his assistance with statistical concepts. Rubén Darío Montoya from the Universidad de Medellín, I would like to thank for his cooperation. For their valuable support with the use of different software and software licensing, I thank Wim Obbels, Samuel de Souza, and Jan Ooghe from Ludit (KULeuven). I would also like to acknowledge the administrative support from Mr. Bob Geivers and Mrs. Veronique Cortens from the KULeuven.

I am very grateful for the cooperation, help, discussions, fellowship and friendship of my colleagues from the KULeuven. Especially Eliz-Mari Lourens, Alessio Giardino, Mireya Matos, Elsy Ebraim, Ilse Vitse, Mark Rubarenzya, Luis Timbe, Pablo Vanegas, Suhyb Salama, Alessandro Toffoli, Raul Vásquez, Marcelo Heredia, Mauricio Villazón, Efrén Santos, Stefanie Adam, Rolando Celleri, Jan Ronsijn, Jaime Bojorque and Andrea Stenti.

Many thanks are also expressed to friends who, at one stage or another, made my stay in Belgium not only comfortable, but also delightful: especially Leen Janssens, Eef Vervecken, Thomas Adriansen, Mishel Vervecken, Consuelo Pérez, Ann Eijkelenburg, Jos and Staf Janssens, Cristina Mejia, Solomon Tuccu, Jining Peng, Lisa Copin, Ellen Nys and the families Ávila-Zeman, Florizone-Rodríguez, Martínez-Vanwildemeersch, De Clerck-Galera, and Baeyens-Basile.

I am also very grateful for my many friends in Ecuador, especially Diego Velasco and Javier Mideros, for their selfless support and encouragement. Heartfelt thanks also to Andres Barrera, Wilmer Portilla, Luis Narváez, Darwin Villagómez, Marcia Delgado, Graciela Delgado, Maria Elena López, Roberto Sánchez, and César Muñoz for their unconditional friendship. Finally, I express my love and gratitude for my parents Sonia Yandún and Segundo Jesús Portilla E. and to my sisters and brothers: Damaris, Mishel, Henry and Favio.

Table of contents

Chapter 1

Introduction

1.1 Framework

The study of wind waves deserves major attention at present time, because not only they are a fascinating phenomenon that grants big satisfactions to many people but mainly because they play a very important role in human's live. Offshore oil exploitation, fishery, and marine transportation, for instance, are activities that remind us of the fact that each of us (and therefore today world's economy) is dependent on the oceans (UNCTAD 2008, Langley et al. 2009). Moreover, coastal engineering and planning have deserved increasing attention in the last years due to the danger associated to sea level rise and the devastating impact of recent natural disasters as hurricanes and tsunamis (Collins et al. 2006, Eisner 2005). From a more fundamental point of view, wind waves play a crucial role in other important marine processes like the exchange of gases, heat, and momentum between the atmosphere and the sea, the interchange of energy with currents, the interaction with the bottom, the sea thermodynamics, etc (Stewart 2008). These are processes affecting the earth's climate and marine live.

Thanks to development of the theoretical basis governing the evolution of wind waves carried out during the last few decades (Komen et al. 1994), and in parallel thanks to the accelerated computing and technological progress of these last years, at present wave estimates are available at global scale and with an outstanding degree of accuracy (Cavaleri et al. 2007). Numerical modelling is therefore one of the most important tools for wave estimation, not only for practical operational use but also for further conceptual and engineering studies.

The importance of wave measurements must also be emphasized. Especially in-situ measurements are expected to have high accuracy (Young 1999). Previously, measurements have served mainly to verify and calibrate models, but more recently the abundance of in-situ and remote sensed data have motivated a line of research dedicated to incorporate wave measurements into the models with the aim of

improving models performance (Hasselmann et al. 1997). The procedure by which measurements are incorporated into the model computations is known as data assimilation and aims at minimizing the differences between computed and observed magnitudes.

The assimilation of data has deserved special attention in many forecasting problems. Mainly in meteorology, data assimilation schemes have been extensively developed, tested, and employed (Daley 1991, Evensen 2006). Other fields where this methodology is important are for instance astrophysics, economics, computer sciences, navigation and geo-positioning technology, and also wave and hydrological modelling (e.g., Barrero-Mendoza 2005, Zhang and Luh 2005, Abbott and Powell 1999). Until now, the application of data assimilation in numerical wave modelling has been rather modest. This is due to the limited availability of good quality data, and a common assumption of wave modellers that most errors in wave predictions are due to driving wind fields errors. Although, it is also recognized that relevant processes of the wave evolution as white-capping dissipation and wave-bottom interaction processes are still poorly understood (Holthuijsen 2007). Therefore, the data assimilation mechanism can prove useful for improving wave estimates in several situations.

1.2 Scope and Objectives

The focus of this study is the application of an assimilation scheme in the nearshore scenario. Most marine activities take place in nearshore environments. Moreover, in most circumstances, observations are available only at sparse locations and wave modelling can be very challenging. Therefore, improving wave estimates nearshore by data assimilation is a crucial issue as several users of both operational and hindcast products can benefit from. The aim of this study is not the improvement of the wave forecast as such but the improvement of wave hindcasts for wave climate studies.

The main objectives are:

- To implement a buoy data assimilation scheme into a numerical spectral wave model for nearshore applications.
- To study the impact of assimilating buoy observations in numerical wave estimation.
- To improve the wave estimates nearshore

The wave model used is the WAM-PRO model (Monbaliu et al. 2000), a version of the WAM model (Günther et al. 1992, Komen et al. 1994) adapted to run efficiently in nearshore applications. WAM is a state-of-the-art third generation wave model, which solves the action balance equation accounting for sink-source terms as wind input, quadruplet interactions, white-capping dissipation, and bottom friction.

The data assimilation approach considered in this study is the so called Optimal Interpolation (OI) approach (Lionello et al. 1992, Voorrips et al. 1997, Hasselmann et al. 1997, Bouttier and Courtier 1999). In this approach, sparse data available from observations are interpolated into the model in such way as to obtain an analyzed field that is the optimum in statistical terms, considering the relative accuracy of measurements and model output. This method has a number of advantages over other methods mainly because it is rather straightforward to implement and it is computationally very efficient.

1.3 Outline

This document consists of seven chapters. This introductory chapter gives an overview of the problem and its relevance, the objectives, and the scope of this study. In chapter two a revision of spectral wave modelling is given. This revision intends to give the framework for the wave modelling setup used in the data assimilation experiments. Other modelling components like the bathymetric data used and the wind fields are also explained. In addition, a description is given i) of the data available for model verification and data assimilation, ii) of the wave parameters used for comparison, and iii) of the statistical parameters used for the analysis. The results for the default model run are also introduced.

From the outset of this study, the retrieval of analyzed mean parameters into the wave spectrum was considered fundamental for developing a robust assimilation algorithm. Specifically the identification of wind-sea and swell was given special attention since good model performance is recognized by its ability to properly represent the wave spectral shape and not only the spectral mean parameters. The outcome of that part of the study is documented in Chapter 3. Existing methodologies and algorithms for partitioning wave spectra and for the identification of wind-sea and swell were investigated and tested in a number of different conditions. As an outcome, some improvements to these algorithms where suggested (Portilla et al. 2009). It is recognized that the implementation of such a scheme in the assimilation system requires a robust cross-assignment algorithm to match buoy and model partitions. The robustness of this algorithm is crucial, to

assure continuity of the two wave fields (wind-sea and swell) both in space and time.

In Chapter 4, a revision is given of data assimilation schemes, giving emphasis to applications in spectral wave modelling. Common methods can be placed in three main categories: Variational methods, Kalman filtering, and Optimal Interpolation. The main objective of this overview is to give a framework for the OI approach considered in this study and identify the advantages and disadvantages of this method with respect to other commonly used methods. It is pointed out that the main characteristics of the OI method are its robustness and simplicity. The efficiency of this method and its low computational costs make it very attractive for many practical applications. The need for understanding data assimilation as a statistical problem is emphasized, which make us conclude that a better description of the error statistics of the modelling system is required.

The implementation of the OI approach is studied in Chapter 5. This implementation was first set up according to the parameterizations given by Voorips et al. (1997) for the errors of the system, since the study area and observations used are similar to those studied here. However, the relatively large dimensions of the assimilation domain proved to be inappropriate for nearshore application. Especially the external introduction of energy in very shallow waters caused unrealistic source term rates and lead eventually to run-time computing errors. This shortcoming was overcome using a two-dimensional Gaussian function, adapted to the geographical space, for the error structure. The ability of the scheme to make corrections is evident in situations where the default model performance is low, namely in conditions where wind-sea and swell systems coexist. In wind-sea conditions the model performance is better, and consequently corrections by data assimilation are less effective. Additionally it is shown that by assimilating wave height the scheme has the ability to correct also spectra. Once the energy density is updated, the model dynamics (i.e., source terms and propagation) act on this new state. Consequently, other mean parameters experience corrections. Moreover, the fact that low frequency waves (swell) are simulated better has a positive assimilation memory effect. Corrections prevail during days when wind-sea and swell coexist because spectral corrections act as improved initial conditions for the proceeding wind events. The presence of swell in the model enhances the transfer of energy from wind in moderated conditions.

In Chapter 6, further developments of the assimilation scheme are described. From the previous experiments, it is likely that assimilation results can be improved by the assimilation of more than one buoy data source. However, the transferability of the 2D Gaussian function assumed for the assimilation structure is limited. In this

chapter, two other parameterizations for the errors of the modelling system are tested. The first is based on long-term model estimates, and the second on short-term model estimates. The results using these two other parameterizations are similar to those given in Chapter 5 when using the two-dimensional Gaussian function structure adapted to the geographical space. The advantage however is that these parameterizations can be easily transferred to other buoy locations. Additional experiments were carried out in order to assess the influence of assimilating mean wave period. It is noted that the assimilation of the two variables, wave height and wave period, does not show an advantage over the assimilation of wave height only. One of the reasons is that when assimilating wave height only, the mean period is already corrected. Other causes are related to the simplifications of the retrieval algorithm. Finally, experiments were carried out to assess the effect of accounting for uncertainties in the observations.

The main conclusions derived from this study and the suggestions envisaged for further research are compiled in Chapter 7. It is recognized that the interaction between the model and the assimilation scheme is actually better than expected. On theoretical bases, the OI scheme is only able to operate in the model output while in practice there is evidence of constructive interaction between the wave model physics and the OI scheme. The assimilation step behaves as an extra source term and updates the spectrum. In turn, the model processes the new information according to its dynamics. In addition, results from the numerical experiments show the benefits of data assimilation in nearshore wave modelling. Also the memory effect of assimilation opens the possibility for operational applications. Moreover, the data assimilation exercise helped in the analysis and interpretation of model results. Modelling and assimilation shortcomings are highlighted in order to point to possibilities for further research.

Chapter 2

Spectral modelling of wind waves

2.1 Introduction

Initial attempts to predict wind waves were made by Sverdrup and Munk (1947) during World War II. They developed empirical relationships between wind and wave characteristics in order to predict wave conditions. Some generally, vaguely defined concepts of a wave height and a wave period were predicted. However, in those years that was already a big leap forward. Later on, Pierson et al. (1955) introduced the use of spectral analysis to represent more conveniently ocean waves as the superposition of a number of harmonic waves with random phases. This advance allowed for a stochastic approach in the description of ocean waves and introduced the concept of the wave spectrum. The account for the dynamical evolution of the spectrum was then possible with the introduction of the spectral transport equation by Gelci et al. (1957), giving origin to the so called first generation spectral wave models. In absence of appropriate theoretical basis to describe the source term components, in these models, the net source function parametrically accounted for wave generation and dissipation imposing a universal shape to the high frequency tail of the spectrum. Further studies on the wind generation process by Phillips (1957) and Miles (1957) provided the basis for the representation of the wind input. Moreover, with the derivation of the non-linear (quadruplet) wave-wave interactions function by Hasselmann (1962), and the results of the JONSWAP experiment from Hasselmann et al. (1973), the importance of this term in the evolution of the wave spectrum was recognized. However, since the exact calculation of the non-linear wave-wave interactions was computational prohibitive the inclusion of this term in numerical wave models was only possible using simplified parametric expressions. The implementation of all this new knowledge gave rise to the development of the so-called second-generation models. In these models, the tail of the spectrum was still imposed using a universal parametric shape and the net source function accounted for wind input, quadruplet interactions, and white-capping dissipation. Subsequently, the development of an algorithm to compute approximately non-linear wave-wave interactions by

Hasselmann and Hasselmann (1985), the Discrete Interaction Approximation (DIA), allowed for the solution of the energy balance equation without imposing any shape to the wave spectrum and gave way to a new generation of wave models, the so-called third-generation spectral wave models. At present third-generation models have been implemented in many forecasting and research centers around the world, making possible an exhaustive evaluation of new physical theories and improved computing algorithms. In addition, the accelerated technological development of these last years has greatly contributed to the development of wave models. Moreover, although it is recognized that the performance of numerical wave models is nowadays outstanding, many important phenomena of the wave evolution process are far from being well understood (Cavaleri et al. 2007).

In the present chapter, a brief overview is given of theory and concepts concerning spectral modelling of wind waves. Namely, the energy balance equation and the corresponding source terms are presented according to the formulations of the default WAM model Cycle IV version. For a detailed description of wind wave theory and wave modelling the reader is referred to the dedicated literature (e.g., The SWAMP Group 1985, Komen et al. 1994, Young 1999, Janssen 2004, Holthuijsen 2007, Cavaleri et al. 2007)

2.2 The energy balance equation

The evolution of the wave spectrum is described by the spectral energy balance equation, which written in Cartesian coordinates reads:

$$\frac{\partial F}{\partial t}+\frac{\partial}{\partial x}\left(C_x F\right)+\frac{\partial}{\partial y}\left(C_y F\right)+\frac{\partial}{\partial \theta}\left(C_\theta F\right)+\frac{\partial}{\partial \sigma}\left(C_\sigma F\right)=S_{TOT} \qquad [2.1]$$

where: $F= F(\sigma,\theta,x,y,t)$ is the wave spectrum (in m^2/Hz.rad), x and y are the space coordinates (in meters), t is time (in seconds), θ is the wave direction (in radians), σ is the intrinsic radiant frequency (in Hz) and S_{TOT} represents the total source function (in m^2/Hz.rad.s). C_x and C_y are the propagation velocities in geographical space and C_θ and C_σ are the propagation velocities in spectral space

The left hand side of equation 2.1, accounts for energy propagation in space (x,y) and time. The first term represents the rate of local change of energy in time. The second and third terms correspond to the advection of wave energy in space. The last two terms account for depth and current refraction describing the distribution of energy over the spectrum. In the right hand side of equation 2.1 the resulting of the sink/source terms is expressed as S_{TOT}. This term is further explained in the next section.

2.3 Source terms

The total sink/source terms includes non-linear processes that are not accounted for in the advection part of equation 2.1. The main source process is the generation of waves by wind, called wind input (S_{in}). Sink terms include redistribution of energy amongst the different spectral components due to non-linear wave-wave interactions (S_{nl}), and dissipation processes as white-capping dissipation (S_{wc}), and bottom friction dissipation (S_{bf}). This is expressed in equation 2.2.

$$S_{TOT} = S_{in} + S_{nl} + S_{wc} + S_{bf} \qquad [2.2]$$

These four source terms are accounted for in the default version of the WAM model.

As waves propagate in shallow waters, other processes like depth-induced wave breaking, triad wave-wave interactions, reflection, and diffraction become also important. Specially, the relevance of wave breaking due to limited depth should be emphasized as it is expected to be the dominant process in extreme shallow water conditions (Battjes 1994, Padilla-Hernández 2002). Depth-induced breaking has been also considered an essential component to simulate waves at coastal scale, Monbaliu et al. (2000). In order to quantify the effect of depth-induced wave breaking, Luo (1995) performed numerical experiments with the WAM model using the depth-induced wave breaking formulation of Battjes and Janssen (1978). Luo (1995) found a maximum effect of 15% difference in wave height at the location of A2B in storm conditions, where wave heights were of about 3m. No appreciable effects were observed neither at other locations nor in moderated and low wave conditions.

It should be also mentioned that in the presence of non uniform currents wave energy is not conserved due to an exchange of energy between waves and currents. Instead, the wave action, defined as $N = F(\sigma,\theta,x,y,t)/\sigma$ is conserved, and equation 2.1 is written in terms of this magnitude (see for instance Ozer et al. 2000, Osuna 2002).

Despite the expected relevance of the depth-induce breaking, and current and depth variations in shallow waters, these phenomena were not included in the present modelling setup since the introduction of these processes requires further analysis and verification and the emphasis here is the implementation and evaluation of the data assimilation system.

2.3.1 Wind input

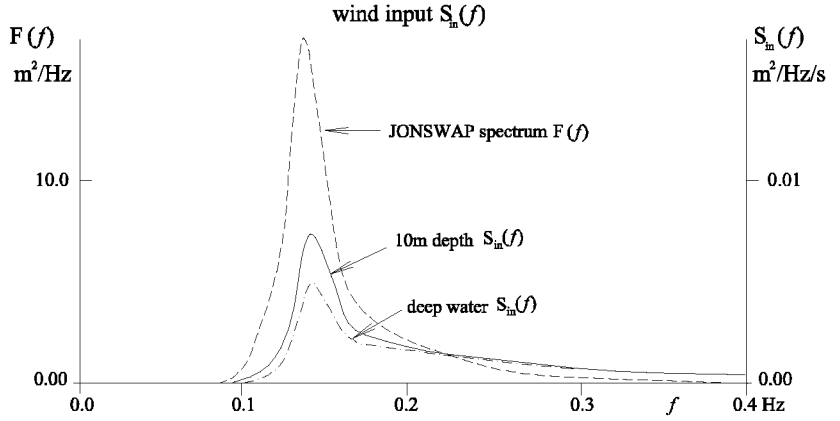

Figure 2.1. Wind input source term for a JONSWAP spectrum in deep and shallow water. After Holthuijsen 2007.

The wind input source term is represented as the sum of a linear term for the initial growth stage given by Phillips (1957), and an exponential term given by Miles (1957). However the linear term is usually neglected as it is overridden by the exponential growth.

$$S_{in}(\sigma,\theta) = a + \beta F(\sigma,\theta) \qquad [2.3]$$

Based on Miles (1957) theory and assuming a logarithmic shape for the wind profile, Janssen (1991) incorporated the effect of the sea state in the wave's growth rate by means of the so-called profile parameter Ω. As Ω depends on the roughness of the air flow, it is considered that the effective roughness length (z_0) depends on the sea state through the wave-induced stress (τ_w). The grow rate of waves thus reads:

$$\beta = \max\left\{0, \varepsilon \frac{\beta_m}{\kappa^2} x^2 \mu \ln^4 \mu\right\}\sigma; \qquad \mu \le 1 \qquad [2.4]$$

where σ is the radiant frequency, ε is the ratio of air to water density, β_m is a constant equal to 1.2, κ is the von Karman constant equal to 0.41,

$$x = \frac{u_*}{c}\cos(\theta - \psi) \;, \quad \mu = \left(\frac{u_*}{\kappa c}\right)^2 \Omega e^{\kappa/x} \;, \qquad [2.5]$$

where c is the phase speed, u_* is the friction velocity, and θ and ψ are the wave and wind directions respectively,

$$\Omega = \frac{g z_0}{u_*^2}, \quad z_0 = \frac{\alpha \tau}{g\sqrt{1 - \frac{\tau_w}{\tau}}} \qquad [2.6]$$

Ω is the profile parameter, z_0 is the effective roughness length, g is the gravity acceleration, α is a constant equal to 0.01, τ is the total stress ($\tau = u_*^2 = \tau_w + \tau_{turbulent}$), which is found from the iterative solution of:

$$\tau = \left[\frac{\kappa U_L}{\ln(L / z_0)} \right], \qquad [2.7]$$

where L is the mean height above the water, U_L the wind speed at that height,

$$\tau_w = \varepsilon^{-1} g \iint S_{in} N(\sigma, \theta) \mathrm{k}\, d\sigma\, d\theta \qquad [2.8]$$

where $N(\sigma,\theta)$ is the action density spectrum.

2.3.2 Nonlinear wave-wave interactions

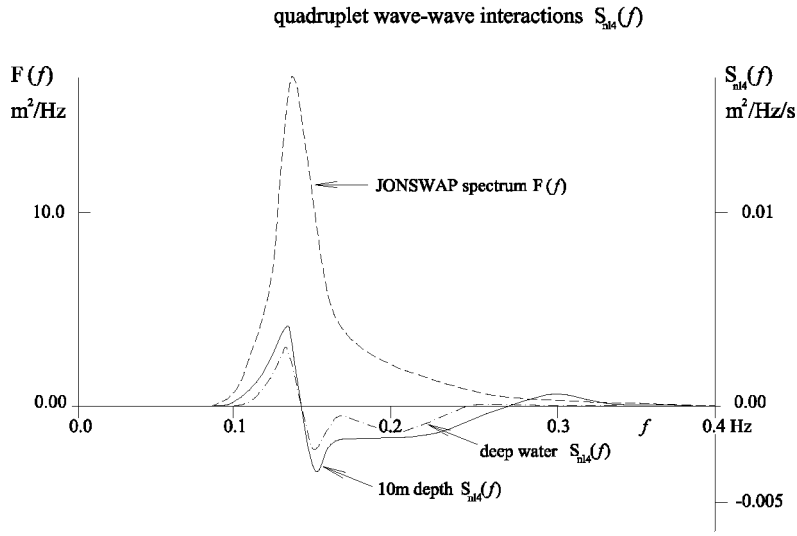

Figure 2.2. Quadruplet nonlinear wave-wave interactions source term for a JONSWAP spectrum in deep and shallow water. After Holthuijsen 2007.

Nonlinear energy transfer occurs amongst waves due to a resonance mechanism that is not accounted for in the energy transport equation due to the linear approximation.

Hasselmann (1962) found that a set of four waves exchange energy if they fulfill the following resonant conditions:

$$\begin{aligned} \mathbf{k}_1 + \mathbf{k}_2 &= \mathbf{k}_3 + \mathbf{k}_4 \\ \sigma_1 + \sigma_2 &= \sigma_3 + \sigma_4 \end{aligned} \qquad [2.9]$$

Where the vectors $\mathbf{k_1}$, $\mathbf{k_2}$, $\mathbf{k_3}$, $\mathbf{k_4}$, represent the wave numbers, and the scalars σ_1, σ_2, σ_3, σ_4, the radiant frequencies of the interacting waves. The rate of change of action density at a particular wave number can be written in the form of a Boltzmann integral.

$$S_{nl}(\mathbf{k}_4) = \iiint G(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4)\,\delta(\sigma_1 + \sigma_2 - \sigma_3 - \sigma_4) \times \left[N_1 N_3 (N_4 - N_2) + N_2 N_4 (N_3 - N_1)\right] \mathrm{d}\mathbf{k}_1 \mathrm{d}\mathbf{k}_2 \mathrm{d}\mathbf{k}_3 \qquad [2.10]$$

where G is a coupling coefficient, δ is the Dirac delta function and N_j is the action density at the corresponding wave numbers.

The exact calculation of the non-linear transfer is computationally too expensive. Therefore, several other methods have been developed to find approximate solutions for operational wave models (Van Vledder 2006a,b). One of the most popular approximations, which is implemented in the Cycle IV version of the WAM model is the DIA algorithm from Hasselmann and Hasselmann (1985). They found that the nonlinear transfer could be well simulated by only one mirror image pair of interaction configurations where two wave numbers are taken as identical and the other two are specified with a parameter λ.

$$\begin{aligned} \sigma_1 &= \sigma_1 = \sigma \\ \sigma_3 &= (1+\lambda)\sigma = \sigma^+ \\ \sigma_3 &= (1-\lambda)\sigma = \sigma^- \end{aligned} \qquad [2.11]$$

A value of 0.25 was found by Hasselmann and Hasselmann (1985) to give good agreement with the exact computations. The nonlinear interactions computed with the DIA are given by:

$$\begin{pmatrix} \Delta S_{nl} \\ \Delta S^{+}_{nl} \\ \Delta S^{-}_{nl} \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} C_{nl} g^{-4} \left(\frac{\sigma}{2\pi}\right)^{11} \left[F^2 \left(\frac{F^+}{(1+\lambda)^4} + \frac{F^-}{(1-\lambda)^4} \right) - 2F \frac{F^+ F^-}{(1-\lambda^2)^4} \right] \qquad [2.12]$$

Where C_{nl} is a constant equal to $3\text{x}10^7$, g is the gravitational acceleration. F is the energy density at the central interacting number $\mathbf{k}$ ($\mathbf{k_1}$ = $\mathbf{k_2}$), F+ and F- are the energy densities at the interacting wave numbers $\mathbf{k_3}$ and $\mathbf{k_4}$. The net source function

is the sum of equation 2.12 over all wave numbers, directions and interaction configurations.

2.3.3 White-capping dissipation

The main source of wave energy dissipation in deep water is thought to be due to wave breaking which is a visible phenomenon characterized by "white-capped" waves. Wave breaking is a strongly nonlinear interaction process and is the least understood in the description of wind waves.

Hasselmann (1974) proposed a formulation for the white-capping dissipation considering that the process could be treated as a random distribution of pressure pulses. It was found that under very general conditions the dissipation process is quasi-linear with respect to the spectrum.

$$S_{wc}(\sigma,\theta) = -\mu k F(\sigma,\theta) \qquad [2.13]$$

where k is the wave number magnitude and μ is a proportionality factor which depends on the wave steepness and represents a statistical property of the white-caps.

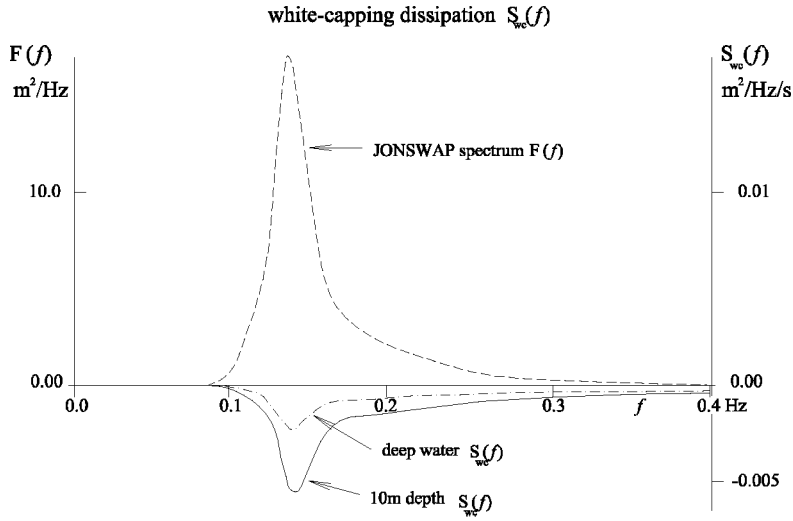

Figure 2.3. White-capping dissipation term for a JONSWAP spectrum in deep and shallow water. After Holthuijsen 2007.

Komen et al. (1984) proposed a dissipation function that allowed the existence of an equilibrium solution of the energy balance equation for fully developed waves. However, when the wind input generation formulation of Janssen (1991) was introduced, that formulation had to be reconsidered in order to obtain a proper energy balance at high frequencies (Cycle IV version).

$$S_{wc}(\sigma,\theta) = -C_{wc}\left((1-n)+n\frac{k}{\tilde{k}}\right)\left(\frac{\tilde{s}}{\tilde{s}_{PM}}\right)^{p}\frac{\tilde{\sigma}}{\tilde{k}}kF(\sigma,\theta) \qquad [2.14]$$

Where $\tilde{s}$ is the overall wave steepness, $\tilde{s}_{PM}$ is the value of $\tilde{s}$ for the Pierson and Moskowitz (1964) spectrum. $\tilde{\sigma}$ is the mean frequency, $\tilde{k}$ is the mean wave number, and C_{wc}, n and p are calibration coefficients. In the Cycle IV version C_{wc}= 4.10x10^{-5}, n= 0.5, and p = 4.

2.3.4 Bottom friction dissipation

In shallow waters, dissipation of wave energy due interaction with the bottom can occur due to different mechanisms. In shelf seas however bottom friction is the dominant process. As the orbital motion of the water particles reaches the bottom, energy is dissipated in the form of turbulent friction. This is often represented as a drag law:

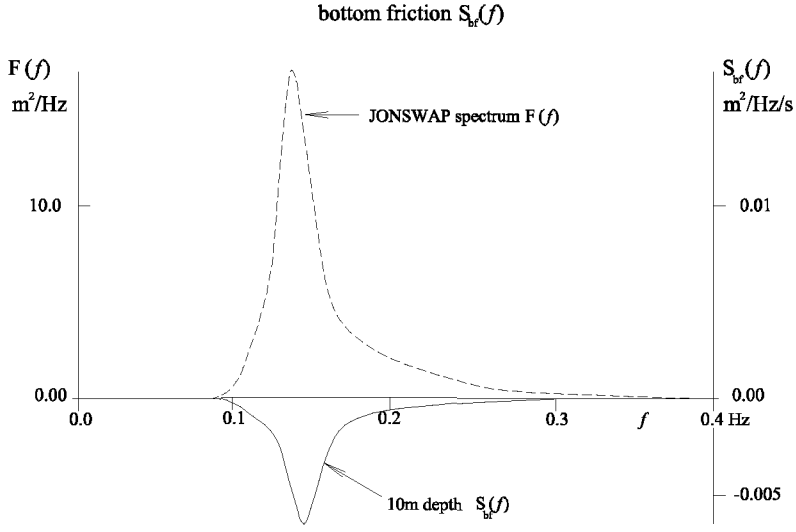

Figure 2.4. Bottom friction dissipation term for a JONSWAP spectrum in shallow water. After Holthuijsen 2007.

$$S_{bf}(\mathbf{k}) = -\frac{C_{bf}}{g}\left\langle U_{\mathbf{k}}^2 \right\rangle \qquad [2.15]$$

Where C_{bf} is a dissipation coefficient, g is gravity, and $U_{\mathbf{k}}$ is the root-mean-square orbital velocity at the bottom, associated to the wave number **k**. Written in terms of the wave spectrum thus becomes:

$$S_{bf}(\sigma,\theta) = -\frac{C_{bf}}{g}\left[\frac{\sigma}{\sinh(kd)}\right]^2 F(\sigma,\theta) u_{rms,bottom} \qquad [2.16]$$

Where $u_{rms,bottom}$ is the root mean square orbital velocity at the bottom, and d is the water depth.

The different formulations for the bottom friction differ mainly in the expression of the dissipation coefficient. A simple expression was proposed by Hasselmann et al. (1973), considering a constant dissipation coefficient.

$$C_{bf} = \frac{\Gamma}{g} \qquad [2.17]$$

where g is the acceleration of gravity and $\Gamma = 0.038$ m^{2}s^{-3}.

2.4 Deep and shallow water processes

Regarding the relative importance of the different processes in deep and coastal waters, Battjes (1994) has given an interesting overview that is recalled in table 2.1.

It is also interesting to note that, as indicated by Holthuijsen (2007) and shown in figures from 2.1 to 2.4, not only the main processes responsible for the wave evolution change in relative importance but also deep water processes are enhanced when waves reach shallow water. This effect has been discussed in some detail by Young (1999). In the case of the wind input term this is mainly associated to the ratio of the wind speed to the wave phase speed (e.g., u_* /c) present in the formulation (equation 2.5). As the phase speed decreases due to water depth decrease, energy transfer increases. However, he remarks that there is no experimental evidence to verify these results. In the quadruplet interactions, two effects are accounted for. For the first, the interacting quadruplets (equation 2.10) are modified in shallow water with respect to deep water (due to the different form of the dispersion relationship). For the second, also the magnitude of the coefficient G in equation 2.10 increases as water depth decreases. In the case of white-capping dissipation the enhancement effect is caused by the dependence on the wave steepness. Waves propagating from deep to shallow water decrease in wave length and increase in steepness (shoaling). As a result white-capping dissipation increases.

Table 2.1. Relative importance of physical processes in deep and shallow waters (After Battjes 1994 and Holthuijsen 2007)

	Oceans	Shelf seas	Nearshore	Harbor
Wind generation	■	■	▨	
Quadruplet wave-wave interactions	■	■	▨	
White-capping dissipation	■	■	▨	
Bottom friction dissipation		▨	▨	
Current refraction	▨	▨	▨	
Depth refraction / Shoaling		▨	■	▨
Depth breaking		▨	■	
Triad wave-wave interactions			▨	▨
Reflection				■
Diffraction			▨	■

LEGEND

Dominant	■
Significant but not dominant	▨
Minor importance	▨
Negligible	

2.5 Integral mean wave parameters

The energy spectrum is described by the mean integral parameters: total energy E_m, or significant wave height H_{m0}, mean direction θ_m and mean frequency $f_{m-1,0}$ (or its inverse value mean period $T_{m-1,0}$).

$$E_m = \int_0^{\infty}\int_0^{2\pi} F(f)\,df\,d\theta \qquad [2.18]$$

E_m is the total energy [m^2], F represents the energy density spectrum (in m^2/Hz.rad), f is the frequency (in Hz), and θ is the direction (in radians).

$$H_{m0} = 4\sqrt{E_m} \qquad [2.19]$$

H_{m0} is the significant wave height based on the moment zero of the spectrum (in meters).

$$\left(T_{m-1,0}\right)^{-1} = f_{m-1,0} = \frac{E_m}{\int_0^\infty F(f) f^{-1} df} \qquad [2.20]$$

$T_{m-1,0}$ is the mean wave period (in seconds) based on the zero and first moments of the spectrum. The inverse value of $T_{m-1,0}$ is the mean frequency based on these moments (in Hz).

$$\theta_m = \arctan\left(\frac{\int_0^\infty F(f)\sin\theta(f)\,df}{\int_0^\infty F(f)\cos\theta(f)\,df}\right) \qquad [2.21]$$

θ_m is the mean vectorial wave direction of the spectrum (in radians), see Kuik et al. (1998)

2.6 Statistical parameters

The following statistical parameters are considered for comparisons between model results and buoy parameters: root mean square error (RMSE), bias, and scatter index (SI). The correlation coefficient is also used to evaluate the correlation between two variables. The formulations of these parameters as given by Van Vledder (1993) are:

$$RMSE = \left\{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - y_i\right)^2\right\}^{1/2} \qquad [2.22]$$

$$BIAS = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - x_i\right) \qquad [2.23]$$

$$SI = \frac{RMSE}{|\bar{x}|} \qquad [2.24]$$

$$r_{xy} = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\left\{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2\right\}^{1/2}\left\{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2\right\}^{1/2}} \qquad [2.25]$$

where x is the measured and y the modelled variable.

2.7 The Wave Model WAM

The WAM (Cycle IV) model is a third generation spectral wave model developed by the WAMDI Group (1988) and described by Günther et al. (1992). The WAM model is one of the most widely used wave models both for operational wave forecasting and for research. It includes 'state-of-the-art' formulations for the description of the physical processes involved in the wave evolution and it is steadily improved. Although the WAM model was originally developed to run at global scale, it has been also adapted to run efficiently with high resolution in shallow water regions (Monbaliu et al. 2000). WAM is coded in standard Fortran (77/90/95) programming language and it is adapted to run on many different computational platforms. Moreover, recent modifications (WAM 4.5.1) allow for parallel multiprocessor computations using High Performance Computing technologies.

2.8 Buoy observations

2.8.1 Measuring technique

A common technique for measuring wave height is to register the heave motion of an object floating on the water surface (i.e., buoy). This can be done directly by measuring the vertical acceleration of the buoy with an electronic instrument called *accelerometer*. The basic principle of the accelerometer is to detect the non-gravitational forces acting on a mass-spring-like system. The vertical displacement is obtained by integrating twice the acceleration with respect to time. In order provide the accelerometer with a vertical axis, the sensor is mounted on a sphere of stabilizing liquid that serves as an artificial horizon (Datawell manual 2006, Holthuijsen 2007).

Directional information can be obtained by measuring the buoy's horizontal accelerations with two extra accelerometers fixed to the north-south and east-west axes of the buoy. Complementary, pitch and roll angles of the stabilizing liquid are determined by an array of coils mounted on the buoy, and the geographic north is determined with an onboard compass (Datawell manual 2006).

Power spectral density is obtained from the times series of the vertical displacements by applying standard Fast Fourier Transform techniques. Some information about the directional distribution can be derived from the vertical and horizontal displacements by applying standard cross-spectral analysis. From that analysis wave directional parameters per frequency (e.g., mean wave direction, directional spread)

can be obtained. (Longuet-Higgins et al. 1963, Kuik et al. 1988, Datawell manual 2006).

2.8.2 Reconstruction of the buoy 2D spectrum

For the reconstruction of the two-dimensional spectrum, the power spectral density can be conveniently expressed as:

$$F(f,\theta) = F(f)D(f,\theta) \qquad [2.26]$$

Where $F(f,\theta)$ is the two-dimensional power spectral density, $F(f)$ is the power spectral density in the frequency domain, and $D(f,\theta)$ is the directional distribution function. A commonly used shape for the directional distribution function is given by Longuet-Higgins et al. 1963 and Kuik et al. 1988:

$$D(f,\theta) = A\cos^{2s(f)}\left(\frac{\theta-\theta_0(f)}{2}\right) \qquad [2.27]$$

Where A is a normalization factor, $s(f)$ is the directional width parameter, and $\theta_0(f)$ is the mean wave direction. The normalization factor A is computed in such a way that the following property of the directional distribution is satisfied:

$$\int_0^{2\pi} D(f,\theta)d\theta = 1 \qquad [2.28]$$

And the directional width parameter $s(f)$ is computed from the directional spreading $\sigma(f)$, following Kuik et al. 1988:

$$\sigma(f) = \left(\frac{2}{s(f)+1}\right)^{1/2} \qquad [2.29]$$

2.8.3 Observations uncertainty

The uncertainty associated to the observed significant wave height can be estimated from the uncertainty associated to the spectral estimates, given that H_{m0} is related to the mean spectral energy (i.e., equations 2.18 and 2.19). Since the estimate of the variance calculated from the spectrum follows a chi-square distribution, it can be sown that (Young 1999):

$$\left[\frac{\upsilon_s}{\chi^2_{\upsilon_s;1-\alpha/2}}\right]^{1/2}\widehat{H}_{m0} \le H_{m0} < \left[\frac{\upsilon_s}{\chi^2_{\upsilon_s;\alpha/2}}\right]^{1/2}\widehat{H}_{m0} \qquad [2.30]$$

where $\chi^2_{\upsilon_s;\alpha}$ is the α percentage point of the chi-square distribution, and υ_s is the number of degrees of freedom of the spectrum calculated as:

$$\upsilon_s = \frac{n\left[\sum_{j=1}^{M}\widehat{F}(f_j)\right]^{\wedge 2}}{\sum_{j=1}^{M}\left[\widehat{F}(f_j)\right]^2} \qquad [2.31]$$

where M is the number of frequency bands and n is a counter specifying the order of averaging in the recorded water elevation in the Fourier analysis.

2.8.4 Data sources

Buoy measurements within the study area are available for the Belgian continental shelf through the monitoring network of the Flemish Community, Waterways and Marine Affairs Administration (Afdeling Kust|MDK-VLIZ 2008). Directional and non-directional measurements are available at a number of locations. The directional data delivered consist of energy spectral density, man wave direction, and directional wave spreading. Non-directional data contain only energy spectral density. Additionally, there are a number of Dutch buoys in the area operated by the Ministerie van Verkeer en Waterstaat of the Nederlands (ZEGE Meetnet 2008). In figure 2.5 some of the buoy locations within the study area are indicated.

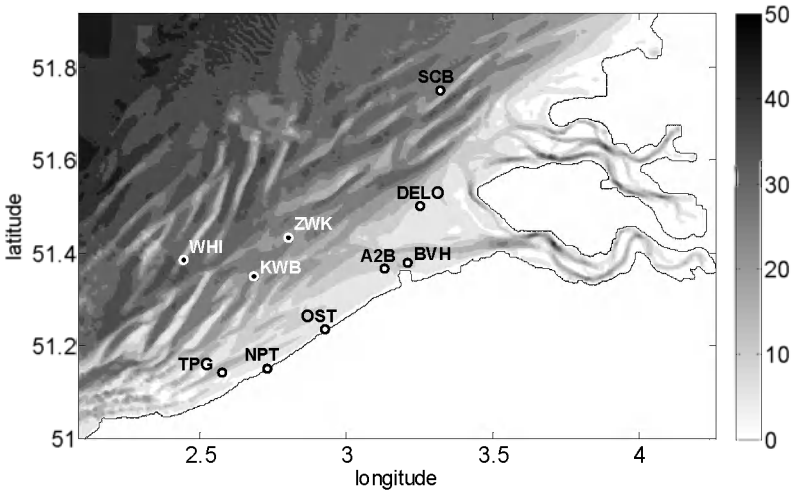

Figure 2.5. Monitoring buoys from the Afdeling Kust (MDK) - VLIZ and ZEGE networks deployed in the study area.

Table 2.2. Buoy locations in the study area.

Buoy	Location	Network	Type	longitude	latitude	Depth MSL
WHI	Westhinder	MDK-VLIZ	Waverider (2D)	2°26'40"	52°23'04"	26.0 m
BVH	Bol van Heist	MDK-VLIZ	Waverider (2D)	3°12'29"	51°22'45"	10.6 m
ONS	Oostende	MDK-VLIZ	Waverider (2D)	2°54'51"	51°14'29"	6.1 m
AKZ	Akkaert Zuid	MDK-VLIZ	Waverider (1D)	2°48'09"	51°25'59"	20.5 m
A2B	A2 Boei	MDK-VLIZ	Waverider (1D)	3°07'43"	51°21'57"	8.7 m
TRG	Trapegeer	MDK-VLIZ	Waverider (1D)	2°34'30"	51°08'30"	10.5 m
DELO	Deurloo	ZEGE	Waverider (2D)	3°14'36"	51°30'13"	15.0 m
SCB	Schouwenbank	ZEGE	Waverider (2D)	3°18'18"	51°44'48"	20.0 m

2.9 Wave modelling in the Belgian continental shelf

2.9.1 Study area

The Belgian continental shelf is located in the southern North Sea. It is constituted by a relatively complex and shallow bathymetry characterized by the presence of several sand banks. From the socio-economical point of view, this area is subject to extensive pressure from different marine related activities.

Wave conditions in the study area are characterized by the presence of wind-seas driven by the moderated local meteorological conditions, and the presence of swells generated in the North Atlantic, that reach the area from the north flank. Often, the coexistence of these two systems results in rather complex wave spectra where the wind-sea and swell systems overlap in their spectral domains (mixed seas) or give origin to bimodal spectra.

2.9.2 Model grid

As swell systems are often present in the study area, a proper modelling setup must account for its representation. For this reason, the wave model domain is extended towards the north to cover the whole North Sea and part of the Norwegian Sea where swells are generated. From the practical point of view, a detailed spatial resolution grid over the whole North Sea is undesirable in view of the high computational demands, but it is required, on the other hand, at the study area to account for bathymetric effects in the wave evolution. Considering that the wave model used (WAM) is a finite differences model, this issue is solved by a series of

nested grids. For the present study, three nesting levels are considered (see figure 2.6 and table 2.3).

Wave boundary spectra are passed from the Coarse grid to the first nested grid (Local 1) and from the Local 1 grid to the second nested grid (Local 2). Initial wave conditions are generated from user-defined JONSWAP parameters. It is expected that the initial (cold start) condition will quickly be overridden by the wind-generated waves. Further details regarding this implementation are addressed in Osuna (2002).

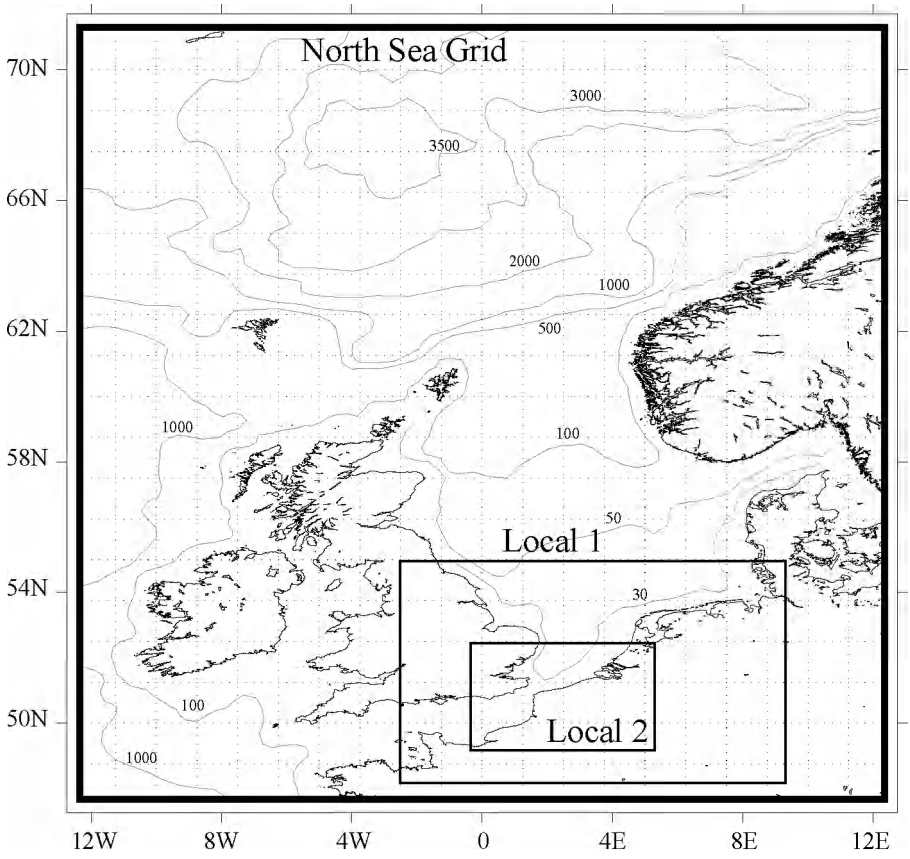

Figure 2.6. The North Sea model grid domain and nested grids Local 1 and Local 2. Gray continuous lines are bathymetric contour lines at the indicated depths given in meters. The dotted grid lines correspond to the UKMO wind fields grid used for the modelling tests.

Table 2.3. Model nested grid characteristics

	North Sea		Local1		Local2	
RESOLUTION	**(°)**	**(Km)**	**(°)**	**(Km)**	**(°)**	**(Km)**
Latitude resolution	1/3	~36.6	1/15	~7.2	1/45	~2.4
Longitude resolution	1/2	~47.7	1/10	~8.7	1/30	~2.8
COORDINATES						
Most Northern Lat. (°)	71.17		55.50		52.63	
Most Southern Lat. (°)	12.25		48.50		49.23	
Most Eastern Long. (°)	47.83		9.25		4.75	
Most Western Long. (°)	-12.25		-2.75		0.05	
COMPUTATIONAL SETTINGS						
Propagation time step (s)	600		200		50	
Source time step (s)	600		600		600	
Grid Size	50x71		121x106		142x154	
Run time for 12h*	00:01'		00:04'		00:23'	
* using a single processor machine (Pentium III 734 MHz)						

2.9.3 Model options and parameters

The following model setup has been used for the runs carried out in this study:

Spectral resolution:
number of frequencies: 25
minimum frequency: $(1.1)^{15}/100 = 0.041772$
frequency resolution: $f_{i+1} = 1.1 * f_i$
directional resolution: 30°

Wind input:
wind input time step: 6h
wind field grid resolution: 1.25°

Physical mechanisms:
propagation option: spherical
shallow water option: on
refraction option: on
bottom friction: JONSWAP empirical formulation
depth induced wave breaking: off
wind input formulation: Janssen (1991)

nonlinear wave interactions:	DIA, Hasselmann et al. (1985)
white capping dissipation:	Günther et al. (1992)
currents fields:	No
water depth variations:	No

2.9.4 Bathymetry

The bathymetric data used correspond to the CSM bathymetry (Flather 1981), derived from the digitalization of navigation charts, which has been tested and verified in former modelling studies in this area (Luo 1995, Bidlot 1995, Monbaliu 2000, Osuna 2002). At present a finer resolution bathymetry (~250m), product of different measuring campaigns is available for this region (Van Lancker et al. 2005). This bathymetry was not used in the present study because the grid implementation using that bathymetry has not been yet systematically verified. In addition, preliminary tests using a high-resolution bathymetry and model grid showed similar results compared to the results of the Local 2 grid, Van Lancker et al. 2004. Moreover, the computational cost of a high-resolution model grid increases considerably. In view of these considerations, model refinement up to the level of the Local 2 grid was considered sufficient for testing the assimilation scheme.

2.9.5 Wind fields

The wind fields correspond to analyzed winds from the United Kingdom Meteorological Office (UKMO) available through the Management Unit of North Sea Mathematical Models (MUMM). Wind fields are provided at synoptic hours (i.e., every 6h), at the reference height of 10m with a spatial grid resolution of 1.25°. The wind field grid domain covers an area from 32.5°N to 75°N and from 70°W to 35°E.

It is important to note that in the standard model setup used, UKMO wind speed is increased by 10% (wind factor 1.1) in consideration of former implementation results (Monbaliu et al. 1999), and calibration and validation tests (Anonymous 2002). Comparisons of model results using UKMO winds versus altimeter data, made by Monbaliu et al. 1999, show that wave heights from the model are in general lower. They also indicate that using 6 hourly winds results in underestimation of wave heights for the largest waves compared to estimations using hourly winds. In the calibration and validation tests made by Anonymous 2002, emphasis is given to the wind factor as it is the principal model calibration parameter for the coarse grid. Using a wind factor of 1.0, they obtained systematic underestimation of wave height for the three stations considered (AUK, K13, and WHI), although overestimation

was also present in some time periods. With a factor of 1.1 they obtained slight overestimations. The conclusion was that the optimal wind factor was between 1.05 and 1.1 but the factor of 1.1 was chosen in order to better represent extreme conditions.

2.9.6 Simulation period

The period of January 2007 has been considered as a study case for the data assimilation experiments. During this period, considerable wind and wave activity are present. In addition, the occurrence of moderated wind speeds, and the occurrence of low frequency waves are registered. This is important in order to evaluate the effect of assimilation under different wind and wave conditions. In figure 2.8, predicted wind conditions from the UKMO are given for the location of Westhinder. In figure 2.8 the values of U_{10} shown are already affected by the factor 1.1 as referred in section 2.9.5 and used by the standard setup of the WAM model used.

A comparison between measured and estimated wind speeds is given in the scatter plot of figure 2.7. Model and observations are compared at every model output time step. No processing is applied to wind speed measurements (e.g., filtering). Looking at figure 2.7, it can be seen that wind speeds tend to be underestimated (bias= -0.4193) even with the 10% compensation. Nevertheless, there is in general a good agreement between data and model winds.

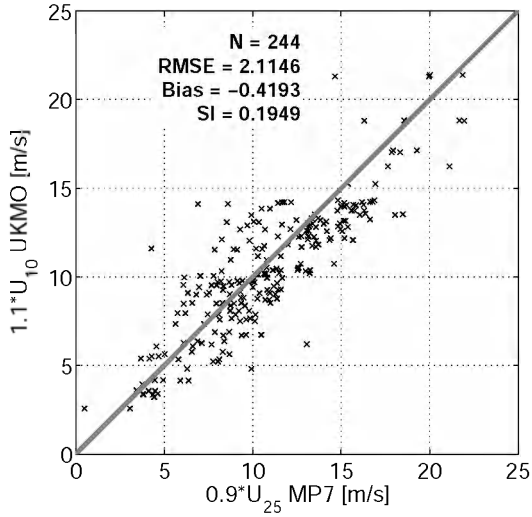

Figure 2.7. Scatter plot of observed versus predicted wind speed (U_{10}) at WHI (MP7). Data are compared every model output time step (3h).

Observed wind conditions from the meteorological station Meteo Paal 7 (MP7), located nearby Westhinder, are shown in figure 2.9. Wind speeds at MP7 are measured at 25m, U_{10} has been approximated as $U_{10} = 0.9*U_{25}$.

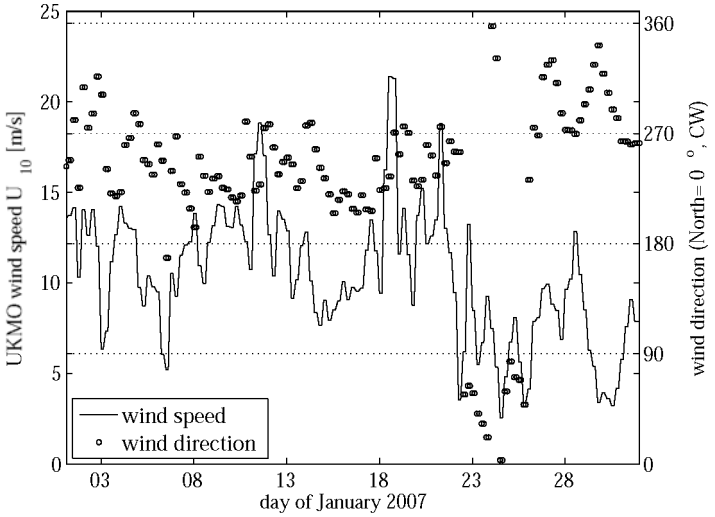

Figure 2.8. Estimated wind from the UKMO at WHI (MP7), for the period of January 2007. U_{10} is multiplied by a factor of 1.1 as used in the normal model run. Time step corresponds to wave model output time step (3h).

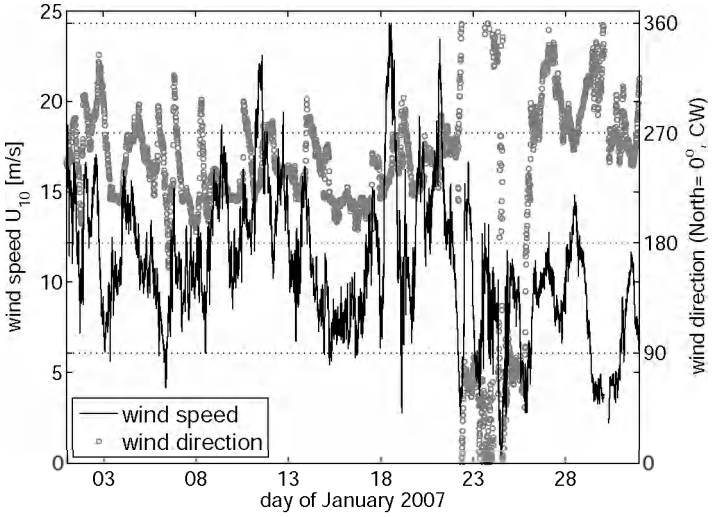

Figure 2.9. Measured wind speed ($U_{10} = 0.9*U_{25}$) and wind direction at MP7 for the period of January 2007. The time step is 10min.

2.10 Reference simulation (Normal run)

The WAM model has been run for the period of January 2007 (without data assimilation: *normal run*). The results of this run for the output locations of Westhinder (WHI), Bol Van Heist (BVH), and Oostende (ONS) are presented in figures from 2.10 to 2.17.

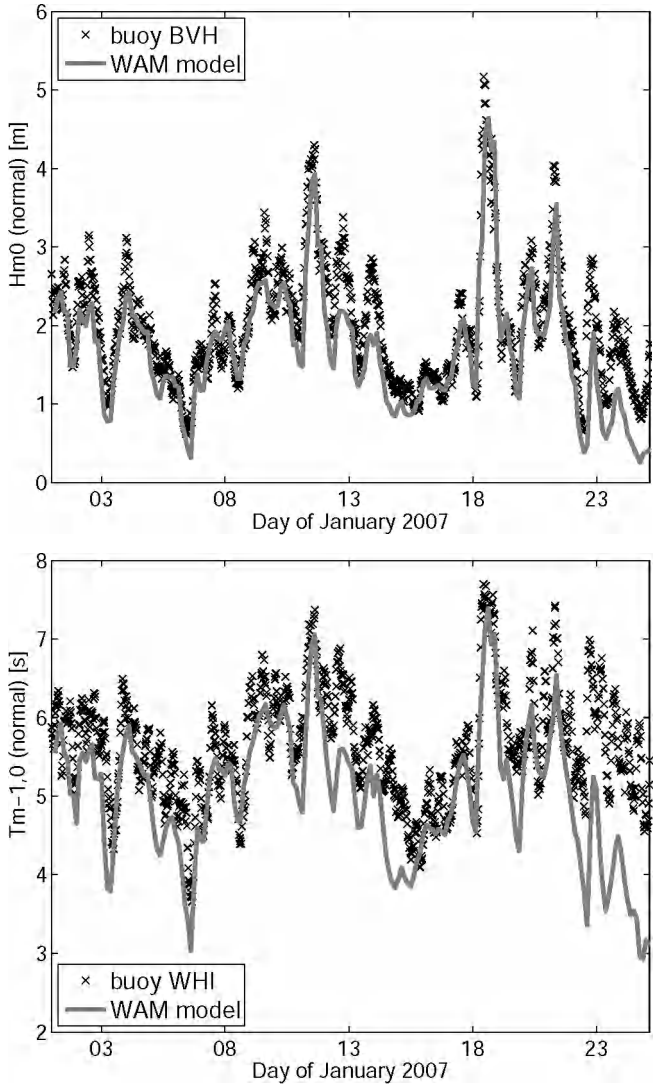

Figure 2.10. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Westhinder for the simulation period of January 2007.

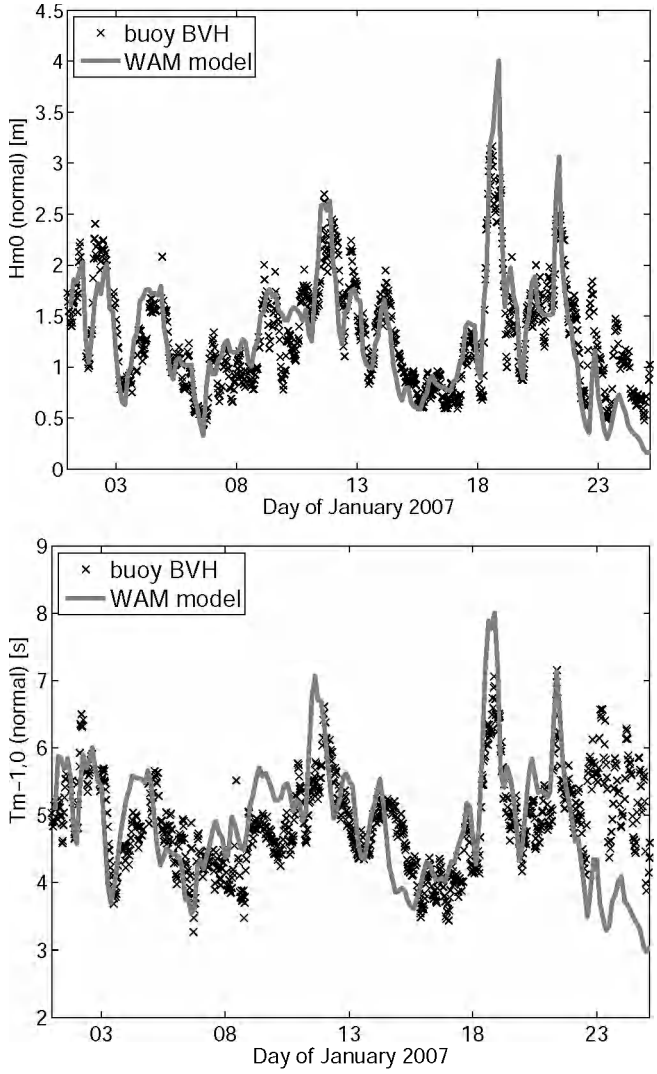

Figure 2.11. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Bol van Heist for the simulation period of January 2007.

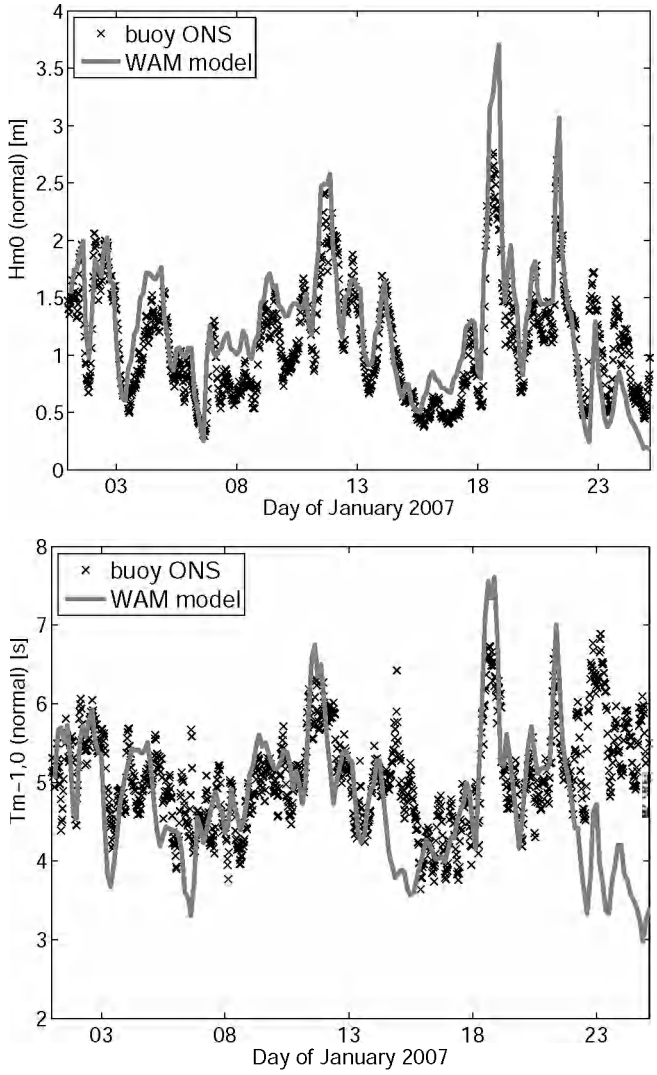

Figure 2.12. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Oostende for the simulation period of January 2007.

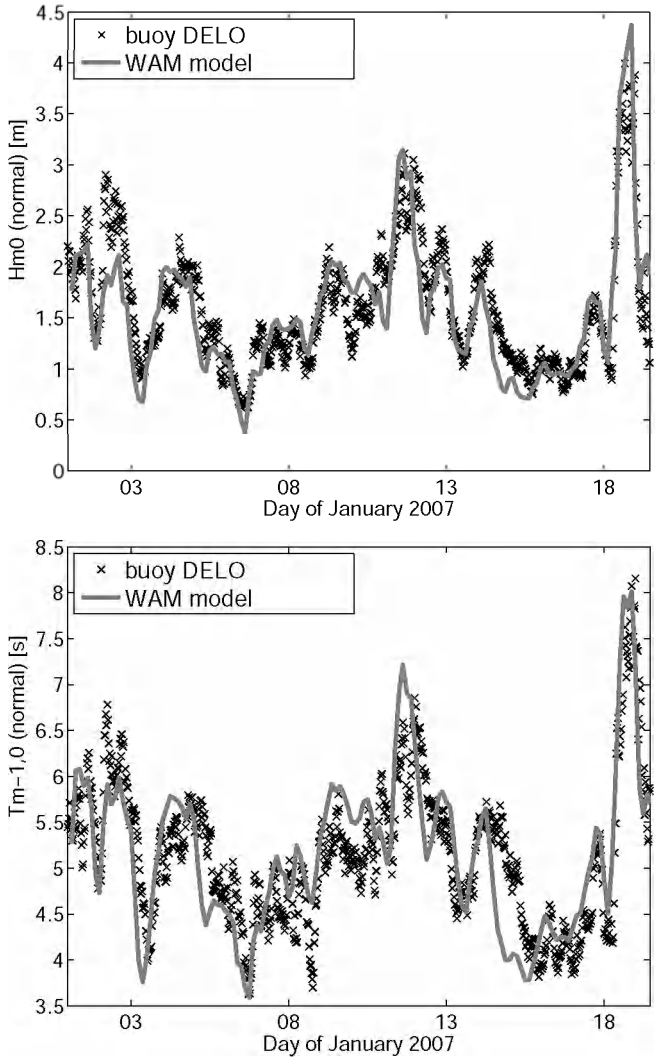

Figure 2.13. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Deurloo for the simulation period of January 2007.

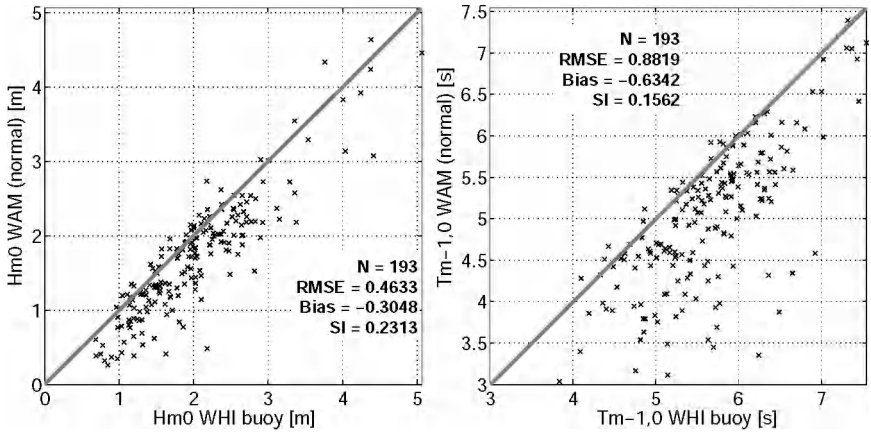

Figure 2.14. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Westhinder for the simulation period of January 2007.

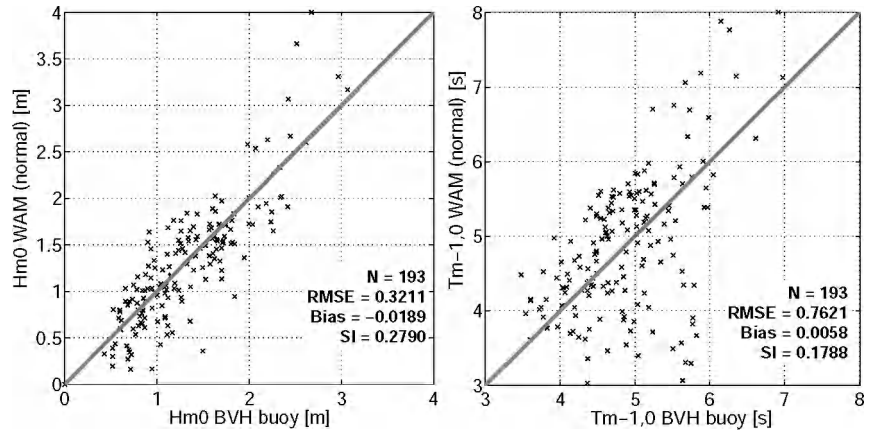

Figure 2.15. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Bol van Heist for the simulation period of January 2007.

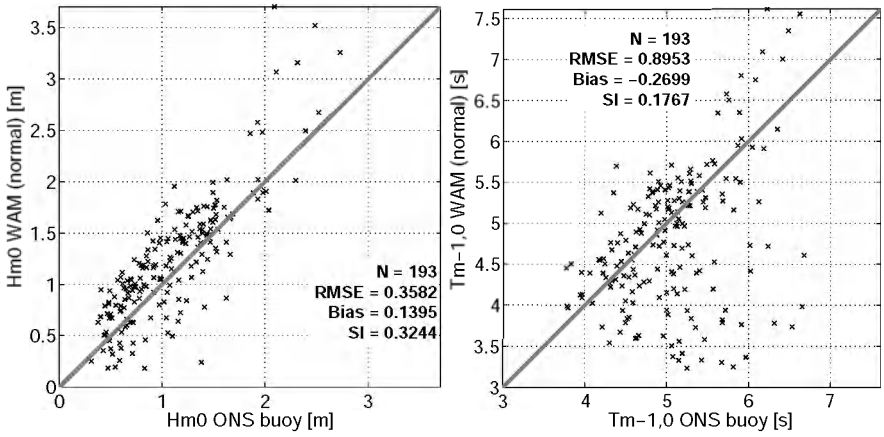

Figure 2.16. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Oostende for the simulation period of January 2007.

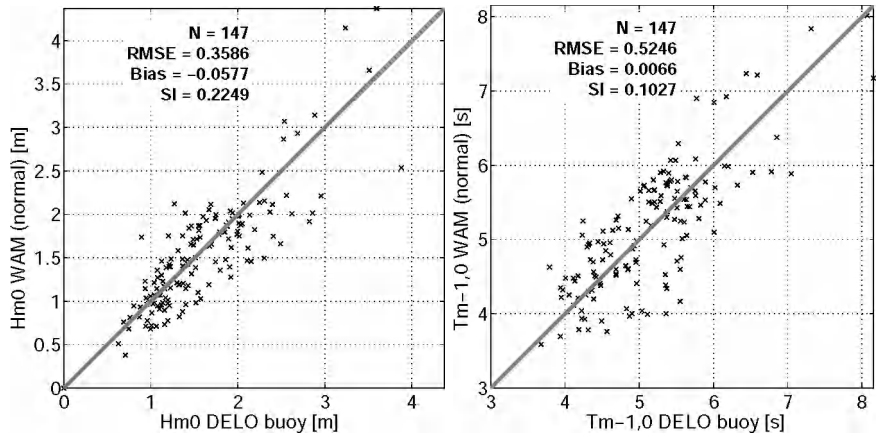

Figure 2.17. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Deurloo for the simulation period of January 2007.

As can be seen from figures 2.10 to 2.17, in general model results are good when compared with buoy data. Scatter indexes are about 30% and lower for significant wave height (H_{m0}), and lower than 20% for mean wave period ($T_{m-1,0}$). A remarkable behaviour of the model is the systematic underprediction of H_{m0} and $T_{m-1,0}$ at WHI, while at BVH that underprediction is less notable. Contrarily, at ONS, H_{m0} is overpredicted.

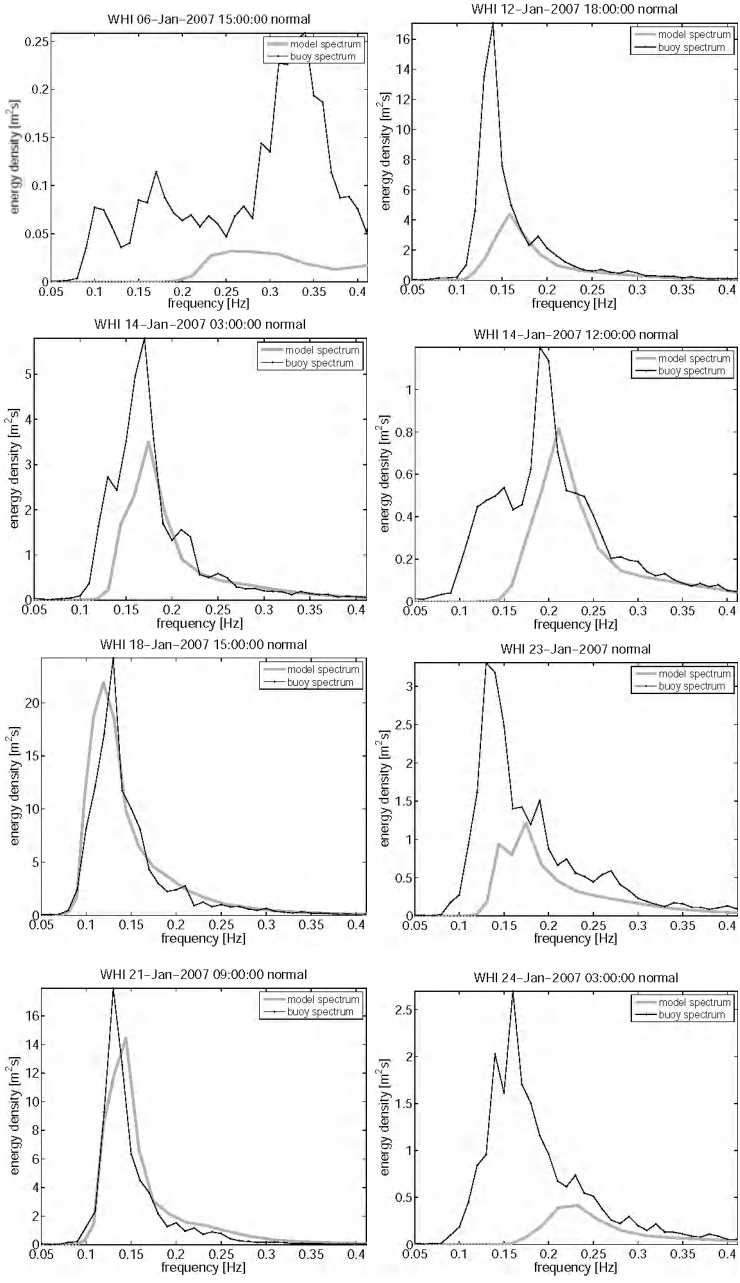

Figure 2.18 Comparison between observed and model spectra from Westhinder (WHI) at different instants of the simulation period.

The most notable model mispredictions occur at periods around the 14th and after the 21st of January 2007, these periods are very likely to contain swell waves, given the combination of relatively low wind speeds, high wave heights and high wave periods, see for instance figures 2.18. The apparent better statistics from DELO appear because the period with considerable swell after the 21-Jan-2007 is not included in the comparison since there is no data for that period at this location. For this reason, measurements from this buoy are not considered for evaluating the effect of data assimilation.

Comparing observed and model spectra at WHI (figures 2.18), it can be seen that the model performs better in high wind and wave conditions than in low conditions. For instance, the correspondence between model and observed spectra is very good at the peak instants of 18-Jan-2007-15:00:00 and 21-Jan-2007-09:00:00. For lower wind and wave conditions the correspondence is less good (i.e., 12-Jan-2007-18:00:00, 14-Jan-2007-03:00:00, 23-Jan-2007-00:00:00). In conditions of energy decay and for lower wave conditions, in which swell is present in the buoy observations (i.e., 06-Jan-2007-15:00:00, 14-Jan-2007-12:00:00, 24-Jan-2007-03:00:00), the model performance is low. In these conditions, the model is not able to represent properly the energy at the low frequency components (i.e., swell).

2.11 Summary and Conclusions

Spectral wave models as the WAM model solve a linear approximation of the energy balance equation. In deep water, sink-source terms comprise the energy input by wind, the interchange of energy amongst interacting resonant waves (quadruplets), and the dissipation due to wave breaking (white-capping). In shallow water and in the presence of currents, as wave energy is not conserved, the balance equation is written in terms of the action density. In addition, other sink mechanisms come into play as bottom friction, depth-induced breaking, and triad interactions.

At present spectral wave models perform very well. However, some of the mechanisms affecting the wave evolution are still poorly understood and they are the subject of further research and observations.

In the present model setup, the WAM model Cycle VI is used to hindcast waves in a nearshore environment. Four sink-source mechanisms account for the total source term: wind input, quadruplet interactions, white-capping dissipation, and bottom friction. Although other mechanisms, as depth-induced wave breaking and the influence of water depth variations and currents, are expected to play an important

role in extremely shallow conditions, they have not been tested within this study since the main attention is given to the data assimilation scheme.

In the present WAM model setup, analyzed wind fields from the UKMO are used. UKMO wind speeds are expected to be slightly underestimated and are multiplied by a factor of 1.1. This relatively high wind factor is meant to obtain better model results in storm conditions.

The performance of the WAM model in the present study area is good. The mean spectral parameters evaluated, significant wave height and mean wave period, show good agreement compared with buoy measurements. The agreement between model and observations is the best under high wind and wave conditions, both in terms of mean parameters and spectral shapes. For lower wind and wave conditions that agreement is less good.

Chapter 3

Spectral partitioning and identification of wind-sea and swell

3.1 Introduction

An ocean wave spectrum describes the distribution of the total wave variance over frequency and direction. Such a distribution is the result of the occurrence of a certain number of individual wave systems originating from different meteorological events. For the interpretation and archival of large data sets, integral parameters rather than whole spectra are preferred. However, while integral parameters describe suitably a wave spectrum composed of a unique wave system the simultaneous occurrence of different wave systems turns integral parameters less meaningful, unless they refer to individual wave components. Partitioning of wave spectra into independent wave systems provides an excellent tool for data reduction. Also for the comparison of data sets or when evaluating model performance, the analysis at the level of wave systems gives more insight in processes than the analysis of mean parameters of the whole spectrum (e.g., Portilla and Monbaliu 2007). For data assimilation purposes, the use of spectral partitioning has given rise to the development of more robust sequential algorithms (Hasselmann et al. 1996, Young and Glowacki 1996, Voorrips et al. 1997). Previous assimilation schemes faced constraints for updating the individual bands of the spectrum, as there is no reason to change the partial contribution of each individual system in absence of additional information (Thomas 1988, Lionello et al. 1992). Additionally, spectral components can be associated in space and time to trace the evolution of wave systems originating from remote storms (Hanson and Phillips 2001, Quentin 2002).

One of the potential applications of partitioning in third generation wave modelling is the determination of wind-sea and swell. As these components are no longer computed separately, and model users have traditionally disposed of such information, there is a need for splitting the spectrum in order to provide this

information as part of the output products (Bidlot, 2001). On the other hand, in wave studies where the identification of wind-sea and swell is relevant nearly each author has adopted his/her own criteria based on some physical properties of wind and waves and therefore several methods can be found in literature (e.g., Wang and Hwang 2001, Violante-Carvalho et al. 2002, 2004).

For consistency, in the rest of the chapter *partitioning* will be seen as the mechanism to detect wave systems looking at morphological features of the spectral signature only. *Identification* on the other hand refers to labelling with wind-sea or swell as a supplementary designation taking into account environmental and physical characteristics.

For the present study, different partitioning-identification schemes available in the literature are implemented and compared. The description of the different methods is structured in two main blocks, the first considering partitioning methods only (section 3.2), and the second considering the identification step (section 3.3). Within each of these two blocks, 2D and 1D schemes are treated. The analysis points at strengths and shortcomings and where possible a more robust scheme is proposed. Examples at the end of each section illustrate the findings (Portilla et al. 2009).

3.2 Partitioning methods

3.2.1 2D partitioning schemes

The first conceptual partitioning algorithm was presented by Gerling (1992). In his algorithm, the lowest energy threshold value is found at which upper parts of the spectrum get disconnected. This process is repeated until all systems are detected. In order to determine if partitions are significant or not, integral mean parameters are compared with spectral components of neighbouring points and subsequent times (pattern-extraction algorithm). A partition is considered significant if it is persistent in time and space.

Many partitioning schemes (e.g. Hanson and Phillips 2001, Voorrips et al. 1997, Lefevre et al. 2005) are specific implementations of the scheme described by Hasselmann et al. (1996). The basic idea of this scheme is the same as the one of Gerling (1992) although the concept of the algorithm differs slightly. According to Hasselmann et al. (1996) a wave spectrum can be regarded as an inverted catchment area, making analogy with hydrological concepts (see also Brüning et al. 1994 and Hasselmann et al. 1994). The different sub-catchments of that main catchment area are determined associating each spectral grid point to a unique neighbour, namely

the one with the highest energy level. Grid points corresponding to the same local peak are clustered and each of these clusters defines a partition (watershed algorithm). In order to assess the significance of the partitions, some of their morphological characteristics are inter-compared. In Hasselmann et al. 1996, two partitions are merged into one:

- if two peaks are one grid cell apart,
- if the trough between them is not sufficiently pronounced (i.e., the lowest point between two partitions is greater than 85% of the smaller peak),
- or if the square spectral distance between two peaks is shorter than the spread of any of the two systems (see table 3.1 for definitions).

Other authors (e.g., Voorrips et al. 1997, Hanson and Phillips 2001) have used the scheme of Hasselmann et al. (1996) with different settings for the combining parameters. These implementations are briefly described in the next section.

3.2.1.1 Discussion

Actually, partitioning results from the two methods above are similar. However, assessing whether those systems are significant or not is less straightforward. This is especially the case for observed spectra as these contain considerable random variability. Model spectra do not contain such random variability, as they are not subject to noise interference. In addition, the generating wind field in the model is more uniform, therefore the signal is not affected by subscale processes. However, assessing the significance of partitions in model spectra will become more problematic as spectral resolution increases.

Gerling's (1992) approach is consistent if several observations of the same network or if model spectra are processed. In both cases, wave systems can be inter-compared and their persistence can be assessed. A practical limitation is that it demands the availability of other spectra. Moreover, the number of partitions detectable in an observed spectrum is typically of the order of tens, thus associating several wave components with neighbouring components at different times becomes very if not too intricate. Gerling (1992) already pointed at this when he states "… *It does not appear possible to obtain completely satisfactory results with the simple metric just defined. …* "

Hasselmann's et al. 1996 approach does not suffer from these limitations. In their scheme, they propose to inter-compare spectral features within the spectrum. However, the criteria used for merging partitions rely on rather arbitrary parameters that need to be adjusted from situation to situation. Moreover, different users of this scheme have adopted different parameters (Voorrips et al. 1997, Hanson and

Phillips 2001). To this end, Hanson and Phillips (2001) suggested the need of an additional routine that optimizes the choice of the parameters by an iterative procedure. Moreover, they emphasized the need for removing partitions with energy under a threshold value determined by the spectral fall-off given by Phillips (1985). However, it is not evident that those small partitions should actually be removed. Voorrips et al. (1997) simply merge partitions with low energy (i.e., lower than 0.0025m^2). Table 3.1 summarizes the parameter settings for the different implementations described above.

Table 3.1 Summary of parameters for combining partitions in the 2D spectrum according to different implementations

	Low energy threshold	contrast	$\Delta f^2 / \delta f^2$
Hasselmann et al. (1996)	-	0.85	1
Voorrips et al. (1997)	0.0025 m^2	0.70	0.5
Hanson and Phillips (2001)	$e \leq A/(f_p^4+B)$ **	0.65 – 0.75	0.4 – 0.5

* The squared distance between two spectral peaks (1) and (2) is defined as:

$\Delta f^2 = (f_x^{(1)} - f_x^{(2)})^2 + (f_y^{(1)} - f_y^{(2)})^2$ where $f_x = f_p \cos\theta_p; f_y = f_p \sin\theta_p$,

f_p is the peak frequency of each wave system and θ_p is its direction

* The average spectral spread of a wave system is defined as:

$\overline{\delta f^2} = \overline{(f_x - \overline{f_x})^2} + \overline{(f_y - \overline{f_y})^2}; \overline{f_x} = \overline{f\cos\theta}; \overline{f_y} = \overline{f\sin\theta}$

** e is the wave system energy, f_p is the peak frequency, and A and B are calibration parameters

While spectral partitioning is conceptually a robust and simple method, the need for continuous calibration becomes tedious especially in operational or automated conditions. Moreover, inappropriate choice of combining parameters renders the method unstable and unreliable. The combining mechanism is crucial because it determines which partitions are significant and how those partitions are merged to determine the resulting wave systems. Similar remarks referring to the complexity of determining significant partitions were pointed out before by Aarnes and Krogstad (2001).

3.2.1.2 *Proposed 2D partitioning algorithm*

The previous section pointed out that the calibration of the combining algorithm is the main difficulty in producing meaningful partitions. In general, adjusting parameters for one situation produces deficient results in others. This will be illustrated with example 1. In this section, an image-processing tool is introduced in the combining algorithm, aimed at alleviation of the parameterization dependence. The 2D wave spectrum is thus treated as an image. As in many cases (either for observed or model spectra) the main problem is the presence of spurious partitions. A 2D noise removal (smoothing) filter has been implemented and tested with satisfactory results. This filter consists of a 2D discrete convolution operation between the spectrum and an equally weighted convolution kernel that averages all immediate neighbours of a central bin. That operation is mathematically expressed as:

$$\widehat{F}(i,j) = \kappa(m,n) \otimes F(i,j) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} \kappa(m,n) F(i-m, j-n) \qquad [3.1]$$

where $\widehat{F}$ is the filtered spectrum, F is the raw spectrum, both having dimensions i x j. The operator $\otimes$ indicates a convolution. The convolution kernel κ is chosen as a constant *3 x 3* matrix with coefficients summing together to unity (i.e., κ (m,n) = $^1/_9$, $\forall$ m,n).

Obviously, different possibilities exist for the choice of the kernel and also the *spectral image* might be subject to more elaborated image processing. However, a setup including this filter seems to perform well in most typical circumstances. Note that this filtering process can be repeated, and an important aspect to be addressed is to which extent the wave spectrum has to undergo repeated filtering. It is clear that more spurious partitions are present in observed spectra than in model spectra and will require more filtering. On the other hand, excessive filtering causes *blurring,* that may render patterns indiscernible. Two measures are taken to tackle this aspect: the first is to indicate *a-priori* a number of expected significant systems in the spectrum, and the second is to merge partitions with low energy by setting a noise energy threshold (*thresholding*).

The partitioning-combining method advocated here is set up as follows:

1. The spectrum is partitioned with the watershed algorithm;
2. Low energy partitions are merged (thresholding);
3. If the number of partitions is higher than the prescribed number the spectrum is filtered, partitioned and low energy partitions are merged (thresholding);

4. Step 3 is repeated until the number of partitions detected is equal or lower than the prescribed number;
5. Finally, low energy partitions are merged (combining).

The degree of filtering is thus determined implicitly by the prescribed number of partitions and the noise threshold. Note that the thresholding (step 2) and the last combining due to low energy (step 5) are carried out by the same subroutine, but are conceptually two different operations. Thresholding aims at suppressing noise since if the reduction of partitions is carried out by filtering alone, excessive filtering would be needed and the blurring effect would be stronger. Combining aims to disregard small systems that are probably real systems but are not necessarily important. The thresholding-combining subroutine merges the target partition with the closest adjacent partition in the frequency-direction space. Tests in a number of different circumstances with buoy and model spectra suggested the following settings:

- number of expected partitions: between 4 and 6
- energy level for noise thresholding: between 1 and 2% of the total energy in the spectrum.
- energy level for last combining: between 5 and 8% of the total energy in the spectrum.

3.2.1.3 Example 1

The 2D partitioning-combining procedures are illustrated using wave spectra from the NDBC buoy 41013. The period from 10-Apr-2006 00:00:00 to 12-Apr-2006 06:00:00 UTC was chosen. During this period the wave conditions are characterized by double peaked spectra, shown in figure 3.1 (for clarity 1D spectra are shown).

These spectra have been partitioned and combined according to the criteria of Hasselmann et al. 1996 (Table 3.1), and also using the combining algorithm proposed in this study (section 3.2.1.2). Time series of wave energy and mean frequency are presented in figure 3.2. Note that in order to draw figure 3.2, partitions in consecutive spectra need to be numbered in a consistent manner. For all combinations of partitions of the current and of the previous time step the difference between the mean frequency is calculated. Combinations closest in mean frequency are assigned the same partition number.

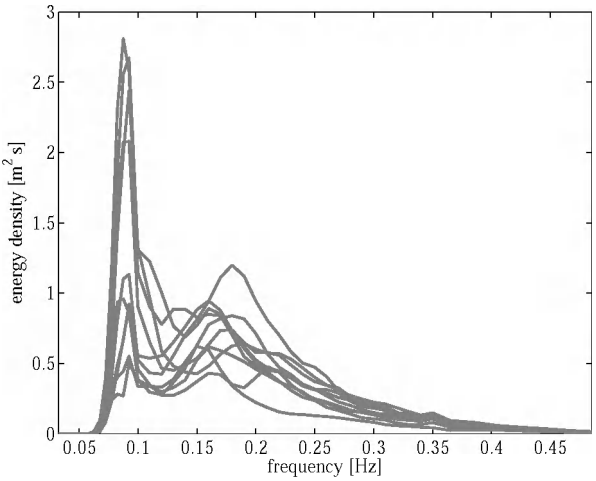

Figure 3.1. 1D energy spectra from NDBC buoy 41013 (33°26'11" N 77°44'35" W) from 10-Apr-2006 00:00:00 to 12-Apr-2006 06:00:00 UTC. Spectra are drawn every 6 hr.

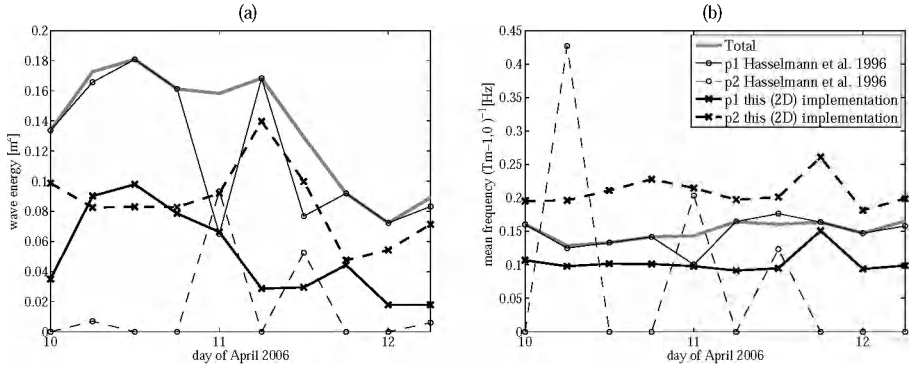

Figure 3.2. Time series of (a) wave energy (b) mean wave frequency ($f_{m-1,0}$), for NDBC buoy 41013 (33°26'11" N 77°44'35" W) for the period 10-Apr-2006 00:00:00 to 12-Apr-2006 06:00:00 UTC, for the whole spectrum (gray thick line). For partitions calculated with Hasselmann et al. (1996) scheme: First partition (continuous circle-marked line) and second partition (dashed circle-marked line). Results from this study 2D implementation. First partition (thick continuous cross-marked line) and second partition (thick dashed cross-marked line).

Using Hasselmann's et al. (1996) scheme, there is only one main partition most of the time (thin continuous circle-marked line in figure 3.2), although a second partition appears and disappears at some occasions. The mean parameters of the first

partition are relatively stable and agree well with those of the total spectrum, while two systems are discernible from the spectra (figure 3.1). The mean parameters of the second partition look more like pure noise. The combining algorithm proposed here detects two wave systems equally significant. Their evolution in time is quite stable. This is in agreement with what it is expected from the spectra (figure 3.1).

In order to analyze details, the spectrum at 10-Apr-2006 18:00:00 UTC is shown in figure 3.3. Visual inspection indeed suggests that two main wave systems are present (note in fact that also, a third one is distinguishable in the 2D spectrum, but its energy is much lower). In order to facilitate the discussion, one of these systems has been contoured (thick red line) in figure 3.3a.

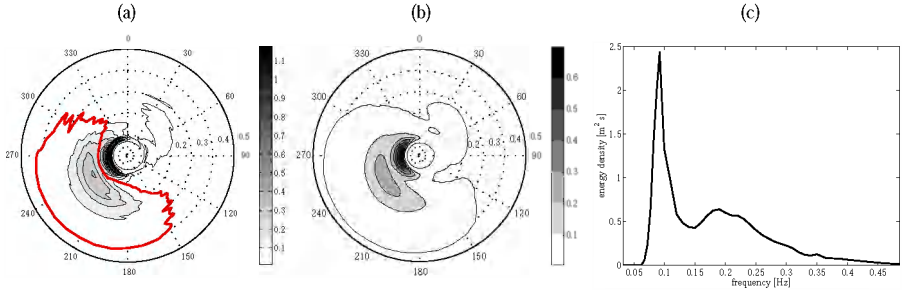

Figure 3.3. Spectrum from NDBC buoy 41013 (33°26'11" N 77°44'35" W) at 10-Apr-2006 18:00:00 UTC. (a) 2D spectrum, (b) 2D smoothed spectrum, and (c) 1D spectrum.

The watershed algorithm detects 16 partitions in this spectrum. According to the Hasselmann's et al. 1996 criteria, all systems are merged into one partition. The condition for which the two (indicated) main partitions are combined is that the square distance between the two peaks (i.e., 0.0088 Hz^2) is lower than the spread of either system (i.e., 0.0183 and 0.0259 Hz^2). On the other hand, according to the 2D implementation of Voorrips et al. 1997 (see Table 3.1) the spectral peak distance must be lower than 0.5 the spread of either system but even that condition is not yet sufficient to keep these two main systems uncombined. Alternatively, following Hanson and Phillips 2001 that distance must be lower than a factor 0.4 the spread of either system. Although this factor seems appropriate in this case, these two partitions would be combined in a further step with their set of parameters because the trough between the two peaks (contrast) is required to be less than 0.65 the energy of the lower peak. The contrast level corresponds in this example to 0.67, which on the other hand satisfies Hasselmann's et al. (1996) and Voorrips' et al. (1997) contrast condition (i.e., 0.85 and 0.70 respectively). Conveniently, one could

choose other factors for this case (i.e., 0.4 for the spectral spread and 0.70 for the contrast) without guarantee that these factors will work for the other spectra.

Using the combining procedure proposed here, the sensitivity to parameter settings is alleviated and the ability of the method to detect relevant spectral features is enhanced. For the present example, after the spectrum has been smoothed once (figure 3.3b), the watershed algorithm detects 5 partitions instead of 16. From those 5 partitions, 3 have energy lower than 2% of the total energy and are merged by the thresholding step, the low energy combining threshold was set to 5% but it doesn't operate in this particular spectrum. This results in the two main wave systems shown in figure 3.3a.

3.2.2 1D partitioning schemes

The 1D partitioning and combining scheme introduced by Voorrips et al. (1997) is a straightforward adaptation of the 2D scheme of Hasselmann et al. (1996). Similarly, each local peak represents the peak of a wave system. The minima between adjacent peaks constitute the partition limits. The combination of partitions is also done under similar criteria:

- if two peaks are within their spectral width (i.e., peaks are closer than half the width at half the maximum of either of the two peaks)
- if the trough between them is not sufficiently pronounced (i.e., the lowest point between two partitions is greater than 50% of the smaller peak),
- or if the partition energy is lower than a threshold (i.e., $0.0025 m^2$)

Additionally, two extra conditions are adopted to detect significant partitions. The first aims at identifying mixed sea states by comparing the mean direction with the peak direction of two potentially merging partitions in order to split them again. The second also considers wind information in order to combine all potential pure wind-sea states. These two extra conditions should not (or cannot) be considered if neither directional information nor wind information is available.

Rodríguez and Guedes Soares (1999) presented also a method to detect significant peaks. Taking into account the energy variability of the spectrum, they consider that some spurious peaks appear due to natural random fluctuations of the spectral estimates. Significant peaks lie outside the confidence interval of those estimates. That is, if the height of a peak, measured from the previous minimum, is greater than the width of the confidence band. The confidence interval is computed considering a chi-square distribution. The parameters of the chi-square distribution (i.e., number of degrees of freedom and the level of confidence) determine the magnitude of significant and spurious peaks.

Violante-Carvalho et al. (2002) presented another set of criteria to detect significant peaks:

- If two peaks are very close to each other (i.e., closer than twice the spectral resolution, 0.03 Hz),
- if the ratio between two adjacent peaks is lower than 15,
- or if the trough between them is not sufficiently pronounced (i.e., lower limit of the 90% confidence interval of the greater peak higher than the upper limit of the 90% confidence interval of the trough between the peaks).

3.2.2.1 Discussion

Similarly as in the case of the 2D spectrum, the combining mechanism based on contrast and the ratio of peak square distance to spectral spread is deficient and depends strongly on the parameterizations. Therefore, the 1D combining algorithm of Voorrips et al. (1997) based on these criteria suffers of the same shortcomings associated to the 2D scheme of Hasselmann et al. (1996). Moreover, in the scheme of Violante-Carvalho et al. (2002) the comparison between energy levels of adjacent peaks and the magnitude of the trough between peaks are conditions analogous to the contrast criterion and have the same limitations.

The criterion of Rodríguez and Guedes Soares (1999) is also similar to the contrast criterion. However, increasing the number of degrees of freedom of the spectrum has a similar effect as filtering. In that sense, this approach is consistent with the idea used in section 3.2.1.2 to improve the 2D scheme. However, the scheme of Rodriguez and Guedes Soares (1999) was not investigated further in this study, mainly because tests using a convolution function to smooth the 1D spectrum showed that the blurring effect was too aggressive in the case of the 1D spectrum. Consequently, spectral patterns quickly became indiscernible resulting in unsatisfactory overall performance of the scheme. Since the smoothing approach did not contribute to the improvement of the 1D algorithm results are not presented here. However, satisfactory partitioning results were obtained by a mechanism aiming to combine most likely spurious peaks. This scheme is presented in the next section.

3.2.2.2 Proposed 1D partitioning algorithm

In order to disregard the most likely spurious peaks and eventually concentrate efforts in detecting more complex features a simple scheme was implemented. It turned out that once these (most likely spurious) peaks are disregarded the so

determined partitions are rather consistent and these criteria are considered sufficient for the 1D combining mechanism. The procedure to detect peaks as spurious is the following:

1. Partitions having the peak frequency above a certain threshold (i.e., 0.35 - 0.4 Hz). The reason for this measure is that in the tail of the spectrum usually high variability is present. That variability is very difficult to treat while in reality peaks in the tail belong to the wind-sea part.
2. Partitions with low total energy (i.e., lower than 5 - 8% of the total energy).
3. Partitions having few spectral bins before or after the peak (i.e., less than 2 bins).
4. Partitions that are placed between two other (neighbouring) partitions and have lower peak energy level than these two neighbours.

3.2.2.3 Example 2

In the present example the 1D partitioning-combining procedure of Voorrips et al. 1997 (without using wind nor directional information) is compared to the 1D implementation given in this study (section 3.2.2.2). The data set is the same as used in the illustration of the 2D scheme (figure 3.2). The resulting time series of the 1D partitioning for wave energy and mean frequency ($f_{m-1,0}$) are shown in figure 3.4 (a and b respectively).

Similarly as in the case of the 2D scheme (Hasselmann et al. 1996), one main wave system is detected using the 1D scheme of Voorrips et al. 1997 (thin continuous circle-marked line). This first partition contains most of the energy and its main frequency agrees with that of the entire spectrum as well. A second partition appears sporadically. From the spectra (figure 3.1), two significant partitions are expected.

The 1D combining procedure outlined here in section 3.2.2.2 detects the two systems present in the spectra. These time series are also quite consistent with the time series of the 2D scheme (figure 3.2). Obviously certain differences exist between results of the 1D and the 2D schemes (i.e., 11-Apr-2006 18:00:00), mainly because not all of the features that are visible in the 2D spectrum are also visible in the 1D spectrum. However, the evolution of the two systems is very similar.

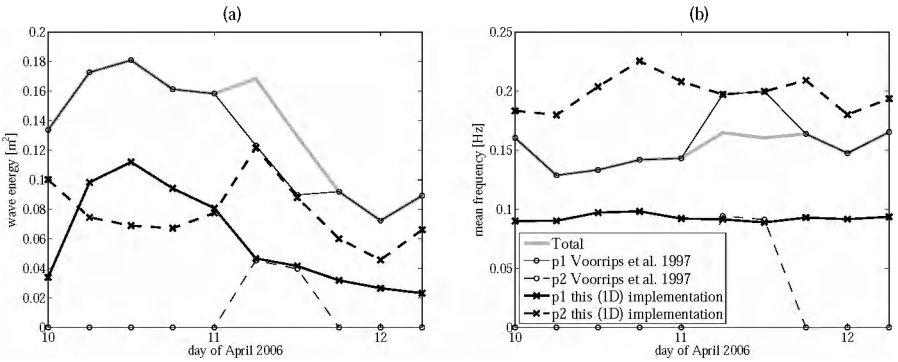

Figure 3.4. Time series of (a) wave energy (b) mean wave frequency ($f_{m-1,0}$), for NDBC buoy 41013 (33°26'11" N 77°44'35" W) for the period 10-Apr-2006 00:00:00 to 12-Apr-2006 06:00:00 UTC, for the whole spectrum (gray thick line). For partitions calculated with Voorrips et al. (1997) 1D scheme: First partition (continuous circle-marked line) and second partition (dashed circle-marked line). Results from this study 1D implementation. First partition (thick continuous cross-marked line) and second partition (thick dashed cross-marked line).

3.3 Wind-sea / swell identification methods

In section 3.2, the detection of different wave systems was done exclusively based on morphological features. In this section, also environmental and physical features are regarded in order to assess the character of their meteorological origin. Locally generated waves growing actively under the influence of wind (wind-sea), and remotely generated waves (swell) arriving to the measuring site are distinguished. Following Holthuijsen (2007), wind-sea waves are more irregular and short-crested. They respond quickly to wind variations and are characterized by a rather broad spectrum. Swell consists of rather regular long-crested waves whose evolution is not as strongly affected by wind. A swell spectrum is narrower and as wind drops, or when waves leave the generation area, their steepness reduces sharply due to frequency-direction dispersion.

From a more practical point of view the energy of wind-sea waves is contained at higher frequencies (i.e., between about 0.1 and 4Hz) while swell waves have lower frequencies (i.e., between 0.03 and about 0.2 Hz). In wave modelling, wind-sea is the part of the spectrum subjected to a positive wind input term (Bidlot 2001).

The distinction between wind-sea and swell is often not obvious. Under changing winds (both magnitude and direction), wave systems can overlap in the frequency-

direction domain giving origin to a rather continuous spectrum where the presence of two or more distinct systems is not clearly discernible. Waves systems in these situations are referred to as *mixed-sea* states and are particularly difficult (if not impossible) to detect and/or identify by automated procedures.

It is evident that more objective and reliable identification algorithms can be constructed when the full 2D wave spectrum and the wind speed and direction are considered. However, in cases when only the 1D spectrum is available, extracting some extra information from it is also advantageous. In the following sections different wind-sea-swell identification methods reported in the literature are studied. Both 2D and 1D spectra are considered.

3.3.1 Wind-sea – swell identification using 2D spectrum and wind data

If 2D spectrum and wind information is available, a straightforward step to identify wind-sea and swell is to apply a definition for wind-sea. Suitably, the definition from numerical modelling might be adopted for which a wind-wave generation formulation must be considered. In the WAM model (Komen et al. 1994) in particular, although the wind wave generation mechanism actually implemented is the one given by Janssen (1991), the identification of wind-sea and swell is based on the formulation of Komen et al. 1984 (equation 3.2). While Janssen's (1991) mechanism takes into account the sea state to compute the wind input, Komen's et al. (1984) formula simply defines a region in the 2D spectrum for the wind input (figure 3.5).

$$\beta \frac{U_z}{c_p} \cos(\theta - \psi) > 1 \qquad [3.2]$$

where U_z is the wind velocity at height z, c_p is phase speed (i.e., $c_p = g/2\pi f$ in deep water), θ is the wave direction, ψ is the wind direction, g is the gravity acceleration, f is the wave frequency, and β is a calibration factor.

Similar criteria based on wave age (U_z/c_p) are used by others to identify wind-seas, see for example Donelan et al. 1985 and Drennan et al. 2003.

The magnitude of the factor β in equation 3.2 is not irrelevant as it directly affects the extent of the wind-sea area in the spectrum. A value of $\beta \leq 1.3$ has typically been applied to characterize the region of *"pure wind-sea"* (Hasselmann et al. 1996, Voorrips et al. 1997, Bidlot 2001). Moreover, Hasselmann et al. (1996) consider *"old wind-sea"* systems those having the peak within the region where $1.3 < \beta \leq 2.0$. Additionally, in order to identify *"mixed-sea"* states produced by a (fast) wind rotation, they impose that either the peak parameters or the mean parameters

(frequency and direction) of a wave system must fulfil the criterion for *"old wind-sea"*.

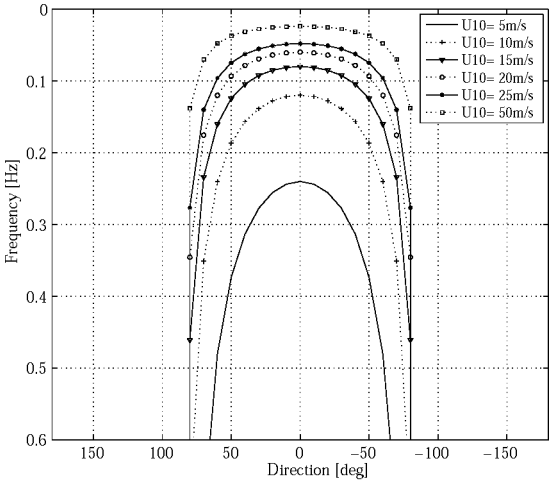

Figure 3.5. Limit of wind-sea and swell in the frequency-direction domain according to equation 3.2 for different values of wind speed (U_{10}) and β = 1.3. The wind-sea area is the area under the curves.

3.3.1.1 Discussion

In practice, equation 3.2 can be applied in two ways: either 1) considering the partitioned 2D spectrum in which case the phase velocity and direction are those of the peak (or mean) of the partition; or 2) without partitioning the spectrum in which case each spectral grid point is evaluated independently by equation 3.2.

The disadvantage of applying equation 3.2 to each grid point is that the frequency-direction area that fits the wind-sea-swell criterion (figure 3.5), does not necessarily corresponds with a wave system and the spectrum is split even in situations where it is composed of only one wave system. Not surprisingly, more consistent results in terms of wave systems are obtained if partitioning is used.

3.3.1.2 Example 3

The difference between applying equation 3.2 to the partitioned and non-partitioned spectrum are illustrated considering six hourly spectra from the ECMWF MARS archive (limited area deterministic system wave model WAM using the assimilation

system) at Westhinder in the southern North Sea (51.50°N, 2.50°E) from 26-Jan-2007 00:00:00 to 31-Jan-2007 18:00:00 UTC. Note that the ECMWF WAM model applies equation 3.2 to each frequency-direction bin for identifying wind-sea.

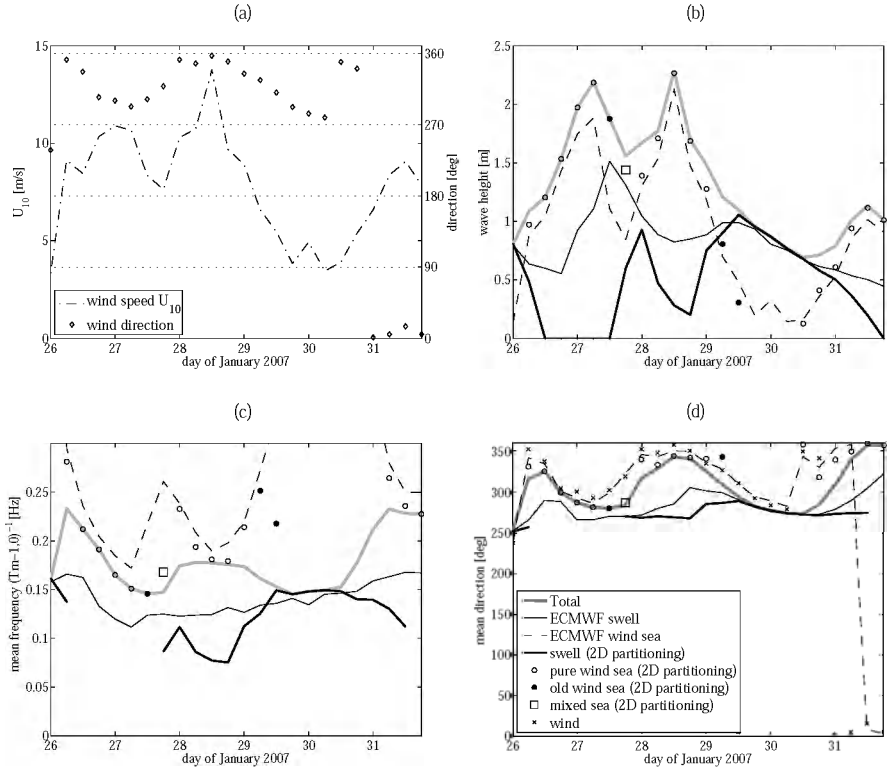

Figure 3.6. Wind and wave characteristics from the ECMWF MARS archive (WAM model) at Westhinder in the southern North Sea (51.50°N, 2.50°E) from 26-Jan-2007 00:00:00 to 31-Jan-2007 18:00:00 UTC. (a) Wind speed and direction, (b) significant wave height, (c) mean wave frequency ($f_{m-1,0}$), and (d) mean wave direction. In (b), (c), and (d), integral parameters of the entire spectrum (grey thick line), of the ECMWF swell estimates (thin black line), of the ECMWF wind-sea estimates (dashed line), of swell estimates using 2D partitioning (thin black dot-marked line), of pure wind-sea estimates using 2D partitioning (white circles), of old wind-sea estimates using 2D partitioning (black circles), and of mixed wind-sea estimates using 2D partitioning (squares).

The ECMWF give a continuous occurrence of swell (figure 3.6b) with a notorious peak at 27-Jan 12:00:00, while with the support of partitioning there are two successive swell events one at 28-Jan 12:00:00 and the second at 29-Jan at 12:00:00 originating from the two wind activity events (figure 3.6a). In any case, the ECMWF swell estimates are of larger magnitude. In addition, maxima of swell

energy are not occurring at the same time. For instance for the first swell event the ECMWF identification reaches its maximum 12 hours earlier than when working with partitions. Note the evolution of the decaying wind-sea system when partitioning is used. For example the wind-sea system present at Jan 27 12:00 evolves from a pure wind-sea into an old wind-sea, then into a mixed-sea (due to wind rotation from north-west to north) and finally into swell.

One can also follow the evolution of wave systems from looking at the time series of frequency and direction (figures 3.6, c and d). While the swell frequencies from the ECMWF estimates are quite continuous and appear rather constant, the evolution of the systems using also partitions looks more episodic. For instance in the swell event of 28-Jan starting at 18:00:00, low frequency swell waves arrive first. The swell mean frequency increases progressively. The energy in the swell systems from the two storm events decay faster (and even extinguish) than the ECMWF assigned swell energy (figure 3.6b).

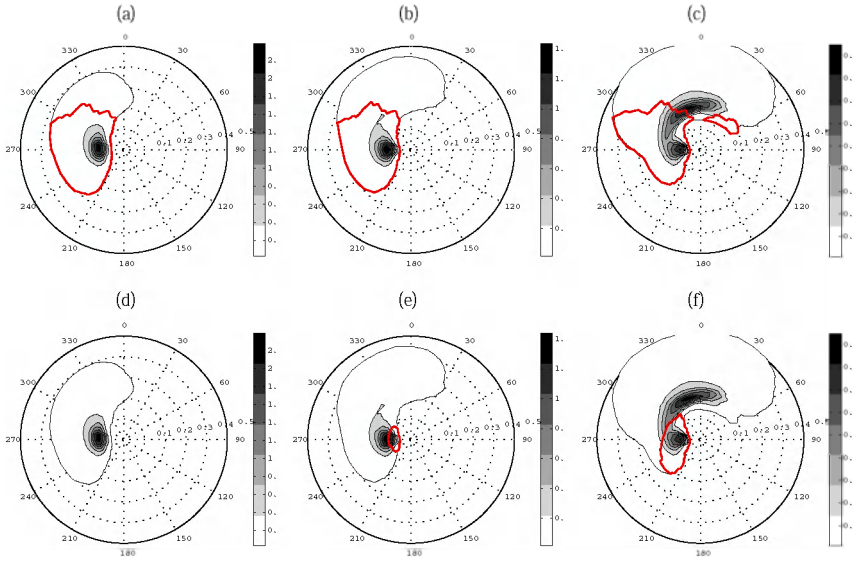

Figure 3.7. Wave spectra from ECMWF MARS archive (WAM model) at Westhinder (51.50°N, 2.50°E): (a) and (d) at 27-Jan-2007 12:00:00, (b) and (e) at 27-Jan-2007 18:00:00, and (c) and (f) at 28-Jan-2007 00:00:00 UTC. The swell part is contoured (thick continuous line) for ECMWF swell estimates (a, b, and c), and for swell estimate using 2D partitioning (d, e, and f).

In order to analyze these differences, wave spectra from the swell system of 28-Jan 12:00:00 are shown in figure 3.7. Figure 3.7a suggests that part of the spectrum is wind-sea and part is swell, while partitioning indicates that the spectrum is composed by a single (old wind-sea) system. There is little change in figure 3.7b with respect to 3.7a, while figure 3.7e shows that some energy of the previously old wind-sea system was transferred to lower frequencies forming a second (swell) system. The old wind-sea part in figure 3.7d has now, figure 3.7e, become a mixed wind-sea due to the rotating wind from 300° over 320° to 350°. In figure 3.7c the ECMWF swell part is composed of the low frequency wave system plus a significant part of the higher frequency system, while in figure 3.7f the more natural (morphological) decomposition into two wave systems is kept in a consistent manner.

3.3.2 Wind-sea – swell identification using 1D spectrum only

A simple method often used in practice to identify wind-sea and swell because of the sensitivity of ships to large period waves, is to set a constant splitting frequency or period (i.e., 10s). Although this method might be reliable in zones where wind-sea and swell occur markedly separated in the frequency domain, in many circumstances this method is not consistent since frequency as such does not determine whether a wave system can be considered wind-sea or swell.
Another common practice is to split the spectrum close to the peak frequency of the Pierson-Moskowitz (PM) spectrum (1964), equation 3.3.

$$f_{PM} = 0.13\frac{g}{U_{10}} \qquad [3.3]$$

A factor of 0.8 is commonly applied to indicate the splitting frequency (f_s=0.8f_{PM}) to account for uncertainties in the actual sea state or in the angular shift between wind and waves (Earl 1984, Quentin 2002).

Wang and Hwang (2001) use a splitting frequency f_s based on wave steepness. They define the wave mean steepness as (equation 3.4):

$$\alpha(f_*) = \frac{8\pi\left[\int_{f_*}^{f_{\max}} f^2 F(f)\,df\right]}{g\left[\int_{f_*}^{f_{\max}} F(f)\,df\right]^{1/2}} \qquad [3.4]$$

where $\alpha(f_*)$ is the steepness function at frequency f_*, $F(f)$ is the 1D spectrum, f is frequency, $f_{\max}$ is the upper-frequency limit of the spectrum, and g is the acceleration

due to gravity. Because of the f^2 in the formula the mean wave steepness is more related to the higher frequency waves and is less affected by lower frequency waves.

Wang and Hwang (2001) evaluated this steepness function for the PM spectrum at different wind speeds and found that the peak frequency of the steepness function can be related to the wind speed U through the regression equation $U=0.379 f_M^{-1.746}$. The separation frequency (f_s) was then set at the frequency where the wave phase speed equals the wind speed: $f_s = g/2\pi U$ (note that U is used without a reference level. In their plots, Wang and Hwang (2001) use U_5 but it is not explicit from the text whether U_5 or U_{10} are used for the calculations). In order to rescind of wind speed these (the regression and separation) equations were combined to obtain an expression for the separation frequency as function of the peak of the steepness function (equation 3.5):

$$f_s = 4.112\left(f_M\right)^{1.746} \qquad [3.5]$$

Where f_s is the separation frequency, and f_M is the peak of the steepness function.

Violante-Carvalho et al. (2002) proposed to fit a JONSWAP spectrum (Hasselmann et al. 1973), equation 3.6, to the high frequency spectral components to detect the peak that corresponds best to wind-sea. For more complex situations, when more than two peaks are present, they however extend this criterion by two other conditions. One condition looks at the wind and wave directional information and the other condition looks at the equilibrium range parameter α (Phillips 1958). Fitted α-values above 0.001 were considered wind-sea

$$F\left(f\right)=\alpha g^2 (2\pi)^{-4} f^{-5} e^{-\frac{5}{4}\left(f/f_p\right)^{-4}} \gamma^{e^{-\frac{(f-f_p)}{2\sigma^2 f_p^2}}} \qquad [3.6]$$

where $S(f)$ is the energy spectrum, f is frequency, α is the Phillips' constant, g is the acceleration of gravity, f_p is the peak frequency, γ is the peak enhancement factor, and σ is the spectral width factor, $\sigma = \sigma_a$ if $f < f_p$, and $\sigma = \sigma_b$ if $f > f_p$.

Although another 1D method to consider is the one of Voorrips et al. (1997), this method uses both wave directional information and wind information and it becomes a sort of 2D. Therefore it will not be used further here.

3.3.2.1 Discussion

It should be mentioned that methods like those of Wang and Hwang (2001) and the "non-extended version" of the method of Violante-Carvalho et al. (2002) have the

advantage of rescinding of wind data. Moreover, when dealing with 1D spectra wind data is of lower value since the wind and wave velocity vectors cannot be compared.

Criteria based on the PM peak might overestimate wind-sea, especially in growing wind-sea conditions where swell is also present. Consequently, the method of Wang and Hwang (2001) is implicitly affected by the two shortcomings mentioned above as it implicitly compares wind and wave velocities through a criterion obtained from the PM spectrum.

Gilhousen and Hervey (2001) indicate that the steepness method of Wang and Hwang (2001) overestimates wind-sea under certain conditions. They replaced equation 3.5 by $f_s = 0.75 f_M$ and introduced an extra mechanism similar to the one of equation 3.3 to complement the algorithm. This approach has not been considered further here because of the rather arbitrary decision to use the higher of the splitting frequencies calculated from the two criteria used.

Following the methodology of Violante-Carvalho et al. (2002) it was found that by fitting a JONSWAP spectrum to the higher frequency part of a wave system helps to identify the peaks that are correlated to that particular wave system. The first tests using this method showed a rather good agreement compared to the 2D scheme (section 3.3.1). Unfortunately, the fitting criterion by itself is not sufficient to decide what is wind-sea and what is swell. Therefore, a criterion related to the magnitude of the fitting parameter γ is introduced below.

3.3.2.2 Proposed 1D identification algorithm

In the JONSWAP formulation (equation 3.6) the peak enhancement factor γ says that the spectrum is sharper than the PM spectrum at the peak frequency which is considered to be an indication of active wave growth. The Phillips' 'constant' α was also found to depend on wind and wave conditions (Hasselmann et al. 1973). However, assuming that the energy at the peak frequency of a swell system cannot be higher than the value of a PM spectrum with the same peak frequency (i.e., α is fixed to its PM value, $\alpha_{PM} = 0.0081$), a simple 1D identification algorithm is set up as follows:

- The ratio (γ^*) between the peak energy of a wave system and the energy of a PM spectrum at the same peak frequency (equation 3.6 with $\gamma = 1, f = f_p$ and $\alpha = \alpha_{PM} = 0.0081$) is calculated.
- If γ^* is above a threshold value ($\gamma^* > 1.0$) the wave system is considered wind-sea, otherwise it is considered swell.

Note that in fact the spectrum no longer needs to be fitted. This criterion was tested here showing good agreement with the results of the 2D scheme. In the following sections, two rather different situations are considered in order to illustrate the operation of different identification methods.

3.3.2.3 Example 4

The data set of this example corresponds to buoy measurements from the Gulf of Tehuantepec, southern Mexican coast at the Pacific Ocean (16°N, 95°W), taken at about 30Km offshore (García 2006). The relevant feature there is a particular combination of meteorological and wave conditions. Due to a geographical depression in the mountain range that crosses the isthmus a particular wind system "Tehuanos" is formed. Tehuanos winds blow offshore generating fetch limited northerly wind-sea in a region where the wave climate is to a great extent characterized by open ocean southerly swells. Consequently, during Tehuanos wind events wind-sea and swell systems are very distinct in the wave spectra in both frequency and direction. The period considered here goes from 03-Mar-2005 23:22:00 to 05-Mar-2005 16:51:00 UTC (figure 3.8). The 1D wave energy spectra are shown in figure 3.9.

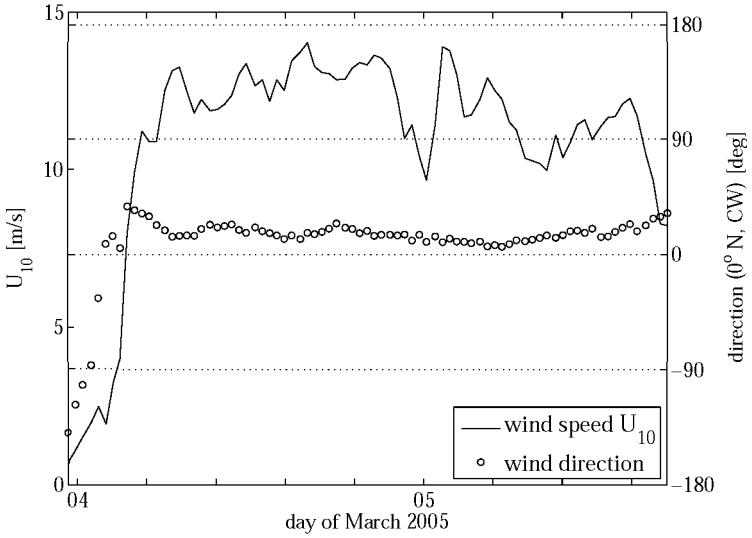

Figure 3.8. Wind conditions in the Gulf of Tehuantepec (16°N, 95°W) southern Mexican Pacific coast for the period from 03-Mar-2005 23:22:00 to 05-Mar-2005 16:51:00 UTC.

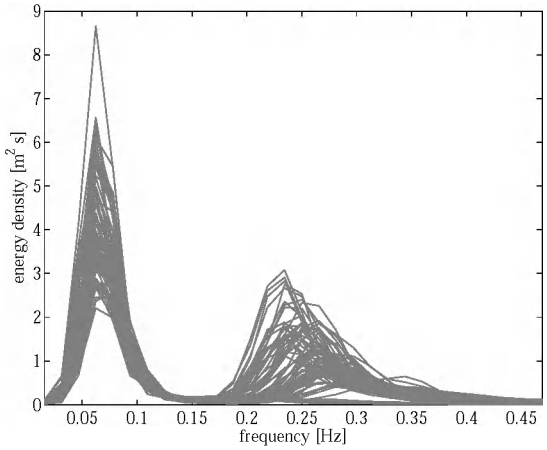

Figure 3.9. Wave spectra obtained at the Gulf of Tehuantepec (16°N, 95°W) for the period from 03-Mar-2005 23:22:00 to 05-Mar-2005 16:51:00 UTC. Spectra are given every 30 min (gray lines).

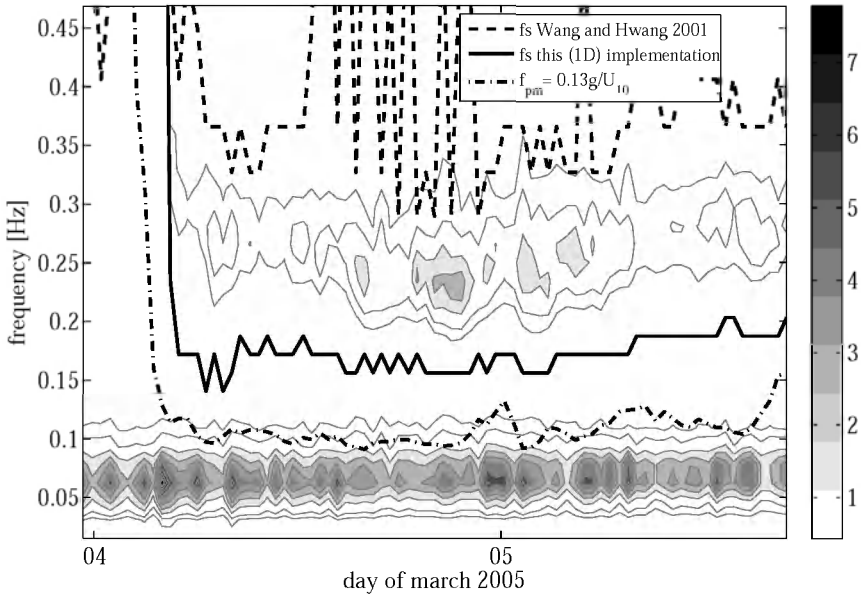

Figure 3.10. Time series of 1D energy density spectra (gray levels) obtained at the Gulf of Tehuantepec (16°N, 95°W) for the period from 03-Mar-2005 23:22:00 to 05-Mar-2005 16:51:00 UTC. And wind-sea – swell separation frequencies obtained using Wang and Hwang (2001) method (thick dashed line), the 1D method proposed in this study (thick continuous line), and using the PM peak frequency (thick dash-dot line).

From figure 3.9 it is clear that these spectra can be split conveniently (at a rather constant frequency) at the trough of the two systems (around 0.15Hz). Thus, the wave systems present in the spectra are known. Three wind-sea and swell identification methods have been applied to these spectra, namely Wang and Hwang (2001) method (equations 3.4 and 3.5), the PM peak (equation 3.3), and the method described here (section 3.3.2.2).

The separation frequencies from these methods were plotted on top of the time series of 1D spectra. By looking at the gray levels in figure 3.10, the wind-sea and swell systems can be clearly discerned. In these conditions, the separation frequency according to the steepness method of Wang and Hwang 2001 (dashed line) is systematically at higher frequencies than the splitting frequency (0.15 Hz) which consequently results in swell overestimation. The PM peak frequency (dash-dot line) is systematically at lower frequencies than the splitting frequency (0.15 Hz) but it seems rather consistent. Note that using factors lower than one would bring the separation frequencies to even lower values causing more overestimation of wind-sea. With the implementation given in this study (section 3.3.2.2) the separation of the systems (continuous line) is very consistent. Actually, partitioning of these spectra results in a very clean detection of the two main systems. Regarding the γ^* values of the wave systems, factors corresponding to the swell partitions are in any case lower than 1.0 while γ^* factors corresponding to wind-sea are above 3.

In order to analyze details the spectrum at 04-Mar-2005 21:00:00 UTC is shown in figure 3.11. There, the separation frequencies using these three methods are also indicated. The main features observed in the time series are also visible in the spectrum. According to the method of Wang and Hwang (2001) the wind-sea part only takes part of the tail of the actual wind-sea part (dashed line). The PM peak frequency corresponding to the present wind conditions (i.e., U_{10} = 13.5 m/s) is 0.1 Hz, taking part of the tail of the swell system as wind-sea (dash-dot line). With the 1D scheme outlined here (section 3.3.2.2), two peaks are detected by partitioning. The limit of the two partitions is indicated (filled diamond). The PM spectra corresponding to the two main wave systems are also indicated (dot line). In the case of the swell system the peak of the PM spectrum has larger magnitude than the observed swell system (γ^* = 0.0306) while the peak of the wind-sea has larger magnitude than a PM spectrum at that peak (γ^* = 14.3644). Note that using partitioning in combination with the PM peak frequency would yield results very similar to those obtained looking at the value of γ^*, but the associated disadvantage is the need of wind speed.

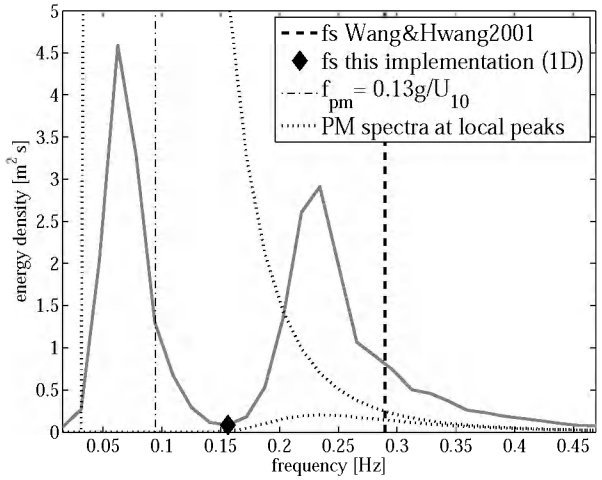

Figure 3.11. Wave spectrum from the Gulf of Tehuantepec (16°N, 95°W) at 04-Mar-2005 21:00:00 UTC and separation frequencies using Wang and Hwang (2001) method (thick dashed line), using the 1D wind-sea / swell identification method propose in this study (black diamond), using the frequency of the PM peak (thick dash-dot line). The PM spectra corresponding to the two main local peaks (dot lines) are also indicated.

3.3.2.4 Example 5

The present data set was measured by a directional Waverider buoy at Westhinder in the southern North Sea (51.38°N, 2.44°W), where wave conditions are characterized by the presence of local wind-sea and occasional swells coming from the north. Wind-sea and swell were present in the period from 12-Oct-1997 00:30:00 to 15-Oct-1997 18:30:00 UTC, which is a period of moderated winds in turning wind conditions (figure 3.12).

Contrary to the previous example, the wave systems in this case are not markedly separated. The spectra are rather complex and the splitting and identification procedures become more complex as well. Moreover, the true systems are not known. Therefore, in this case estimates of the true systems are obtained from the 2D scheme (section 3.3.1).

Time series of significant wave height of the whole spectrum (thick grey line) and of swell estimates of the 2D and of the three 1D identification methods used in the previous example are shown in figure 3.13.

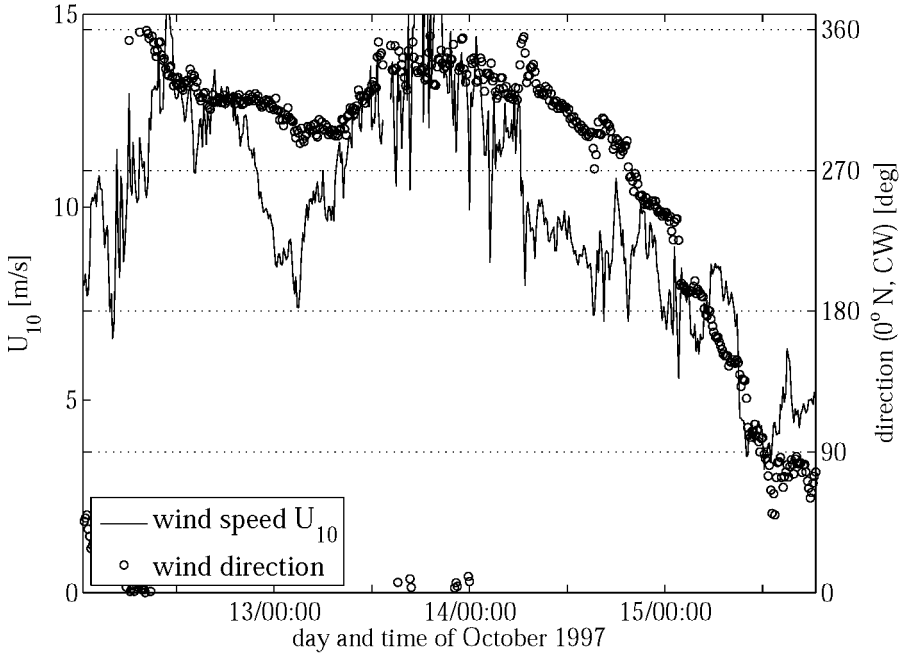

Figure 3.12. Wind conditions at Westhinder (51.38°N, 2.44°W), southern North Sea for the period from 12-Oct-1997 00:30:00 to 15-Oct-1997 18:30:00 UTC.

In figure 3.13 it can be seen that the results of the 1D method given here (section 3.3.2.2, continuous thick red line) are in good agreement with the results of the 2D scheme (section 3.3.1, continuous thin circle-marked line). The method of Wang and Hwang (2001), dashed line star-marked, provide similar estimates in cases where swell is dominant and there is little or no wind-sea (period around the 13-Oct-1997 00:00:00 and period after the 14-Oct-1997 12:00:00). In periods of wind-sea, this method systematically overestimates swell. There is rather good agreement using the PM peak frequency method (dash-dot diamond-marked line) especially in the swell dominated period after the 15-Oct-1997 00:00:00 but in general this method tends to underestimate swell in typical swell events.

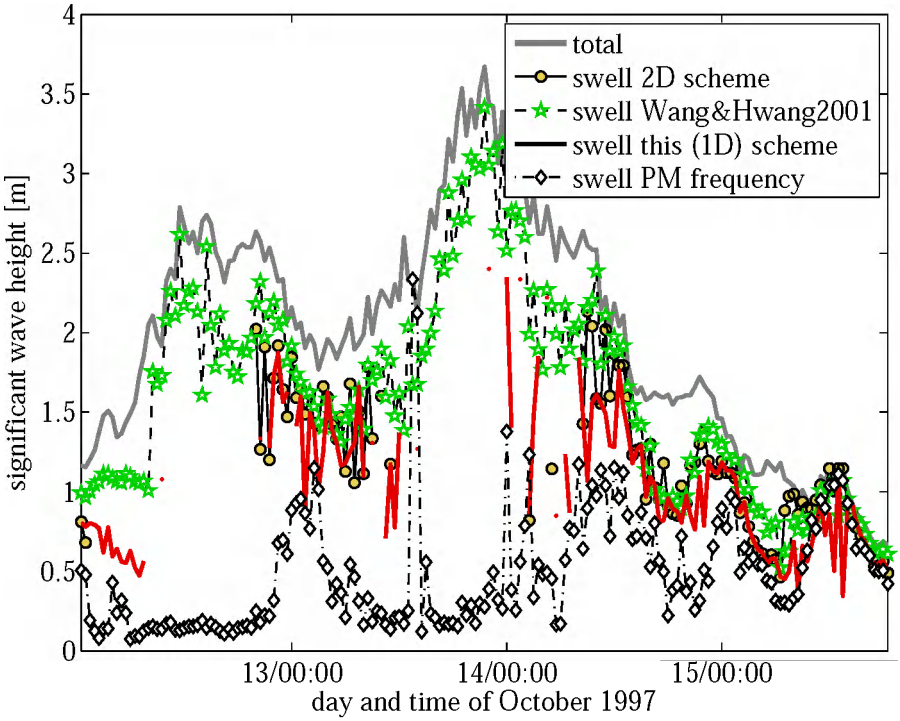

Figure 3.13. Time series of significant wave height at Westhinder (51.38°N, 2.44°W) for the period from 12-Oct-1997 00:30:00 to 15-Oct-1997 18:30:00 UTC and swell estimates: using Wang and Hwang 2001 method (dashed line star-marked), using the 1D wind-sea – swell identification method proposed in this study (thick continuous red line), using the PM peak frequency (dash-dot diamond-marked line), and using the 2D scheme outlined in this study (thin continuous circle-marked line).

In order to analyze the schemes in more detail the spectrum at 15-Oct-1997 00:00:00 UTC is shown in figure 3.14 from where it can be seen that the steepness method splits the spectrum somewhere at the tail of the wind-sea component underestimating wind-sea, while the PM peak frequency tends to be at lower frequencies than those obtained for the 2D scheme resulting in wind-sea overestimation. The 1D scheme presented here (section 3.3.2.2) splits the spectrum more consistently thanks to the use of the partitioning step. The values of γ^* for the two main peaks in the present example are 0.27 and 5.52 respectively. In general, the trends are similar to those from the previous example.

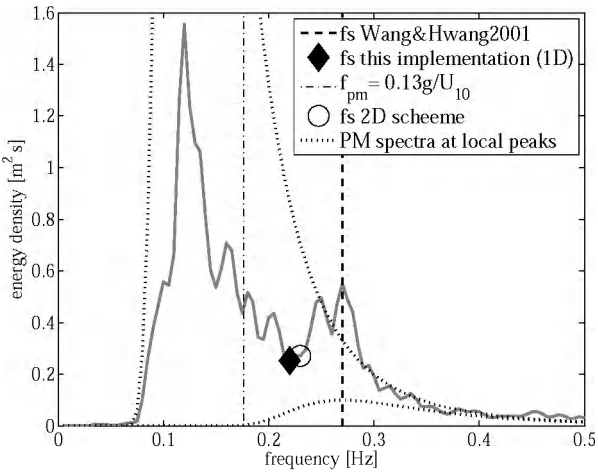

Figure 3.14. Wave spectrum at Westhinder (51.38°N, 2.44°W) at 15-Oct-1997 00:00:00 UTC and separation frequencies using Wang and Hwang (2001) method (thick dashed line), using the 1D wind-sea-swell identification method proposed in this study (black diamond), using the frequency of the PM peak (thick dash-dot line), and using the 2D scheme presented here (empty circle). The PM spectra corresponding to the two main local peaks (dotted curves) are also indicated.

3.4 Conclusions

Different spectral partitioning techniques have been investigated, emphasizing the fact that the different existing methods differ mainly in the way they assess whether partitions are significant or not, which implies basically the use of different combining strategies. It was found that the current mechanisms used for combining partitions reported in literature are not very robust. Moreover, they demand the use of arbitrary parameterizations. As a consequence, the existing spectral partitioning methods deliver rather inconsistent output for wave systems.

The introduction of an image processing tool based on a 2D low-pass filtering step aiming to reduce noise was found to improve the robustness of the 2D partitioning scheme considerably. The detection of wave systems is more consistent and the method is not very sensitive to parameter value settings.

Also a more robust partitioning scheme for 1D spectra has been proposed. The method aims of removing the most obvious spurious peaks. The criteria used for this purpose proved to be sufficient to reduce the number of partitions to a reasonable value.

Wind-sea and swell can be identified from looking at different environmental and physical characteristics of wave systems. However, results from different methods reported in literature sometimes differ largely.

For the identification of wind-sea and swell using the 2D spectrum plus wind speed and direction, a wind wave generation mechanism in combination with 2D partitioning uses all available information and gives the most consistent estimates.

Regarding the 1D wind-sea – swell identification methods, it is pointed out that the method of Wang and Hwang (2001) used at the NDBC center tends to overestimate swell especially during wind-sea periods. The PM peak frequency method is more consistent but underestimates swell systematically. Quite consistent results were achieved using 1D spectra only by looking at the ratio (γ^*) between the energy at the spectral peak of a partition and the energy at the peak of a PM-spectrum with the same peak frequency.

The identification of wind-sea and swell both in the 2D and 1D spectra is found more consistent in combination with partitioning.

Chapter 4

Overview of wave data assimilation schemes

4.1 Introduction

In this chapter, a brief overview of different assimilation methods available in the literature is given, focusing on applications to wave spectral modelling. This revision aims to give a framework for the assimilation approach followed in this study, which is an Optimal Interpolation approach. Three main data assimilation approaches are considered: Variational, Kalman Filtering, and Optimal Interpolation. These last two fall together under the same category of sequential methods. The basic difference between variational and sequential methods is the consideration of the time variable. In the variational approach the data assimilation process is defined for a time interval, using all information available in that interval. Moreover, the solution is obtained by correcting variables or parameters of the model input. In sequential approaches data assimilation is applied at the instant for which data becomes available (Evensen 2006), and corrections are applied to the model output variables.

In this chapter, some of the main advantages and disadvantages of these three approaches are discussed. The variational method is robust in the sense that the resulting analyzed state is fully consistent with the model dynamics. That is not totally the case of sequential approaches. They correct model output (wave fields) and have to use rather ad-hoc mechanisms to transfer the updated information to the wave spectrum (retrieval algorithm). From the statistical point of view, the advantage of the Kalman filter lays in its ability to compute the errors of the system (error covariance matrix), which is the key element of the assimilation system. In the Optimal Interpolation approach, the errors of the system are prescribed a-priori. This raw simplification brings up an important reduction in computational demands, which make the method suitable for practical implementations. Most implementations in wave data assimilation have followed one of these three approaches, although other types of schemes can be also found in the literature (e.g., Mandal et al. 2005, Coli et al. 2006, Theocharis et al. 2006).

The mathematical notation used to present the conceptual bases of each approach is the same as it is usually presented in the dedicated literature. The algorithm of the linear Kalman estimator is written in vector notation in order to illustrate how the filter is recursive and how the errors of the system are explicitly calculated. The variational approach is presented as an optimization problem as it is usually understood. The Optimal Interpolation method is written in matrix notation to give emphasis to the computation of the analyzed field and to illustrate the construction of the covariance matrices. However, it should be kept in mind that assimilation methods based on optimality are closely linked to one another, and could be presented with uniform notation (e.g., Wikle and Berliner 2007, Hollingsworth 1986, Daley 1991, Bouttier and Courtier 1999, Apte et al. 2007).

The basic task of any assimilation system is to introduce observations into the model with the aim of improving results. Observations are introduced at the corresponding model location (with a particular weight), and distributed smoothly in the spatial domain. In early data assimilation practices, both in meteorology and in oceanography, this spatial distribution was based mainly on the experience of the modeler. Some of the physical restrictions of the system were also taken into account. Often, exponential decay functions based on distance were used. This approach is nowadays referred as Cressman analysis (Cressman, 1959). Later on, the need was recognized to incorporate the statistics of the system in the assimilation procedure, not only with the aim of obtaining a more objective analysis, but also for automation. At present, the problem of assimilation is understood as a statistical problem and nearly any assimilation approach follows or tries to follow that principle.

In the data assimilation process, both model output (background) and observations are regarded as stochastic variables with their inherent accuracy. They both represent a physical variable whose true value is unknown. The combination of both pieces of information (regardless of their accuracy) produces a third one: the analysis, with enhanced accuracy compared to any of the two components. From the point of view of the probabilities, the joint probability distribution $p(x|x_b,x_o)$ is obtained from the conditional distribution of x, given the background distribution $p(x|x_b)$, and the observation distribution $p(x|x_o)$, Maybeck 1979, Wikle and Berliner 2007.

$$p(x \mid x_b, x_a) = p(x \mid x_b) p(x \mid x_o) \qquad [4.1]$$

This concept is illustrated for the case of a univariate system in figure 4.1. The prior distribution of the state variable x conditioned on the background value x_b, $p(x|x_b)$ is considered Gaussian with mean x_b and variance σ_b^2. The distribution of x

conditioned on the observed value x_o, $p(x|x_o)$ is also considered Gaussian with mean x_o and variance σ_o^2. Under the assumptions of linearity, whiteness, and Gaussianess, it can be shown that the best estimate of x conditioned on x_b and x_o, $p(x|x_b,x_o)$ is also Gaussian with mean x_a and variance σ_a^2 (Daley 1991, Maybeck 1979, Wikle and Berliner 2007).

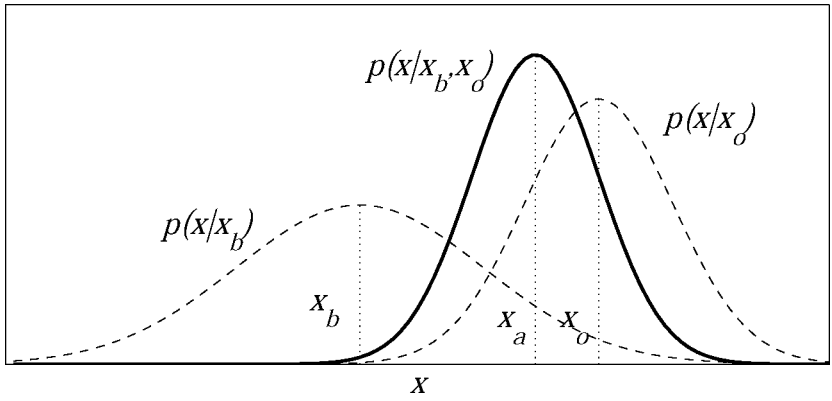

Figure 4.1. The data assimilation concept. The joint probability distribution $p(x|x_b,x_o)$ is obtained from the conditional distribution of x, given the background distribution $p(x|x_b)$, and the observation distribution $p(x|x_o)$.

$$x_a = \left[\frac{\sigma_o^2}{\sigma_b^2+\sigma_o^2}\right]x_b + \left[\frac{\sigma_b^2}{\sigma_b^2+\sigma_o^2}\right]x_o = x_b + K(x_o - x_b) \qquad [4.2]$$

$$\sigma_a^2 = \frac{\sigma_b^2\,\sigma_o^2}{\sigma_b^2+\sigma_o^2} = (1-K)\sigma_b^2 \qquad [4.3]$$

where K is the assimilation gain, $K = \sigma_b^2/(\sigma_o^2+\sigma_b^2)$.

The analysis x_a is the result of a linear combination of the background x_b, and the observations x_o, and the variance σ_a^2 is lower than both σ_b^2 and σ_o^2.

4.2 Variational Methods

4.2.1 Definition

Variational methods are closely related to the least square estimation theory where the assimilation problem is regarded as an optimization problem (Daley 1991). The differences between the model output and the observed parameters are the quantities to be minimized using the model equations as a dynamic constraint. The cost

function J is usually defined as a weighted sum of quadratic differences between model output and observations.

$$J(F,\psi)=\sum_{i,j} M_{i,j}^{-1}\left(x_i^o - x_j^m(F,\psi)\right)^2 \qquad [4.4]$$

where J is the cost or objective function, x represents the wave output parameters under consideration (e.g., H_{m0}, θ_w, f_p, $T_{m-1,0}$, α), F represents the set of spectral variables, ψ represents the control or tuning parameters. $M_{i,j}$ is the weighting factor representing the confidence interval of x (e.g., error covariance matrix between model m and observations o). The subscripts i and j indicate respectively observation points and grid points in the discrete space domain.

Optimal model results are then found by minimizing the cost function J with respect to the control variables ψ, by setting:

$$\nabla J\left(F,\psi\right)=0 \qquad [4.5]$$

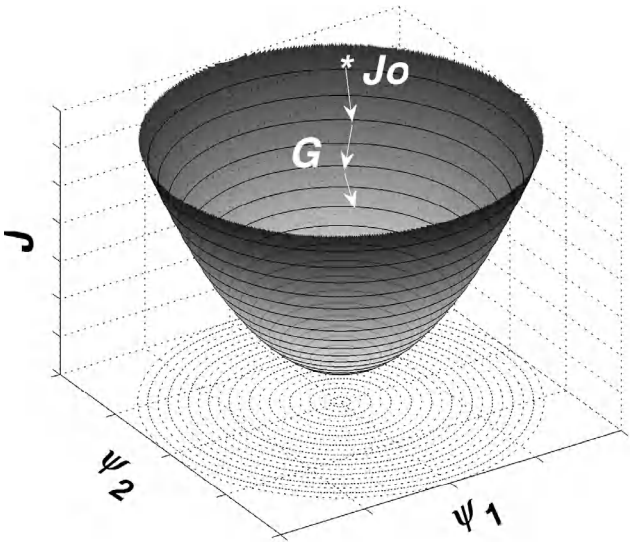

Figure 4.2. Graphical representation of the variational approach. The cost function J is represented against the control variables ψ_1 and ψ_2. The search for the minimum according to the steepest descent method is also illustrated. The search starts at Jo, then the optimization algorithm follows the steepest descent line (arrows) looking at gradient information of J.

Given that $\nabla^2 J(\psi) > 0$, to assure that the solution is indeed a minimum, and constrained by the model equations, this can be written in the form:

$$G(F, \psi) = 0 \qquad [4.6]$$

Note that when the temporal dimension is considered the method is referred to as four-dimensional variational analysis and it is defined for a time window in which the assimilation process takes place. Otherwise, the method reduces to a least square estimation problem and it is often referred to as three-dimensional variational analysis. In fact one of the major advantages of variational methods (i.e., four-dimensional variational methods) over sequential methods is that they are able to perform corrections in the time window in which the assimilation occurs, being able to carry out corrections in the wave model using information from past observations. Another advantage of this kind of approach is that the solution obtained from the minimization problem posing the model equations as a constraint is consistent with the model dynamics, which is not necessarily the case in sequential methods.

A major difficulty of this method is associated to the large computational cost involved in the minimization of the cost function. Especially for large-scale problems and for a large number of control variables several model runs may be necessary. Another significant difficulty of this method is the need to design a model for $M_{i,j}$ for all pair of model variables. The background and observations error covariance matrices must be realistic, because this determines the weight of the observations in the analysis (Bouttier and Courtier 1999). Besides, because the source and propagation functions in the wave model are non-linear functions, several local minima of the cost function may exist (Monbaliu 1992). In some cases, these local minima can be relatively small and show as noise of the cost function. However, the risk also exists of having more than one solution for the optimization problem in the time window defined. In those cases, the problem can be avoided by reducing the length of the assimilation time window (Ngodock et al. 2007).

4.2.2 Applications in wave data assimilation

There are several methods and algorithms for solving the optimization problem and extensive literature can be found on this topic in the area of optimization. However, the scope of this section is to give a short overview of previous implementations in wave data assimilation applications.

One way of minimizing the cost function is by an iterative procedure, applying a number of small perturbations to the control variables (i.e., linear perturbation analysis) in order to obtain gradient information of J to localize the minimum.

However, this approach is not very efficient and there are specialized algorithms (e.g., quasi-Newton type algorithms) to solve this in a more efficient way (Bouttier and Courtier 1999). Monbaliu (1992) has studied both a linear perturbation analysis and finite differences optimization algorithms using a one dimensional version of the WAM model (ONEMOD). He optimized the model parameters in order to obtain growth curves that would be in agreement with those Kahma and Calkoen (1991) obtained from observations. His results pointed out that an optimization with respect to all output variables was not possible since the minima of J for the different output variables were not located at the same value of the control variable. Besides, he indicated the problem that the cost function and its gradient were not smooth functions (Monbaliu 1992).

Similarly Holthuijsen et al. (1997) followed a variational approach to assimilate significant wave height into a second-generation wave model (DOLPHIN-B). They used wind input vectors as control variables. They performed a number of small (linear) perturbations to the control variables and approximated the computation of ∇J using first-order finite differences in order to obtain a more efficient scheme.

Another way of solving minimization problems in which the cost function is constrained by other set of equations is by means of the Lagrange multiplier method. The minimization problem (eq. 4.5) constrained by the model equations (eq. 4.6) is replaced by an unconstrained problem of the form:

$$L(F,\psi)=J(F,\psi)+\lambda^T G(F,\psi) \qquad [4.7]$$

where L is the unconstrained cost function, and λ represents the unknown Lagrange multipliers. If λ is chosen to satisfy the adjoint equation:

$$\frac{\partial J}{\partial F}-\lambda^T\frac{\partial G}{\partial F}=0 \qquad \Rightarrow \qquad \left(\frac{\partial G}{\partial F}\right)^T\lambda=\left(\frac{\partial J}{\partial F}\right)^T \qquad [4.8]$$

it can be shown that (Giles and Pierce 2000):

$$\frac{dL}{d\psi}=\frac{\partial J}{\partial \psi}-\lambda^T\frac{\partial G}{\partial \psi} \qquad [4.9]$$

This means obtaining the solution of an unconstrained minimization problem. This approach is also referred to as the adjoint technique and its main advantage with respect to direct methods lays in the fact that the computational load may be significantly reduced, especially when the number of control variables is large. Actually, two model integrations (for each iteration of the descent algorithm) are needed independently of the number of control variables, one forward in time to

calculate the partial derivatives, and one backwards in time to compute the multipliers (Barzel and Long 1994, Vos 2002).

This type of approach has been studied previously in wave data assimilation exercises. Unfortunately, the construction of an adjoint model is not a straightforward task. In all practical applications, simplifications must be carried out. De Valk and Calkoen (1989) and de Valk (1994) have done an implementation of the adjoint technique. They used a second-generation wave model in order to cope with computational costs and other practical limitations. A comparison of their implementation with an Optimal Interpolation scheme (using the WAM model) does not show the superiority of the variational method (Voorrips and de Valk 1997). However, the simplifications done to the variational scheme must be considered in order to make an objective judgment.

Barzel and Long (1994) have performed a model optimization exercise similar to that of Monbaliu (1992), using an implementation of the adjoint method. De las Heras et al. (1994) and De las Heras et al. (1995) have worked out an analytical adjoint version of the WAM model. However, their application is limited to a single grid point. In addition, the variations of the control variables are restricted to a limited area.

Bauer et al. (1996) implemented an assimilation scheme for the WAM model following an adjoint related approach. They approximated the response function of the wave model to the wind input (Green's function) by a Dirac delta function. This approximation implies physically, the assumption that the input perturbations that cause a response of a specific spectral component are strongly localized in time and space.

Furthermore, the idea of the adjoint method has been extended to generate from the computer code itself a code for the adjoint problem. Since the analytical inversion of the model equations is a very challenging task, this approach regards the computer code as the set of model equations. The adjoint code is then generated by inverting the computer code line by line. Hersbach (1997) has built an adjoint version of the WAM model (The ADWAM model) using an automatic adjoint generator (The AMC compiler, Giering and Kaminski 1998).

This adjoint model has been used by Vos (2002). He carried out data assimilation experiments with emphasis in wind field corrections. He compared results from the ADWAM model and a variational data assimilation approach based on finite differences optimization using the second-generation model (DOLPHIN-B). His results show a rather low impact of the assimilation in the two cases.

4.3 Kalman Filtering

4.3.1 Definition

The Kalman filter (Kalman 1960) is a recursive optimal estimator of the state of a linear dynamic system discretized in the time domain using measurements that are linearly related to that state. Given that both background and measurements are perturbed by Gaussian random processes, the resulting estimator is statistically optimal with respect to virtually any quadratic function of estimation error. (Grewal and Andrews 2001, Maybeck 1979, Gelb et al. 1994).

4.3.2 The linear Kalman filter algorithm

The time-dependent state and observations can be represented by the following linear stochastic models (Barrero Mendoza 2005):

$$x_{k+1} = A_k x_k + B_k u_k + G_k w_k \qquad [4.10]$$

$$y_k = C_k x_k + v_k \qquad [4.11]$$

where k indicates the time index, x is the state variable vector, u is the input vector of the system, y is the observations vector, and w_k and v_k are respectively the state and measurements stochastic noise. A_k, B_k, G_k, and C_k are matrices of orders nxn, nxp, nxl, and mxn respectively, with n the order of the system, p the number of inputs, l the number of uncertain parameters, and m the number of observations.

Equation 4.10 says that the model state variable x at the discrete time step $k+1$ is a linear combination of the state variable at the previous time step k and the input variable u, plus some random noise given by w. Equation 4.11, express that a linear relationship exist between the observed and modelled state variable (plus a noise component v).

From the definition of covariance, the model error covariance matrix Q_k, and the observations error covariance matrix R_k can be represented by:

$$Q_k = \langle w_j w_k^T \rangle \qquad [4.12]$$

$$R_k = \langle v_j v_k^T \rangle \qquad [4.13]$$

where the angle brackets indicates the expectation value operator. The orders of matrices Q_k and R_k are nxn and mxm respectively.

Besides, it is assumed that model and observation errors are uncorrelated. That is, their covariance is zero $\langle v_j w_k \rangle = 0$.

In order to obtain a recursive filter, errors at subsequent time steps have to be calculated. This is done in the Kalman filter by propagating the errors in time (their probability density functions) using the model dynamics (equation 4.10) and a linear state estimator of the form:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \left(y_k - \hat{y}_{k|k-1} \right) \qquad [4.14]$$

where K_k is the gain matrix having order *n*x*m*, and $\hat{y}_{k|k-1} = C_k \hat{x}_{k|k-1}$

In equation 4.14, the estimator of the state variable is found as a linear combination of the prior state estimator ($\hat{x}_{k|k-1}$ and $\hat{y}_{k|k-1}$), and the observations y_k.

The first step is to update the state $x_{k-1|k-1}$ in time using equation 4.10.

$$\hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1} \qquad [4.15]$$

Then the prior error covariance matrix $P_{k|k-1}$ is evaluated. The prior state estimator error is defined as:

$$e_{k|k-1} = x_k - \hat{x}_{k|k-1}, \; k > 0 \qquad [4.16]$$

and the prior error covariance matrix is defined by:

$$Q_{k|k-1} = \left\langle e_{k|k-1} e_{k|k-1}^T \right\rangle \qquad [4.17]$$

Substituting 4.15 and 4.10 into 4.16, a model for the error is obtained based on the dynamics of the system.

$$e_{k|k-1} = A_{k-1} e_{k-1|k-1} + G_{k-1} w_{k-1} \qquad [4.18]$$

Equation 4.18 says that the prior state estimator error is calculated as a sum of two noise components (see for instance figure 4.1).

Accordingly, the prior error covariance matrix can be calculated as:

$$Q_{k|k-1} = A_{k-1} P_{k-1|k-1} A_{k-1}^T + G_{k-1} Q_{k-1} G_{k-1}^T \qquad [4.19]$$

The next important step is to obtain an expression for the Kalman gain K_k. This is achieved by minimizing the mean square error. The function J_k is introduced in the form of a cost function. J_k can be conveniently expressed using the *trace* operator, which is defined as the sum of the elements on the main diagonal:

$$J_k(K_k) = trace(Q_{k|k}) \quad [4.20]$$

The Kalman gain is obtained by setting $\partial J_k(K_k)/\partial K_k = 0$

In order to obtain a solution for equation 4.20, the state estimator error and the estimator error covariance matrix are defined similarly as in equations 4.16 and 4.17 by:

$$e_{k|k} = x_k - \hat{x}_{k|k} \quad [4.21]$$

$$Q_{k|k} = \left\langle \left(e_{k|k} - \langle e_{k|k} \rangle \right) \left(e_{k|k} - \langle e_{k|k} \rangle \right)^T \right\rangle \quad [4.22]$$

The state estimator error ($e_{k|k}$) can be then expressed in terms of the prior state estimator error ($e_{k|k-1}$) by substituting 4.14 into 4.21

$$e_{k|k} = (I_n - K_k C_k) e_{k|k-1} - K_k v_k \quad [4.23]$$

And the error covariance matrix for the state estimator ($Q_{k|k}$) can be expressed in terms of the error covariance matrix of the prior state estimator ($Q_{k|k-1}$) by replacing 4.23 into 4.22

$$Q_{k|k} = (I_n - K_k C_k) Q_{k|k-1} (I_n - K_k C_k)^T + K_k \hat{R}_k K_k^T \quad [4.24]$$

The matrix $\hat{R}_k$ is introduced to have a more compact notation, but it is also by definition expressed as:

$$\hat{R}_k = R_k + C_k Q_{k|k-1} C_k^T \quad [4.25]$$

By setting $\partial J_k(K_k)/\partial K_k = 0$, the Kalman gain is found to be:

$$K_k = Q_{k|k-1} C_k^T \hat{R}_k^{-1} \quad [4.26]$$

Note that the Kalman gain given in equation 4.26 in vector notation is equivalent to the assimilation gain matrix K expressed in scalar form in equation 4.3.

The state estimator error covariance matrix is then calculated from the prior state estimator error covariance matrix and the Kalman gain by substituting 4.26 into 4.25.

$$Q_{k|k} = (I_n - K_k C_k) Q_{k|k-1} \quad [4.27]$$

Finally, the model state variable is updated via 4.14.

4.3.3 Implementations in wave data assimilation

There is growing interest in the use of the Kalman filter for data assimilation as this method has potentially some advantages over other methods. One of the main advantages is that the error statistics of the system are calculated explicitly in a recursive way. However, as the Kalman filter is only defined for linear systems, and also because of its computational complexity, it is not directly applicable to large scales non-linear systems, which are actually the most interesting and common applications (Barrero Mendoza 2005, Welch and Bishop 2006).

Nevertheless, several versions of the Kalman filter have been developed to overcome some of its practical problems. The Extended Kalman Filter EKF (Jazwinski 1970) is an extension of the linear filter to cope with non-linear models. The Ensemble Kalman Filter EnKF (Evensen 1994, Evensen 2003) has been developed to overcome both non-linearities and high computational requirements. Several other versions of the Kalman filter can be found in the literature (i.e., Hoang et al. 1997, Madsen and Cañizares 1999, Hoteit et al. 2005, Apte et al. 2007, Julier and Uhlmann 1997). The different versions aim to tackle the referred or other associated problems

Voorrips et al. (1999) carried out wave data assimilation experiments using the EKF. They investigated the standard EKF, a truncated second-order filter, and a fixed-lag Kalman smoother to account for wind correction in past winds. However, the method has been restricted to small-scale tests and moderated departure errors.

Siddons (2007) has implemented an EnKF for the assimilation of wave data from High Frequency Radar (HFR). In his work, Siddons assessed the performance of the EnKF compared to a 3D variational scheme and to an Ensemble Optimal Interpolation method. His results show that under ideal conditions, where the truth and the errors uncertainties are known, the EnKF is superior to the other methods. However, in more realistic scenarios the performance of the EnKF was lower than the other methods. Several factors are thought to have affected the performance of the EnKF. The main was probably related to a poor quantification of the model and initial uncertainties. Also the reduced number of ensembles (16 members) used for the experiments is expected to have an important influence. In addition, the presence of bias is thought to play a negative role. Moreover, it is foreseen that some of the algorithm simplifications have rendered the filter suboptimal (Siddons 2007).

4.4 Optimal Interpolation

4.4.1 Definition

The process of incorporating spatially sparse observations into a background field by linear interpolation, weighted according to the error characteristics of the involved variables, is often referred to as Optimal Interpolation (OI) or statistical interpolation. In fact, as such incorporation can be achieved by means of any arbitrary interpolation function (i.e., subjective analysis), the need for objectivity forces one to take the statistics of the system into account. The aim of OI is to produce an analyzed field that is the best combination (optimum) between the observations and the background field, given the uncertainties of each of these components (Bouttier and Courtier 1999).

4.4.2 Mathematical representation

In the OI method, the analyzed wave field is expressed as a linear combination of the background (first guess) field and the observations:

$$x_a = x_b + \mathbf{K}\left(y - \mathbf{H}\,x_b\right) \qquad [4.28]$$

where x represents the model state variable vector, the superscripts a and b, denote analysis and background respectively, y is the vector of observations, $\mathbf{H}$ is called the observation operator and projects the model state to the observation locations, the weight matrix $\mathbf{K}$ is calculated as:

$$\mathbf{K} = \mathbf{Q}\mathbf{H}^T\left[\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right]^{-1} \qquad [4.29]$$

where $\mathbf{Q}$ and $\mathbf{R}$ are the error covariance matrices of the model and observations respectively, by definition:

$$\mathbf{Q} = \langle w_i w_j \rangle \qquad [4.30]$$

$$\mathbf{R} = \langle v_k v_l \rangle \qquad [4.31]$$

the angle brackets indicate the expected value operator, w and v represent the model and observations errors respectively, the subscripts i, j correspond to model grid indexes, and the subscripts k, l correspond to observation locations.

This definition implies several assumptions, which have to be taken into account for implementation. Both, the computation of the analyzed field and the observation operator are linear, this assumption not always holds in the case of wave models, but

is a suitable assumption when the differences between model and observations are not too large. Besides, background and observations errors are assumed to be uncorrelated. This is in general an acceptable assumption, but might be also violated in some cases (e.g., measurements taken by the same platform). In addition, bias should be removed in order to assure optimality of the solution (Daley 1991, Bouttier and Courtier 1999).

The OI method has some particular characteristics that differentiate it from other assimilation methods. These characteristics constitute to a considerable degree the core of all its advantages and disadvantages. One of them (as in most sequential methods) is the fact that model corrections occur to the model output. This means to a certain extent addressing the effect of the problem and not its cause, as it is contrarily the purpose of variational methods. An associated shortcoming is that due to the manipulation of model output it is not possible to guarantee that the model dynamics are respected. Consequently, the external introduction of data may be rejected by the model if it happens to be unrealistic or inconsistent (Hollingsworth 1986). Another associated disadvantage compared to variational methods is that an additional mechanism is necessary to retrieve the corrected output (e.g., wave height, wave period) into the model actual variables (e.g., the wave spectrum). Contrarily, the fact that corrections are made to the model output, and not to each of the model variables, brings in an important advantage in terms of computational load reduction.

Another important characteristic of this method, and probably the most fundamental one, lays in the computation of the gain matrix **K** (equation 4.29). The error covariance matrixes **Q** and **R** are pre-specified based on a-priori knowledge of the errors of the system. This simplification is acceptable when the magnitudes of the observations and background errors are somehow known. However, that is often not the case and in general, this pre-specification may constitute a major feebleness of the method. It should be noted that the interpolation weights, and therefore the quality of the analysis, rely on the quality of the error covariance matrices (Dee 1995). On the contrary, again, from the computational point of view, this pre-specification means an enormous load reduction compared for instance with Kalman Filtering techniques. Indeed the cost of the assimilation step, in the OI scheme, is almost negligible compared to the cost of the model integration itself. For this reason, and given that under an appropriate specification of the error covariance matrixes the OI performs very good, this method has gained significant popularity in assimilation applications where the computational time is a crucial aspect (e.g., operational meteorology and oceanography).

4.4.3 Specification of the error covariance

It has been mentioned already that the specification of the system error covariance is a crucial aspect of the OI technique. Especially the specification of the background error covariance **Q** is an important one, since the computation of the gain matrix **K** depends largely on the structure of this matrix. Besides, background errors are the most difficult to specify as they are expected to be spatially correlated (Bouttier and Courtier 1999). On the other hand, for what it concerns to most applications of the OI in oceanography the observation errors can be assumed to be uncorrelated. Under such an assumption the observations error covariance matrix **R** becomes a diagonal matrix with the elements in the diagonal being the variances of the observations (i.e., $\mathbf{R} = \sigma_o^2 \mathbf{I}$).

There are several ways to estimate the background statistics of a system in practice but in most cases approximations are mandatory because it is not possible to dispose of sufficient data to perform empirical statistics of a system. A systematic approach suited for dense observing networks has been given by Hollingsworth and Lonnberg (1986). In their method, the so called observational method, background departures are used to infer background errors, under the assumption of uncorrelated background and observations errors. A method to estimate background errors in forecasting systems with data assimilation has been proposed by Parrish and Derber (1992). In their method, short-range forecast errors are used to infer background errors (i.e., NMC method). Another approach proposed by Houtekamer et al. (1996) consists of introducing a series of perturbations in the physical parameterizations of the model (i.e., ensemble or Monte Carlo methods). A more straightforward way to approximate errors is to take the variance as a fraction of the state variable itself (i.e., $\sigma^2 \propto x$). In any case, the background matrix covariance can be calculated by fitting a correlation model to the background errors (equation 4.32). From the definition of the correlation coefficient we have:

$$\rho_{ij} = \frac{\langle w_i \, w_j \rangle}{\sigma(w_i)\sigma(w_j)} \qquad [4.32]$$

where ρ is the error correlation coefficient, the numerator in the right hand side of the equation represents the covariance (e.g., **Q**), w represents the background errors, σ represents the error standard deviations, and the subscripts i and j represent grid points.

In general conditions of homogeneity and isotropy are assumed for the error background structure (Hollingsworth 1986, Bouttier and Courtier 1999, Daley 1991). Previous OI implementations in wave data assimilation have used different

correlation models, considering in general that correlation decays with distance. Often, exponential decay curves according to equation 4.33 have been used:

$$\rho_{i,j} = \exp\left[-\left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{L}\right)^{\lambda}\right] \qquad [4.33]$$

where $|\mathbf{r}_i - \mathbf{r}_j|$ represents the relative distance between the grid points i and j (i.e., $[(r_{i(x)}-r_{j(x)})^2+(r_{i(y)}-r_{j(y)})^2]^{1/2}$), L is the correlation length which controls the width of the exponential bell, and λ is a fitting parameter.

Greenslade and Young (2004) give a good overview of several correlation models previously used in operational oceanography. In their study, they systematically applied an observational method (Hollingsworth and Lonnberg 1986) to study the correlation of wave height errors between the wave model WAM and ERS-2 observations. They emphasized the fact that little is known in practice about the correlation structure of errors in wave data assimilation systems. Additionally, it is noted that most of the correlation models used previously were rather experimental. Some of these correlation models are summarized in figure 4.3.

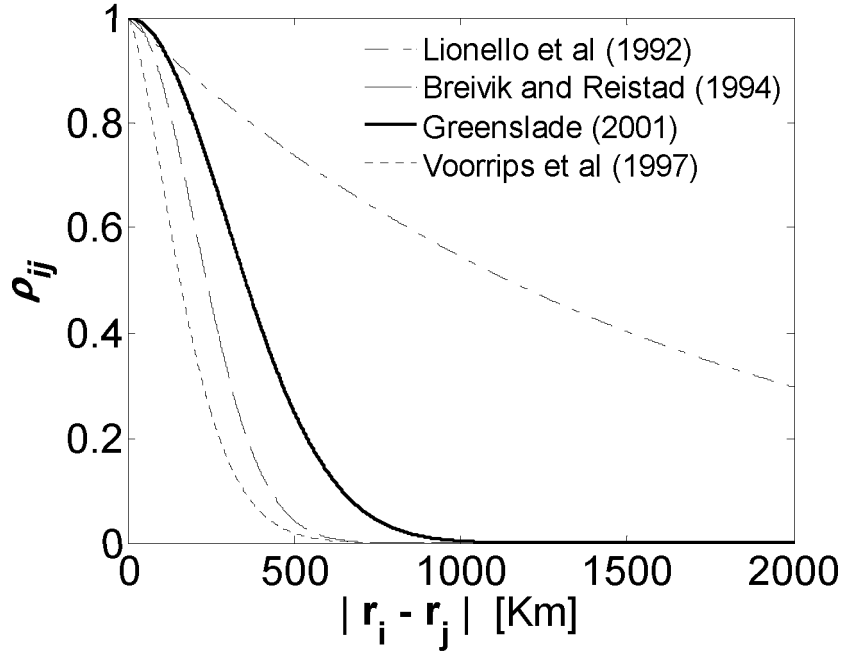

Figure 4.3. Examples of background error correlation functions (after Greenslade and Young 2004).

4.4.4 Applications in wave data assimilation

It has been referred already that some practical characteristics of the OI method like its relative simplicity and low computational demands make it very attractive for practical data assimilation use. In addition, several data assimilation experiments using this method have shown that under an acceptable specification of the system errors the performance of this method is satisfactory even in operational conditions (e.g., Lionello et al. 1992, Greenslade 2001). In view of these advantages, the OI method is probably one of the most implemented methods in practical wave data assimilation problems.

One of the first thorough implementations of the OI method in wave data assimilation systems was presented by Lionello et al. 1992. They assimilated, on a global scale, significant wave height from the Seasat altimeter into the WAM model. In their OI scheme, they used an error correlation model similar to the one given in equations 4.32 and 4.33, with $\lambda = 1$. They calibrated the system for different values of both L, and the ratio between observations and background uncertainties (σ_o/σ_b). They found that typically the forecast improved as L was increased. Saturation was apparent above an average correlation length L of 5 grid points (i.e., 15° or ~1650 Km). For the ratio σ_o/σ_b, they obtained satisfactory results for values of 0 and 1, suggesting that the accuracy of the altimeter was higher or comparable to that of the model.

For the retrieval of analyzed output into the wave spectrum, it is recognized that some physical considerations of the dynamical system may help (Hollingsworth 1986). In this regard, Lionello et al. (1992) considered two different wave systems, wind-sea and swell. As these two systems are not anymore computed separately in third generation wave models, they considered roughly wind-sea as all waves for which the total wave energy was in increment, and swell as waves in periods of decreasing energy. They did not consider appropriate to split the spectrum and update each component separately, as it was formerly proposed by Thomas (1988), given that the limited information provided by the altimeter (i.e., wave height only) did not justify restructuring the whole spectrum. Therefore, they assumed the whole wave spectrum to be either wind-sea or swell. Accordingly, they proposed to update the wave energy spectrum, given the analyzed wave height field, by a straightforward energy rescaling factor α, and a frequency shift factor β, with α and β constant for all frequency bands.

$$F(f,\theta)_a = \alpha F(\beta f,\theta)_b \qquad [4.34]$$

The rescaling and shift factors, α and β are calculated considering the sea state (wind-sea or swell), that is:

wind-sea: $$\alpha = \frac{E_a}{E_b}\beta, \qquad \beta = \frac{f_{mb}}{f_{ma}} \qquad [4.35]$$

swell: $$\alpha = \frac{E_a}{E_b}\beta, \qquad \beta = \Delta\left(\frac{E_a}{E_b}\right)^{1/4} \qquad [4.36]$$

where $F(f, \theta)$ is the wave energy spectrum in frequency f and direction θ. E is the total mean energy of the wave spectrum, f_m is the mean frequency, the subscripts a and b indicate analysis and background respectively, and the factor Δ is a coefficient to account for some eventual variations of the steepness and it is expected to be close to 1.

Actually, by the assimilation of wave height only, the shift of the spectrum in frequency β, is also not supported by the data. In fact, Lionello et al. (1992) had to make some assumptions in order to estimate it. For the case of wind-sea, their hypothesis was that mispredictions of the wave model were mainly caused by mispredictions in the magnitude of the wind field. Contrarily, they assumed that the duration of the storm events was well represented by the atmospheric model. This hypothesis was supported by a general believe of wave modellers and the previous experiments performed by Thomas (1988). Additionally, they obtained from model simulations a time limited growth curve and an energy-frequency relationship. From those relationships, they could estimate an analyzed mean wave frequency field in function of the mean energy. For the case of swell, they observed that the same energy rescaling strategy did not produce satisfactory results. They suggested that the update of swell should be done assuming that the average steepness of the swell system should not change. This approach was supported based on results of previous experiments (Lionello and Janssen 1990).

Lionello et al. (1992) obtained satisfactory results with this scheme in operational conditions. Further studies and implementations adopted similar approaches e.g., Breivik and Reistad 1994, Mastenbroek et al. 1994, Bidlot et al. 1995, Lionello et al. 1995, Ovidio et al. 1996, Young and Glowacki 1996, Hasselmann et al. 1997, Voorrips et al. 1997, Breivik et al. 1998, Dunlap et al. 1998, Greenslade 2001, Sannasiraj et al. 2006.

Previous assimilation studies using sequential schemes had already addressed some of the main issues of this type of schemes. Thomas (1988) for instance, focused on the retrieval of wave height data into the spectrum. Although he did not actually assimilate data into the model, he foresaw the need of distinguishing wave systems in order to perform a more robust analysis of the wave spectrum. This was motivated by two facts. First, he used a second-generation wave model, this type of

models explicitly calculate wind-sea and swell. In addition, he disposed of spectral wave data for verification. He proposed estimating the wind-sea energy from the total observed energy using wind speed observations and parametrical relationships of wave growth. For this purpose, he essayed two different strategies to estimate the wind-sea component. Namely, the conservation of the wave age, and the conservation of the stage of development. In the two cases, the swell part was taken as the remaining from the total energy. The redistribution of wind-sea energy was done according to the model dynamics (i.e., JONSWAP shape), using the analyzed energy and observed wind speed. For the case of swell, a rescaling factor was applied.

Other former studies using sequential schemes, addressed also the problem of spreading the analysis to neighbouring grid points. Esteva (1988) did some experiments preserving the model field slopes as interpolation function, but results were not better than those with single grid point insertion. Janssen et al. (1989) and Francis and Stratton (1990) did already attempt to include the system errors to determine the interpolation weights and to evaluate the correlation lengths of the system. Bauer et al. (1992) used a decay function inversely proportional to the distance.

Although most of the early wave data assimilation studies were motivated by the potential abundance of satellite altimeter data. Later on, the disposal of 2D spectra from SAR (Synthetic Aperture Radar) images encouraged the development of more sophisticated algorithms (Hasselmann et al. 1996). However, some of the inherent limitations of the SAR images have hampered their impact in data assimilation. Firstly, they are highly distorted images of the ocean wave field (Janssen 2008) and therefore their spectra are expected to contain large uncertainties. Besides, the SAR spectra suffer from a 180º directional ambiguity. Additionally, if different wave systems are to be accounted for, observed or modelled wind information is necessary (e.g., scatterometer, altimeter). All this plays a role in detriment of the accuracy. Finally, the retrieval and partitioning algorithms might also be subject to deficiencies, causing a reduction of the final accuracy of the observed spectral data.

In their assimilation scheme, Hasselmann et al. (1996) proposed making use of a spectral partitioning technique (Gerling 1992) to determine wave systems from both the SAR and the WAM model spectra. Cross-assign corresponding wave systems and assimilate wind-sea and swell accordingly following the OI method. But given the limitations of the SAR images, the final accuracy of the data are too low to show a significant impact on the assimilation. In contrast, it is recognized that the highest quality data come from in-situ measurements. Besides in-situ measuring networks are significantly dense in some regions as the North Sea being thus also attractive

for data assimilation. In this regard, Voorrips et al. (1997), following the scheme of Hasselmann et al. (1996) extended the partitioning scheme in order to treat and assimilate 1D spectral wave components from buoy measurements.

Regarding the estimation of errors, many implementations have been achieved by tuning the parameters of the covariance function (equations 4.32 and 4.33). That is, the error correlation length and the ratio of observed to model standard deviations. More dedicated approaches have been more dedicated to estimate the error covariance function. Voorrips et al. (1997) estimated the background errors of the forecasting system according to the NMC approach (Parrish and Derber 1992). Greenslade and Young (2004) studied the correlation structure of the system following the observational method (Hollingsworth and Lonnberg 1986). Sannasiraj et al. (2006) followed an ensemble method, the so called local linear model from Babovic et al. (2005), to generate an ensemble covariance from a series of model runs with perturbed wind input.

4.5 Summary and Conclusions

Wave data assimilation attempts started just after the advent of global wave forecasting systems in the decade of 1980 in view of the availability of large amounts of information available from buoy measurements but especially from satellite measurements.

One of the early motivations to assimilate wave data was not the improvement of the wave forecasting system as such but rather the possibility to give feedback to the forcing meteorological model (Hasselmann et al. 1996). In principle, wind corrections would take a big part of the benefit. However, as more experience was gained in assimilating wave data it has been recognized that such an objective is rather ambitious (Voorrips et al. 1999). A methodical study in the attempt to correct wind fields from the assimilation of wave data is the one presented by Vos (2002). In his experiments, he found that the meteorological model could not benefit from the assimilation of wave data.

It is remarkable that the more sophisticated methods like the Variational and the Kalman filter, do conceptually offer some advantages for data assimilation problems, but so far, their applicability still suffers serious limitations and the existing implementations are not able to handle real situations comfortably.

Although most of implementations have adopted sequential schemes, evolving from Cressman type of analyses, very little attention has been paid to the study of the

errors of assimilation systems. In that regard, some studies like the those from Greenslade and Young (2004), Voorrips et al. (1997), and Sannasiraj et al. (2006) have focused on estimating the background errors. Other implementations have been achieved by calibrating the assimilation parameters.

Little development is noticeable in sequential schemes regarding the retrieval of analyzed wave parameters into the wave spectrum, although from the early studies the distinction of different wave systems was thought to bring in robustness to the assimilation algorithm and palliate in that way the deficiency of the sequential approach.

Chapter 5

Implementation of the OI scheme in the nearshore scenario

5.1 Introduction

In the present chapter, the effect of assimilating offshore buoy data into a local implementation of the WAM model is studied. The aim is to improve wave estimates nearshore and more specifically along the Belgian coast. Wave estimates in this region are fundamental to many coastal and engineering applications and very important for several other marine activities. In practice, wave estimates within a certain domain can only be obtained with the help of numerical models as monitoring is carried out in general only at sparse locations.

When implementing a localized regional wave model, boundary conditions constitute a crucial modelling component together with the wind forcing and the bathymetric data. Boundary information might be obtained either from measurements or from model results from a coarser and larger grid. These two sources of information have advantages and disadvantages. Buoy measurements are expected to be more accurate but they are sparser, while model boundaries might provide a better coverage along the boundary but with reduced accuracy. Ideally, a combination of these two sources shall result in improvements of the wave conditions along the coast.

The appropriate framework for a combination of model and observations is data assimilation. The method followed in this study is a sequential scheme of the Optimal Interpolation type. The advantage of the OI approach is that it is rather straightforward to implement and it is also very efficient in terms of computational demands. These characteristics make it very suitable for practical use. The assimilation of data in the OI approach (and in sequential schemes in general) is done in two steps. In the first, so called *analysis*, wave fields of integral mean

parameters are corrected according to the observed values. In this step the observations are interpolated into the model grid. In the second step, so called *retrieval*, this new information is passed to the model actual variables (the spectrum). Integral mean parameters are only model output (not model variables).

It can be argued whether wave estimates cannot be improved by other mechanisms rather than by data assimilation. For instance, increasing the temporal and spatial resolution of the driving wind fields is expected to have a direct impact on the wave estimates. Also increasing the spectral resolution of the model might improve estimates. In addition, increasing the bathymetry resolution may also help as wave–bottom interactions become important in intermediate and shallow waters. Moreover, the account for specific mechanisms of the wave evolution within the model physics should result in further modelling improvement. However, the data assimilation approach is relevant to gain knowledge of both the wave modelling and assimilation systems. In addition, this methodology is expected to be important for many other modelling applications.

The study area is the southern North Sea and more specifically the Belgian continental shelf domain. The model used is the WAM model, a state-of-the-art third-generation wave model. Observations are available from directional and non-directional buoys deployed in the domain. In particular, the buoy located at Westhinder (WHI) is relevant for data assimilation experiments. WHI is positioned at about 30Km offshore in relatively deep water (30m) and it has been used in former studies to provide boundary conditions for the local model (IMDC and Alkyon 2002). Wind measurements in the area are also available from a nearby meteorological station (MP7). Details about the model configuration and implementation are given in Chapter 2. In the following sections, a description is given of the data assimilation experiments and the results from the different tests. Finally, the main conclusions derived are presented.

5.1.1 Approach and objectives

In the present section, the implementation of an OI scheme is studied. One of the key issues of the OI approach is the specification of the errors of the modelling system. Knowledge about the spatial structure of the errors can be derived for instance from an extensive data collection in the area or from an operational forecasting system (see Chapter 4). A disadvantage in the present case is that the present model is not used for wave forecasting. In addition, the data analysis approach was not in the scope of the study. In order to set up preliminary experiments, the specification of the errors as described by Voorrips et al. (1997) is used. This scheme was previously tested in the southern North Sea area using also

the WAM model. The main purpose is to gain knowledge of the assimilation system from that starting point and test it in a rather empirical way.

A common approach to evaluate the impact of assimilation schemes is to evaluate the ability of the model to retain the introduced data in the time horizon. However, the time horizon is not a target in the present application since the spatial domain is relatively small. The focus of this implementation is not to improve the forecasting ability of the system but the wave estimates in a hindcast mode that can be used to derive long-term statistics. The assimilation point is only about 30 Km offshore. Waves observed by this buoy are reaching the coast in a matter of a few hours (1h for a 0.18Hz wave). In the present experiments, the impact of data assimilation is evaluated in the spatial domain, looking at the ability of the system to correct wave estimates at further (coastal) locations. Thus, data assimilated at Westhinder is evaluated looking at the corrections performed to the coastal locations. The assimilation of data from a single buoy location is particularly interesting because it allows quantifying more clearly the impact of assimilation. Eventually, other locations can be incorporated following a similar procedure.

Another important aspect is the assimilation time step. Different time steps are involved in the computation of waves in the WAM model (i.e., source terms time step, propagation time step, wind input time step). Shorter assimilation time steps will have a stronger forcing effect in the model to make it agree with observations. For these experiments, the assimilation time step is chosen to agree with the source integration time step (10min). Buoy data at WHI are given every 30 min. Therefore, linear interpolation is applied to obtain mean parameters each 10min.

5.1.2 Assimilation parameters

The wave parameters considered for assimilation are significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$). Actually, the availability of directional buoy data offers the possibility to consider mean wave direction as well. However, given the relatively coarse spectral resolution of the model (30° in direction), and given that there is a good agreement between modelled and observed wind directions, assimilation of directional information has been omitted. In this chapter, only significant wave height is considered. The assimilation of mean wave period is addressed in Chapter 6.

It has been mentioned in Chapter 4, that one of the interesting advantages foreseen for the assimilation of wave data was the potential ability to provide feedback to the driving wind fields. This aspect is not considered part of the scope of this study

given that first, wind fields are already analyzed and the level of accuracy is relatively high. In addition, there is not a validated procedure in order to perform this task properly using a sequential assimilation scheme and considering all the relevant meteorological variables and constraints (see Vos 2002).

5.1.3 Retrieval of assimilated parameters into the wave spectrum

Once wave fields of mean parameters are corrected (analyzed) with observations, that information need to be retrieved into the wave spectrum. One of the limitations of sequential assimilation schemes (compared to variational schemes) is the need of a mechanism to retrieve the analyzed integral mean parameters into the wave spectrum. For the retrieval of H_{m0}, Thomas (1988) considers that given the observed and predicted wind speeds, the wind-sea portion of the spectrum can be estimated. Wind-sea can be updated using the JONSWAP spectrum to agree with the analyzed conditions. Consequently, swell energy can be computed as the part of the energy that cannot be supported by the wind and it is rescaled to match the analyzed value. It should be noted that Thomas (1988) used a second generation wave model in which wind-sea and swell are computed separately. This is not the case in third-generation wave models (WAM), although the use of spectral partitioning techniques (see Chapter 3) may facilitate this operation. Another approach, used by Lionello et al. (1992) is to rescale the whole spectrum (or wave component) by a constant factor, forcing the model to match the value of the analyzed H_{m0}. In the present experiments, a scaling factor will be used to retrieve wave height, following the scheme of Lionello et al. 1992. However, it should be taken into account that other (more sophisticated) mechanisms can be used.

5.2 Data assimilation experiments using the OI approach

5.2.1 Estimating the background error covariance matrix

It has been seen from equation 4.28, that the analysis of the wave field can be expressed as a linear combination between the model wave field and the observations. The gain matrix **K**, contains the weighting factors of the linear combination. Since the analysis has to be carried out under the principle of optimality, the matrix **K** takes into account the error characteristics of the system in the form of equations 4.29, 4.30 and 4.31, accounting for the observations and background error covariance matrices **R** and **Q**. In the case of observations, the errors are assumed to be uncorrelated and conveniently matrix **R** becomes diagonal. In the case of the background errors, they are known to be spatially correlated and

therefore **Q** needs to be specified. This is typically done by fitting an error correlation model (equations 4.32 and 4.33). Often, conditions as homogeneity and isotropy are assumed. The condition of homogeneity states that errors in the modelling system are of the same nature within the assimilation domain. Isotropic conditions mean that the magnitudes of errors do not depend on direction.

Preliminary experiments are set up using the specifications of errors applied by Voorrips et al. (1997). Obviously, some of the parameters of that scheme are not expected to be valid in the present case. Those values where obtained by Voorrips et al. (1997) from fitting their own data at specific buoy locations (others than those used in this study). However, since the study area, the model, and the observations used are similar, this scheme is a good reference to start experiments.

5.2.2 Preliminary experiment (assimilation of wave height)

The assimilation of significant wave height is tested using the correlation model given in equation 4.33, with correlation length L = 200Km and exponent λ = 1.5 (Voorrips et al. 1997). Figure 5.1 shows this correlation model applied to the location of WHI in the study area.

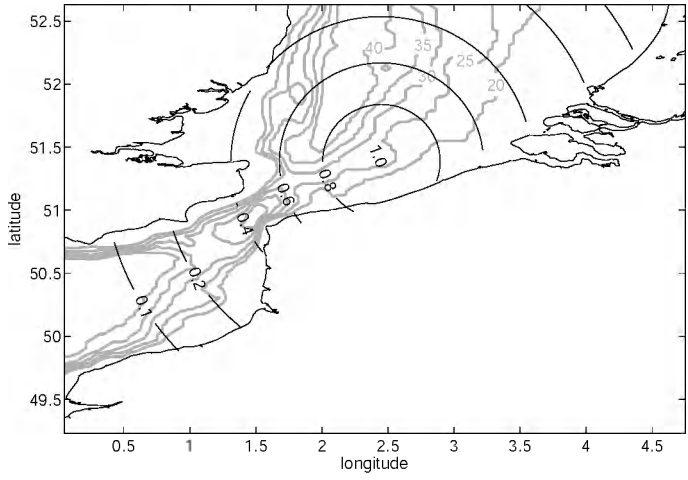

Figure 5.1. Homogeneous-isotropic error correlation function as given by Voorrips et al. (1997), equation 4.33, for the location of WHI on the Local 2 model grid domain. The parameters in equation 4.33 are L= 200Km and λ= 1.5. Bathymetric contour lines are also indicated (in gray).

For the estimation of the mean energy (E_m) background variance (σ_E) Voorrips et al. (1997) obtained:

$$\sigma_E = 0.02 + 0.35E \qquad [5.1]$$

They also considered that the corresponding observation variance is of the same magnitude as that of the background (model), $\sigma_E{}^o = \sigma_E{}^b$.

A preliminary numerical experiment was run using these set of parameters. The simulation was started at 01-Jan-2007 00:00:00 but ended with a model run-time error 3h after initialization. The run-time error itself is due to domain violation of the square root function because after the output time step of 01-Jan-2007 00:03:00 negative values of energy appear at coastal sites nearby the assimilation point.

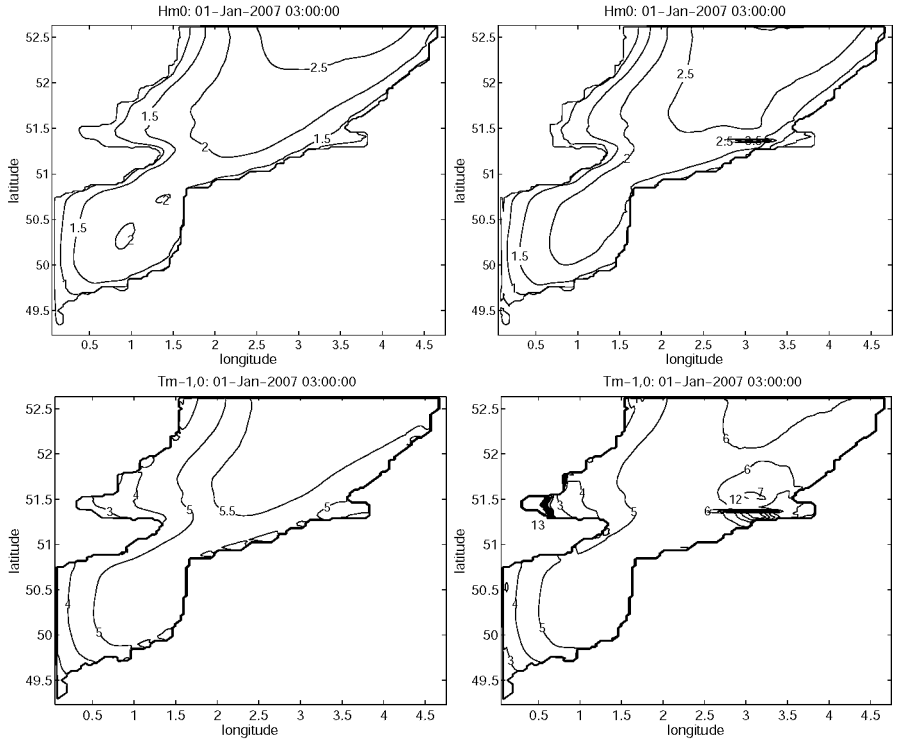

Figure 5.2 Wave fields of H_{m0} and $T_{m-1,0}$ on 01-Jan-2007 03:00:00 (3h after model initialization) for the normal run (left) and for data assimilation test 0 (right).

In order to analyze this problem, wave fields of H_{m0} and $T_{m-1,0}$ are regarded (figures 5.2) for the normal run and for the run using assimilation. In addition, wave spectra

from BVH for the normal run and for the run with assimilation are shown in figure 5.3.

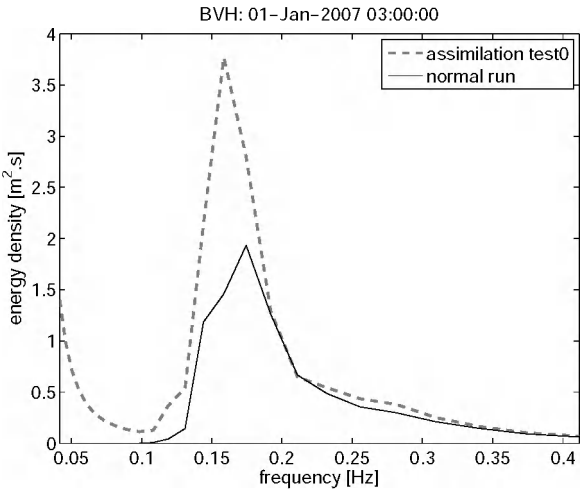

Figure 5.3 Wave spectra at BVH at 01-Jan-2007 03:00:00 for the normal run (continuous line) and for data assimilation test 0 (dashed line).

By looking at the wave fields (figures 5.2) it can be seen that the introduction of external information disturbs the model and creates spectral shapes that are not supported by the model (figure 5.3). The wave field in the assimilation run (figure 5.2) is discontinuous. Values of H_{m0} and $T_{m-1,0}$ are much larger than those from the normal run due to the introduction of data, although wave period as such was not assimilated. Moreover, the model disturbances are present in areas far away from the assimilation location where the wave conditions are not expected to be correlated. For instance values of $T_{m-1,0}$ at the Thames' delta (east coast of England) reach 13s while those from the normal run are of about 3s. Also large values of $T_{m-1,0}$ are present near the assimilation region. Looking at the spectra from BVH (figure 5.3) it can be seen that a spectral feature at low frequencies has developed in the assimilation run with wave energy growing disproportionally.

From these results, one aspect that becomes evident is that influence of WHI data in the domain is unrealistically large. In addition, it is apparent that the creation of discontinuities around the assimilation point is the cause of the run-time error. It should be noted however that this kind of problems is actually expected from sequential schemes, as the corrections are not in agreement with the model dynamics (see Chapter 4). In sequential schemes, the model is forced to agree with observations, causing model instability. This problem as such cannot be eliminated, as it is inherent of the assimilation scheme, but it can be avoided or reduced by adjusting parameters like the correlation length.

5.2.3 Assessing the correlation scale from buoy observations

In order to assess whether the magnitude of the correlation length assumed for the errors is consistent, it is useful to analyze the correlation of mean wave parameters from buoy measurements at some locations. In figures 5.4 significant wave height and second moment wave period from several buoy in the area (i.e., Bol van Heist, A2Boei, Zuid Akkaert, and Deurloo, see figure 2.5) are plotted against the mean parameters at Westhinder. The reason for using wave period T_{m02} in this case, while $T_{m-1,0}$ was used before for the analysis, is that long time series of integral parameters are available for all these buoys. These time series do not contain $T_{m-1,0}$, data and the availability of spectral data is much more limited.

From figure 5.4 it can be seen that the correlation values between the variables at WHI and at the buoy locations in the area are relatively high for both H_{m0} and T_{m02}. These results are consistent with the correlation values assumed for the error correlation model, figure 5.1. The correlation of mean period (T_{m02}) is in general lower than that of H_{m0}. In addition, the comparison of H_{m0} between WHI and BVH for instance, gives an idea of the wave energy dissipation from offshore to coastal locations, the slope of the regression line according to the model y=a + bx is 0.58.

It should be noted that the correlation model from Voorrips et al. 1997 (figure 5.1) was actually obtained from wave energy (wave height) only. For this magnitude the correlation scales of the model are consistent with those from the buoy data. However, some inconsistencies in the spatial structure are visible. In the model used correlation decays exponentially with distance. According to distance only, the correlation value between WHI and DELO data is expected to be lower than that at BVH. However, the value is rather high because water depth conditions are similar. Other remarkable divergence is seen in the correlation values with A2B and BVH. The value of A2B is lower than that of BVH, although A2B is closer to WHI than BVH. That divergence on the other hand can be due to the different type of buoy used or due to particular conditions in the positioning.

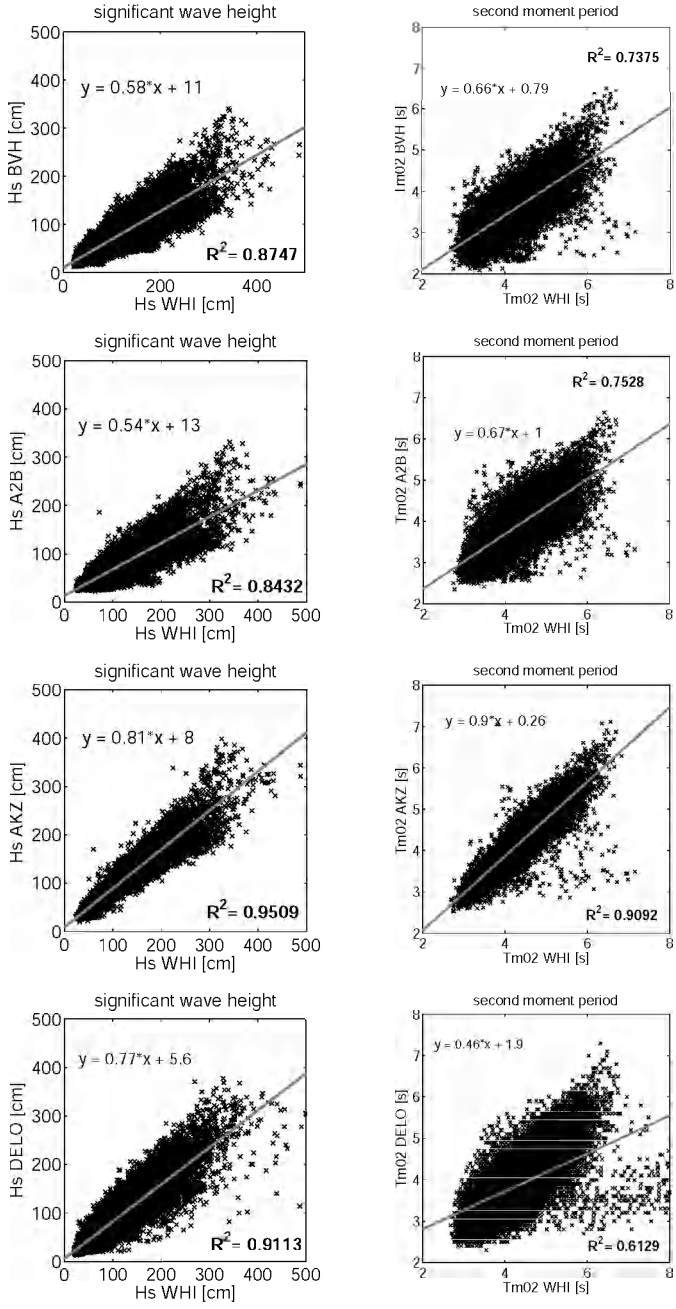

Figure 5.4 Scatter plots of significant wave height and second moment wave period from buoy data between WHI and the locations BVH, A2B, AKZ (periods between 2005 and 2007), and DELO (periods between 2002 and 2005).

5.2.4 Discussion

From equation 4.29 it can be seen that the gain matrix **K**, using the error correlation model from equations 4.32 and 4.33 is composed by four elements. 1) the observation variance, 2) the model variance at the assimilation point, 3) the model variance at the remote location (e.g., equation 5.1), and 4) the model error correlation between the assimilation and the remote locations (covariance, e.g., equation 4.33). All these four elements are actually aspects from the assimilation system that need to be specified or approximated properly, this can be done for instance using one of the methodologies discussed in section 4.4.3. The correlation value can be approximated assuming a correlation structure for the model errors (equations 4.32 and 4.33). For the approximation of the variances, equation 5.1 has been used.

The limitation of the present setup is that none of the elements to compute the gain matrix are well known. Consequently, the assimilation gain matrix is fully parameterized. Using a simplified setup, the gain matrix can be parameterized directly without loosing or gaining any accuracy. The advantage is that the system becomes more flexible for testing and allows a better assessment of the assimilation structure. The disadvantage is that such a scheme turns out to resemble a Cressman type of analysis (Cressman 1959) and the different elements of the OI approach can not be identified explicitly. In view of the relative advantages, this approach is applied in the following sections.

5.3 Parameterization of the gain matrix using an isotropic model

One of the most remarkable problems of the previous setup is the length scale of the assimilation domain. By looking at figure 5.1 it is evident that the correlation length is excessively large for this area. According to this model, wave conditions between Westhinder and Bol van Heist have a correlation value above 0.6. In addition, waves in the basin of the English Channel have a correlation with Westhinder above 0.1. Although the correlation values from some buoy locations seem to confirm these values, it is also clear that water depth conditions play an important role. On the other hand, it is also apparent that the homogeneous isotropic (concentric) correlation model used is rather inconvenient for the area. In any case, one of the most important aspects to assess is the extension of the assimilation domain.

Since the elements of the assimilation system are not well known, in the following experiments observations are assumed to correspond to the true values. In those conditions matrix **R** in equation 4.29 becomes zero and the gain matrix can be

approximated using the correlation model itself, like in a Cressman type of analysis. Later on, in Chapter 6, an assessment will be made to quantify the effect of accounting for observation uncertainties. The numerical experiments of the next section have been done in order to get an idea about the extension of the assimilation domain.

5.3.1 Numerical experiments (assimilation of wave height)

For the present experiment a parameterized gain matrix using the function given by equation 4.33 (with the same parameters as those presented in figure 5.1) is used to assimilate WHI wave height data. Similarly as in the previous experiment, this run resulted in a model run-time error 3h after initialization. This is not surprising because the extension of the domain is the same, and the magnitude of the assimilation gain factors is even larger due to the zero variance assumed for the observations.

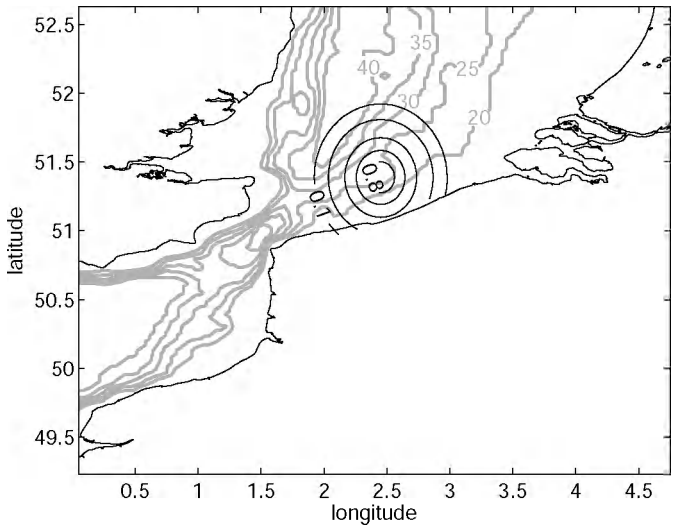

Figure 5.5. Parameterized gain matrix for the location of WHI on the Local 2 model grid domain according to equation 4.33. The parameters of equation 4.33 are L= 25Km and λ= 1.5. Bathymetric contour lines are also indicated (in gray).

Two other tests were carried out reducing the correlation length to ½ and ¼ of the initial value (i.e., 100Km and 50Km). In those experiments, also run-time errors are experienced (6h and 8 days after initialization respectively). The following test experiment was run with a correlation length of 25Km (1/8 the initial value). In this case, the simulation was completed without run-time errors. Results of this run are presented in figures 5.6 to 5.11 and in table 5.1.

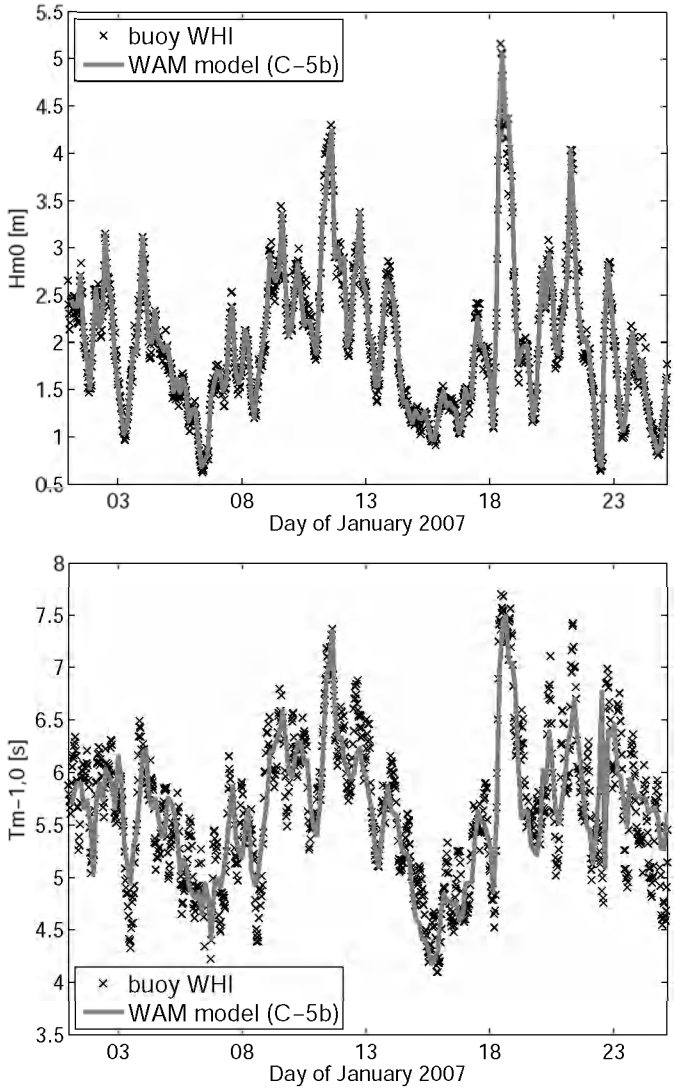

Figure 5.6. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Westhinder (WHI) for the simulation period of January 2007 using data assimilation, test C-5b.

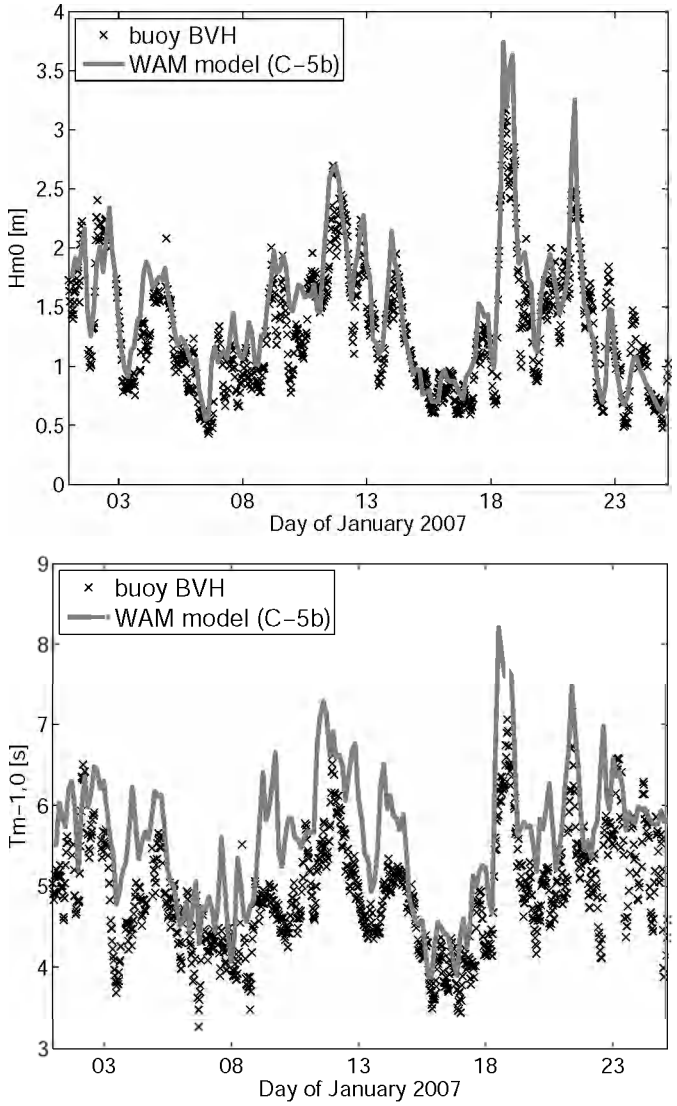

Figure 5.7. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Bol van Heist (BVH) for the simulation period of January 2007 using data assimilation, test C-5b.

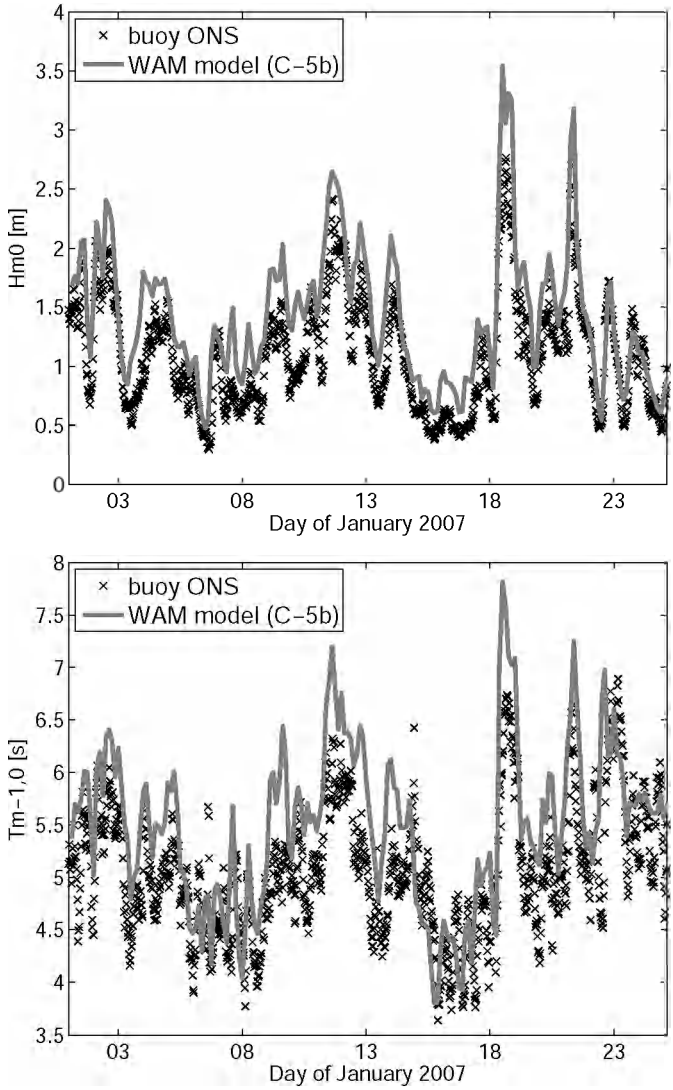

Figure 5.8. Time series of observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Oostende (ONS) for the simulation period of January 2007 using data assimilation, test C-5b.

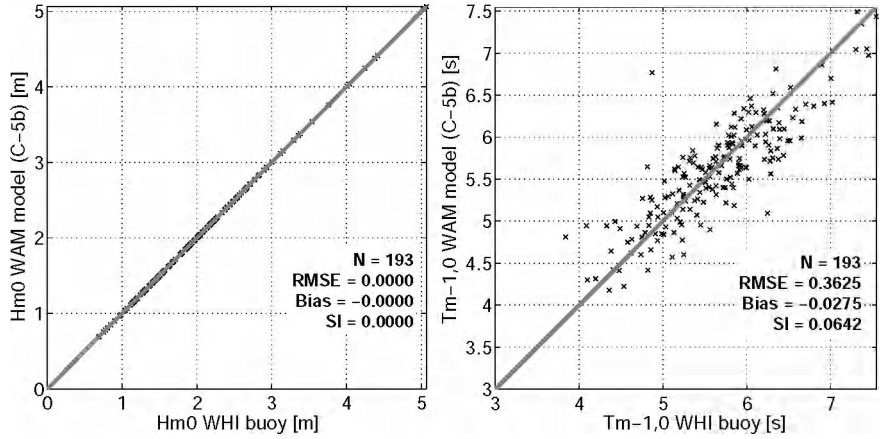

Figure 5.9. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Westhinder (WHI) for the simulation period of January 2007 using data assimilation, test C-5b.

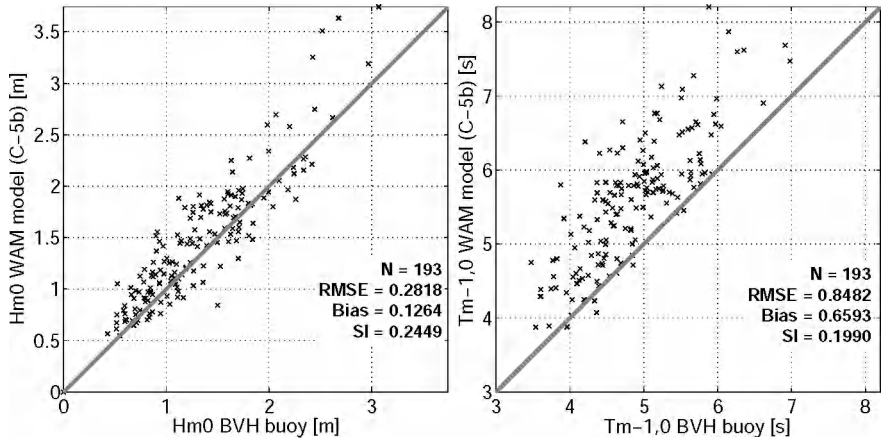

Figure 5.10. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Bol van Heist (BVH) for the simulation period of January 2007 using data assimilation, test C-5b.

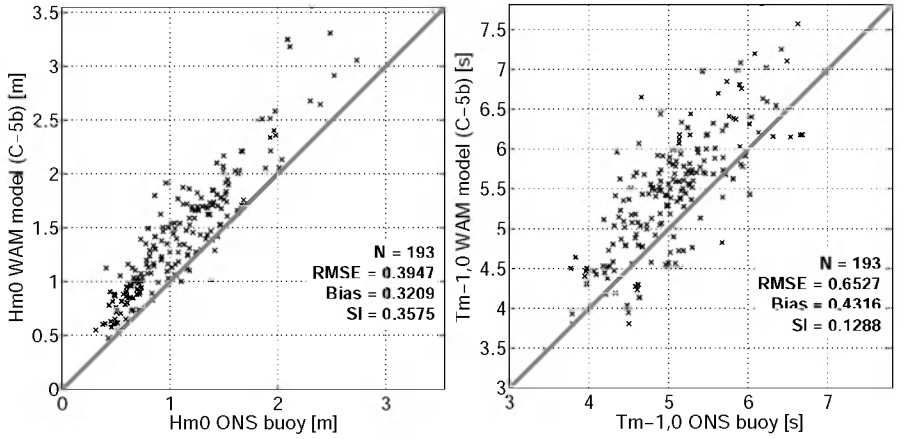

Figure 5.11. Scatter plots between observed and calculated significant wave height (H_{m0}), and mean wave period ($T_{m-1,0}$) at Oostende (ONS) for the simulation period of January 2007 using data assimilation, test C-5b.

Table 5.1 Statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder (WHI), Bol van Heist (BVH), and Oostende (ONS), for the normal run and data assimilation experiment C-5b using a concentric function for the gain matrix. RMSE and Bias are given in meters for H_{m0} and in seconds for $T_{m-1,0}$, SI is given as fraction.

H_{m0} [m]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C5b	0.0000	0.0000	0.0000	0.2818	0.1264	0.2449	0.3947	0.3209	0.3575

$T_{m-1,0}$ [s]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C5b	0.3625	-0.0275	0.0642	0.8482	0.6593	0.1990	0.6527	0.4316	0.1288

From table 5.1 and figures 5.6 and 5.9, it can be seen that values of H_{m0} at WHI are forced to agree with the buoy data. In terms of $T_{m-1,0}$ the statistics at WHI are considerably improved by the assimilation of H_{m0}, although $T_{m-1,0}$ as such was not

assimilated. The most considerable improvement at the coastal locations is seen at the periods around the 14th and after the 21st of January 2007, for which the normal model run was deficient. The improvement of H_{m0} is also appreciable in the statistics. In terms of *rmse* and scatter index improvement is shown for BVH, although bias increased considerably (from -0.0189 to 0.1264). This means that the assimilation is causing considerable overestimation. For ONS all statistical parameters regarding H_{m0} are deteriorating. Regarding $T_{m-1,0}$, results are worse in the assimilation case for the two locations BVH and ONS, except for the SI at ONS that has been appreciably reduced. It should be noted that the weighting factors at BVH and at ONS are 0.16% and 7.2% respectively. Therefore, it is apparent that the worsening in estimates at ONS is due to a combined effect between the direct influence of the assimilation of WHI observations at that specific location plus the fact that waves crossing the vicinities of WHI and reaching the coastal locations have been already subject to corrections.

On the other hand, although the overall results of this set of experiments don't look very positive, interesting results are displayed in the previous plots. Namely, situations in which model results were deficient seem systematically improved, not only in wave height but also in mean period. On the other hand, there is a limitation of the assimilation system regarding the modelling behaviour. Considering wave height, bias at WHI is negative while at ONS this parameter is positive. Bias at BVH is also negative but in lower magnitude, and in many situations wave heights are overestimated. In these conditions, assimilation requirements for WHI and the coastal locations are conflicting. The assimilation of data at WHI demands energy input, while the system requires a reduction of energy to agree with observations at coastal locations. Moreover, the present results show that using a reduced spatial assimilation domain the system is able to run and to produce consistent results.

5.3.2 Discussion

From the comparison of mean parameters between WHI and different buoy data in the domain (figures 5.4) it is clear that there is a high degree of correlation in wave conditions in the area. However, from the experimental results it is appreciable that the gain matrix assumed in the previous tests is not appropriate for the area. Ideally, the assimilation system should affect directly areas that are highly correlated, while coastal locations would be affected only indirectly, from the changed wave conditions at offshore locations. In the OI scheme the variances at the corresponding locations included in the computation of the covariance matrix (in equation 4.32) account for this to a certain extent. However, the manipulation of those parameters without a proper assessment turns the scheme obscure.

When setting up an OI scheme in the nearshore scenario, the considerations of homogeneity and isotropy impose limitations to the system. These limitations are both of physical and practical nature. For the first ones, the fact that waves propagate from deep to intermediate and shallow waters means that the main physical processes responsible for the equilibrium of the wave spectrum are either different or have different magnitudes in a very short spatial range (see figures from 2.1 to 2.4, and table 2.1 in Chapter 2). As a consequence the background errors in that short range are also expected to be of different nature and magnitude. Consequently, the condition of homogeneity assumed is not fulfilled. A similar constraint affects also to the assumption of isotropy. As the bathymetric effects play an important role in the wave evolution nearshore, the errors structure is also expected to be related to the bathymetric shape and to the geographical configuration of the basin (that determine the fetch characteristics of wave generation) rather than to the mere distance from the observing point.

Considering that the correlation structure is one of the key elements of the assimilation system, the following numerical experiments will aim to investigate the effect of assuming a gain matrix that is adapted better to the geographical conditions of the area.

5.4 Parameterization of the gain matrix using an anisotropic model

Obviously, the assumption of any other correlation structure is rather arbitrary and it can only be assessed empirically from the assimilation results. In that sense, the assimilation of data from WHI at this area is particularly convenient since previous studies (IMDC and Alkyon 2002, De Mulder et al. 2004) have shown that wave conditions observed at WHI are representative for the area. They imposed WHI observations as boundary conditions for a local (coastal) model with satisfactory results. In the following experiments, a geometrical shape consisting of a two-dimensional Gaussian function (equation 5.2) is used.

$$\rho_{ij} = \exp\left[-\left\{a\left(x-x_0\right)^2 + b\left(x-x_0\right)\left(y-y_0\right) + c\left(y-y_0\right)^2\right\}\right] \qquad [5.2]$$

The coefficients *a, b,* and *c* are given by:

$$a = \left(\frac{\cos\phi}{\sigma_x}\right)^2 + \left(\frac{\sin\phi}{\sigma_y}\right)^2,\ b = -\frac{\sin 2\phi}{\sigma_x^2} + \frac{\sin 2\phi}{\sigma_y^2},\ \text{and}\ c = \left(\frac{\sin\phi}{\sigma_x}\right)^2 + \left(\frac{\cos\phi}{\sigma_y}\right)^2$$

From equation 5.2 it can be seen that this 2D Gaussian function has three parameters, σ_x, σ_y and ϕ that can be adjusted to fit conveniently the function on the

modelling basin. This is illustrated in figure 5.12. With this shape, it is expected that the assimilation of data from WHI within the coastal zone will have a similar effect as to changing the boundary conditions in a local coastal model.

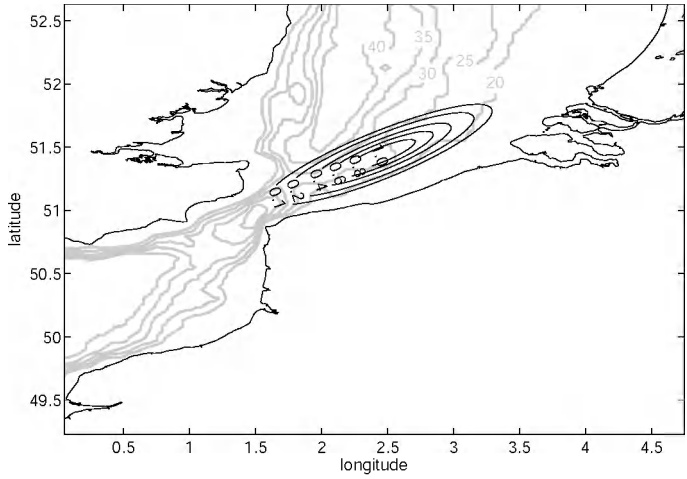

Figure 5.12. Parameterized gain matrix for the location of WHI on the Local 2 model grid domain according to equation 5.2. The parameters of equation 5.2 are σ_x= 100Km, σ_y= 20Km, and $\phi = 20^o$. Bathymetric contour lines are also indicated (in gray).

5.4.1 Numerical experiments (assimilation of wave height)

In the present tests, different values for the parameters σ_x and σ_y in equation 5.2 are tested (table 5.2). The rotation ϕ is fixed to agree with the direction of the bathymetry gradient (~20°). The results of these experiments are presented in the table 5.3.

Table 5.2. Parameters of the two-dimensional Gaussian function (equation 5.2)

	σ_x[Km]	σ_y[Km]	ϕ [°]
C-5b	50	50	-
C-6a	50	25	20
C-6b	50	10	20
C-6c	100	25	20
C-6d	100	10	20
C-6e	150	25	20
C-6f	150	10	20

Table 5.3. Statistical parameters for H_{m0} and $T_{m-1,0}$ at WHI, BVH, and ONS, for the assimilation experiments using a 2D Gaussian shape for the gain matrix. RMSE and Bias are given in meters for H_{m0} and in seconds for $T_{m-1,0}$, SI is given as fraction. Results of test C5b are highlighted with dark gray. The light gray highlights are the experiments using σ_y= 25Km.

H_{m0} **[m]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C5b	0.0000	0.0000	0.0000	0.2818	0.1264	0.2449	0.3947	0.3209	0.3575
C6a	0.0000	0.0000	0.0000	0.2795	0.1163	0.2429	0.3779	0.2917	0.3423
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6c	0.0000	0.0000	0.0000	0.2691	0.1271	0.2338	0.3801	0.3070	0.3442
C6d	0.0000	0.0000	0.0000	0.2742	0.1077	0.2383	0.3747	0.2831	0.3394
C6e	0.0000	0.0000	0.0000	0.2650	0.1339	0.2303	0.3814	0.3119	0.3454
C6f	0.0000	0.0000	0.0000	0.2716	0.1108	0.2360	0.3754	0.2869	0.3399

$T_{m-1,0}$ **[s]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C5b	0.3625	-0.0275	0.0642	0.8482	0.6593	0.1990	0.6527	0.4316	0.1288
C6a	0.3669	-0.0456	0.0650	0.8148	0.6371	0.1912	0.6138	0.4033	0.1211
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6c	0.3704	-0.0299	0.0656	0.8298	0.6522	0.1947	0.6285	0.4336	0.1240
C6d	0.3745	-0.0528	0.0663	0.7791	0.6111	0.1828	0.6011	0.3924	0.1186
C6e	0.3766	-0.0208	0.0667	0.8368	0.6606	0.1963	0.6362	0.4445	0.1255
C6f	0.3761	-0.0428	0.0666	0.7837	0.6164	0.1839	0.6051	0.4019	0.1194

From table 5.3, some conclusions can be derived. Regarding the dimensions of the elliptical structure, length (σ_x) and width (σ_y), a general conclusion is that a narrower (10Km) ellipse produces better results than a wider one (25Km). Test results using 25Km for σ_y are indicated with light gray shaded entries in table 5.3 to facilitate reading. This is the case for all pair of simulations using different width, for the two coastal locations, for the two wave variables presented, and for all statistical parameters. The only exception is H_{m0} at BVH, where slightly better results are

obtained for RMSE and Bias. This indicates that the closer the influence of WHI is in the coastal locations, the worse are the results. For the length of the ellipses, this trend is similar. Wave estimates get worse as the length of the ellipse increases. Again, that is not totally the case for wave height at BVH, where the values of RMSE and SI get slightly improved for larger lengths.

5.4.2 Statistical analysis

Regarding the magnitudes of the statistical parameters, it is interesting to assessed whether the differences are significant or not. In order to give an idea of the statistical significance of these differences, the null hypothesis tests of equal sample variances (F-test) and equal sample means (t-test) are used (see appendix B). For evaluation, results of the experiment C-6b are compared to those of the normal run at the location of BVH.

F-test for testing equal sample variances

The sample variance can be associated to the RMSE by the relationship: $RMSE^2 \approx bias^2 + s^2$. The ratio of the variances follows an F-distribution with v_1-1 and v_2-1 degrees of freedom.

For the normal run (s_1) and test C-6b (s_2), for H_{m0} at BVH. From table 5.3 this gives:
$s_1^2 = (0.3211)^2-(-0.0189)^2 = 0.1027$,
$s_2^2 = (0.2797)^2-(0.1046)^2 = 0.0673$, and the ratio
$s_1^2/s_2^2 = 1.5260$

By looking at table B.1 (Appendix B), the hypothesis of equal variances ratio is fulfilled for less than 150 degrees of freedom for a confidence interval of 5% and less than 110 degrees of freedom for a confidence interval of 10%. The number of degrees of freedom of the sample is 193 but some dependency is expected among the data points. Namely, autocorrelation in the time series decrease the number of degrees of freedom. In order to estimate the effective number of degrees of freedom, Santer et al. (2000) have used a simple approach based on the autocorrelation coefficient. The autocorrelation sequence gives a measure of correlations between points in the time series that are certain interval apart of each other (lag). The autocorrelation sequence can be found by:

$$r_k = \frac{\sum_{i=1}^{N-k}\left(x_i-\bar{x}\right)\left(x_{i+k}-\bar{x}\right)}{\sum_{i=1}^{N-1}\left(x_i-\bar{x}\right)^2} \qquad [5.3]$$

Where r_k is the autocorrelation coefficient for the lag k, x is the variable in the time series, with mean value $\bar{x}$, and N is the sample size. The effective sample size N_e based on r_1, the lag-1 of the autocorrelation regression, is given by:

$$N_e = N\frac{1-r_1}{1+r_1} \qquad [5.4]$$

Calculating the autocorrelation sequence for the time series of H_{m0} for the normal and the C-6b run, one obtains:
r_1= 0.2561
N_e= 193*(1-0.3561)/ (1+0.3561)= 114.3004

Under these conditions, the null hypothesis of equal variances is accepted at 5%, which means that the variances for these two runs are not statistically different at 5% confidence interval. However, the null hypothesis is rejected at 10% confidence interval, in that interval the variances are statistically different.

t-test for testing equal sample means
The differences in bias can be tested using a t-test to determine if two population means are equal. Testing again H_{m0} at BVH gives (see equation B.2, appendix B):

$0.1046-(-0.0189)/\sqrt{(0.1027/v + 0.0673/v)} = 0.1235\sqrt{(v)} / 0.4123 = 0.2995\sqrt{(v)}$

Evaluating for different number of degrees of freedom:

N	$0.2995\sqrt{(v)}$	v_1+v_2-2	$t_5(v_1+v_2-2)$	$t_{10}(v_1+v_2-2)$
10	0.9471	18	2.0894	1.6709
20	1.3394	38	2.0631	1.6675
30	1.6404	58	2.0546	1.6663
40	1.8942	78	2.0504	1.6656
50	2.1178	98	2.0479	1.6652
60	2.3199	118	2.0463	1.6650
70	2.5058	138	2.0451	1.6648
80	2.6788	158	2.0442	1.6646
90	2.8413	178	2.0435	1.6645
100	2.9950	198	2.0430	1.6644

In this case, the hypothesis of equal sample means is fulfilled for 40 and less degrees of freedom at the 5% confidence interval and 30 and less degrees of freedom at the 10% confidence interval. In the case of bias, the null hypothesis is rejected at both 5% and 10% confidence intervals, this means that the differences in bias are statistically different.

Even if differences in *rmse* are not statistically different evaluating for the entire simulation period, for a first general overview of the results, regarding the overall statistical parameters is useful because they allow identifying general trends and spot large deviations. In this sense, we can say that the best results of this series of experiments are obtained with the scheme C-6b although there is not much appreciable difference between C-6b and C-6a for instance, or C-6b and C-6d (the two tests with the immediate larger ellipse in width and length).

Looking at the statistical parameters of test C-6b, it is not evident that the effect of the assimilation is positive. In some cases, the statistics at the coastal locations deteriorate. That is specially the case for bias, which is the parameter that has changed most considerably. Actually, at BVH for instance, the bias of the normal run is remarkably low, both for H_{m0} and $T_{m-1,0}$.

5.4.3 Results analysis

For detailed analysis, it is useful to look at the scatter plots of H_{m0} and $T_{m-1,0}$, these plots are given in figures from 5.13 to 5.18. In these figures, also results from the following section are included in order to have a view of the four experiments for later analysis.

As can be seen from figures 5.13 to 5.18 (upper panels), the very low value of bias from the normal run is rather fortuitous, specially in the case of $T_{m-1,0}$, it does not reflect a nearly perfect simulation. That is also confirmed by the relatively large values of the other statistical parameters. Contrarily, large underestimations for low values and large overestimations for high values of $T_{m-1,0}$ are compensate with each other, resulting in a very low overall value. The same is true for H_{m0} although in less magnitude. Therefore, it is not abnormal that any change in the modelling system would produce worse bias results. However, the overall picture of the assimilation experiment is better than what the statistical figures indicate. Actually, there is a clear trend in the modelling system using assimilation. That trend is a systematic overestimation of the wave conditions, both in H_{m0} and $T_{m-1,0}$. That trend is not clear in the normal run simulation, in which low waves tend to be underestimated while higher waves tend to be overestimated.

The most remarkable effect of the assimilation is the increase in magnitude of the rather low H_{m0} waves, which happen to be low in $T_{m-1,0}$ as well for the model, namely swell waves under low wind conditions. It can be seen that the effect of the assimilation for intermediate and large waves is less notable. Once these low waves are increased by data assimilation, the simulation trend is more regular, and although

intermediate and high magnitude waves have not been changed considerably, the increase of bias is large.

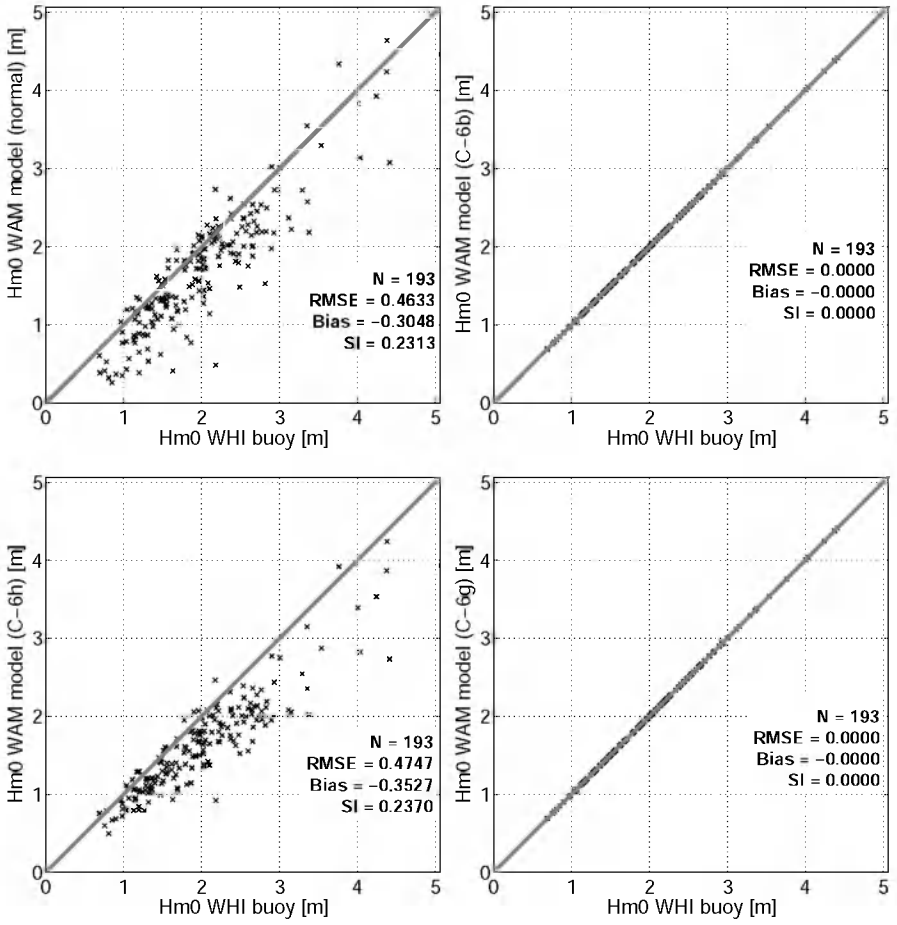

Figure 5.13. Scatter plots between observed and calculated significant wave height (H_{m0}), at Westhinder for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

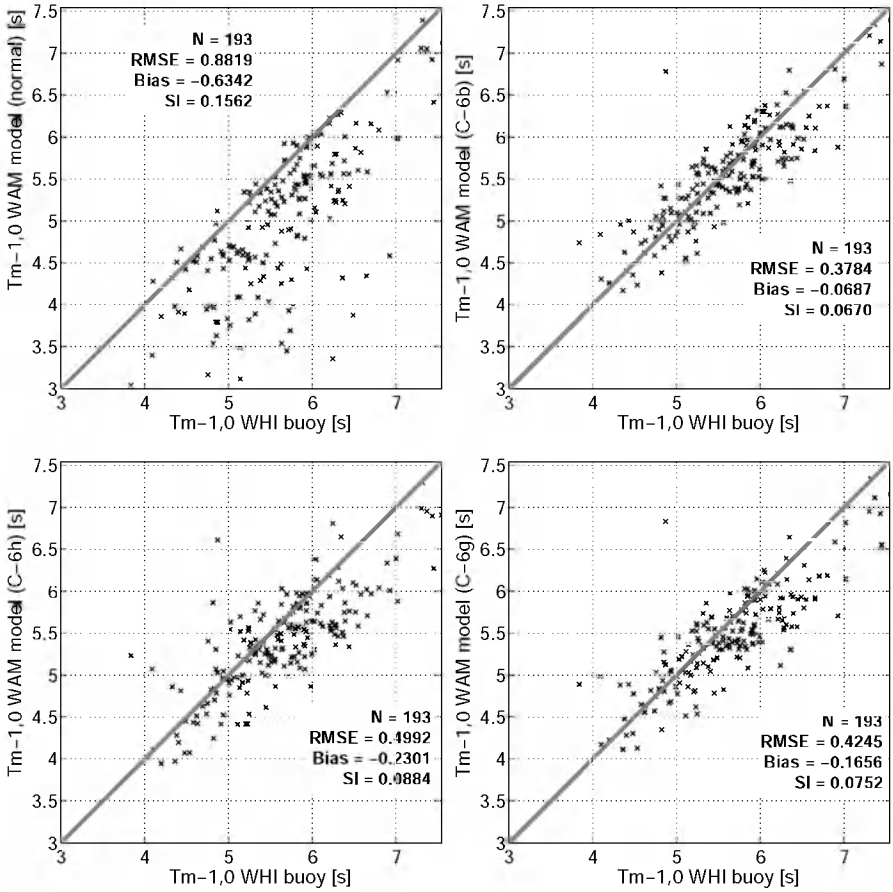

Figure 5.14. Scatter plots between observed and calculated mean wave period ($T_{m-1,0}$), at Westhinder for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

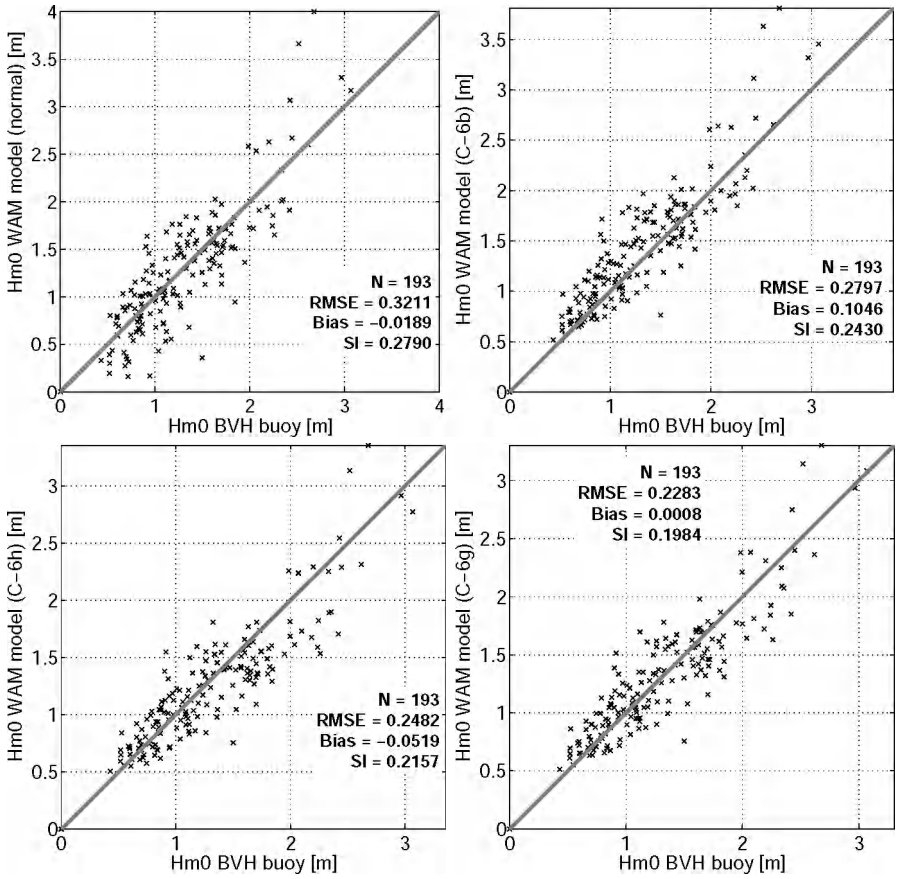

Figure 5.15. Scatter plots between observed and calculated significant wave height (H_{m0}), at Bol van Heist for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

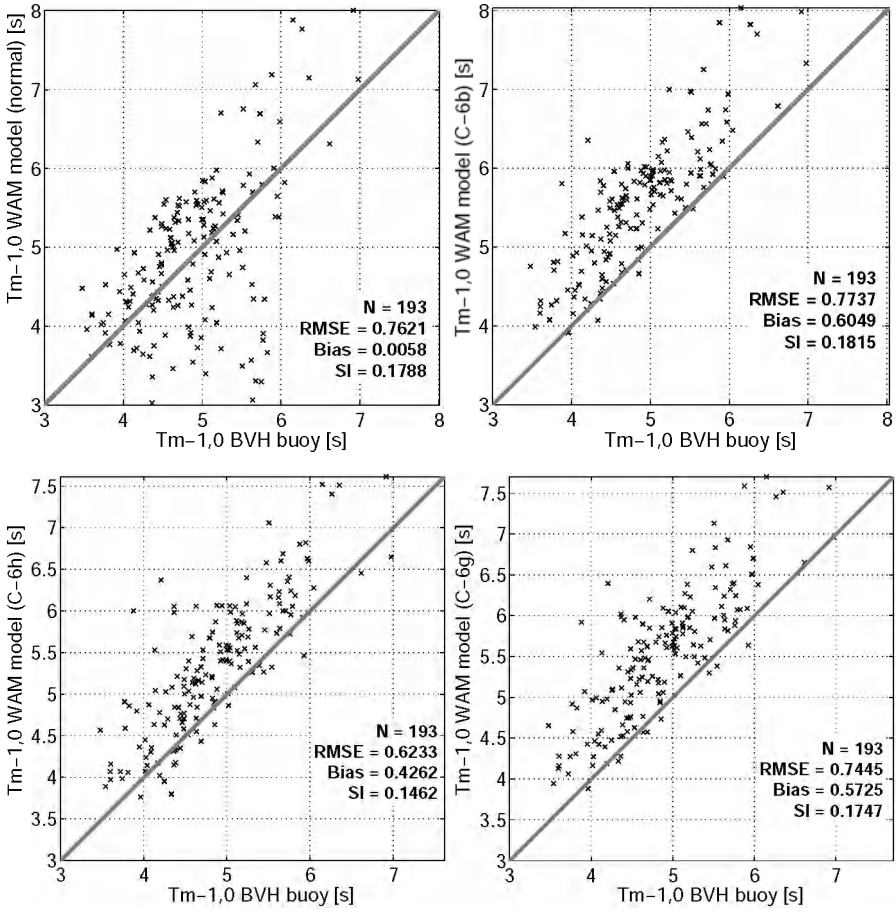

Figure 5.16. Scatter plots between observed and calculated mean wave period ($T_{m-1,0}$), at Bol van Heist for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

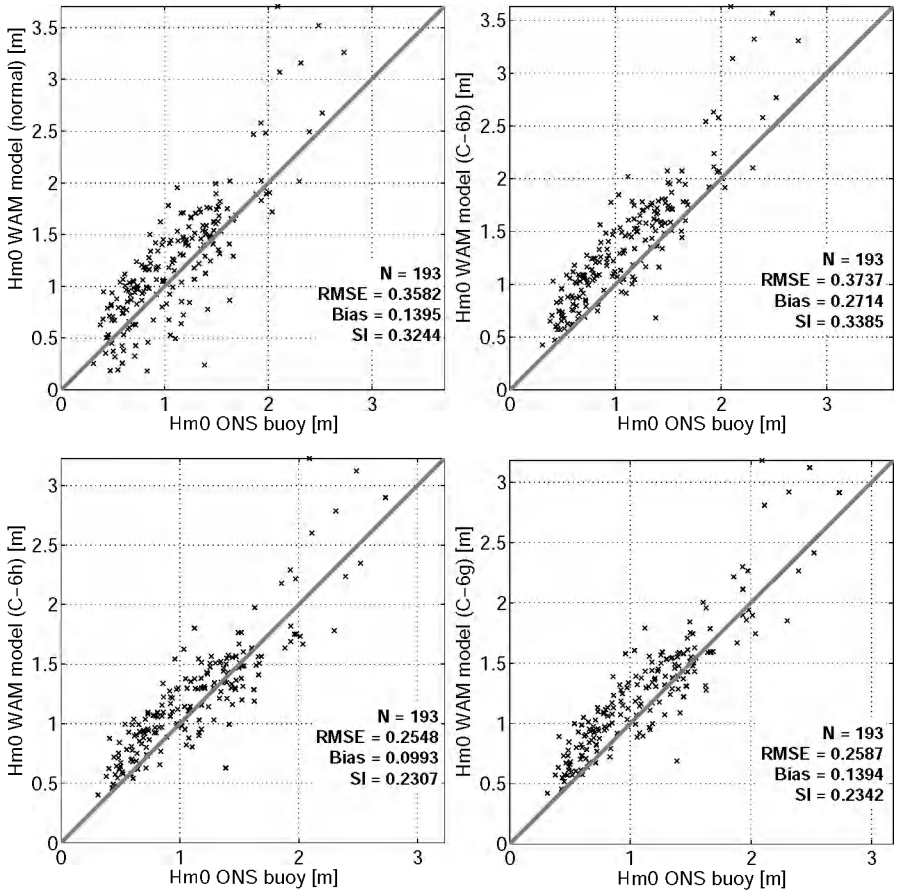

Figure 5.17. Scatter plots between observed and calculated significant wave height (H_{m0}), at Oosetende for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

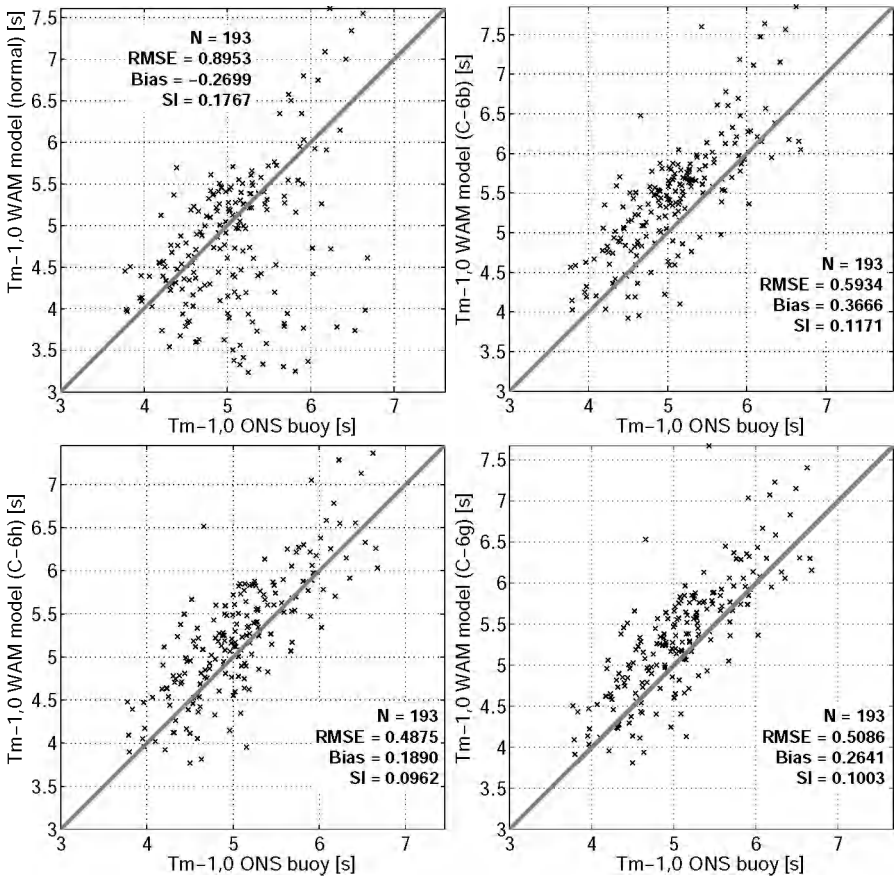

Figure 5.18. Scatter plots between observed and calculated mean wave period ($T_{m-1,0}$), at Oostende for the simulation period of January 2007, for the normal run and using data assimilation, tests C-6b, C-6h, and C-6g.

It should be reminded that only H_{m0} was assimilated in this series of experiments, the increase in $T_{m-1,0}$ is the result of the spectral evolution from model dynamics. In this sense, the assimilation produces the effect of an additional source term. Once high frequency energy is increased, the model is responsible of transferring that energy to lower frequencies mainly by the non-linear interactions mechanism. In this way, $T_{m-1,0}$ gets increased in the whole assimilation domain.

On the other hand, in the assimilation tests (test C6b) waves are overestimated systematically both in H_{m0} and $T_{m-1,0}$. This is not the case for the normal run. However, there is a tendency to overestimate the higher waves. The scatter plots suggest that the trend of the normal run is to underestimate low waves, while high-energy waves are overestimated. It was mentioned before that in the normal model run wind speed is increased by 10% in consideration of former calibration tests. This factor will have a larger effect in the higher winds than in the lower winds. However, looking at figure 2.7 there is a good agreement between modelled and observed wind speed. Nevertheless, since data assimilation acts as an extra source term (an input term in this case), experiments are run to evaluate the effect of adjusting the wind factor in combination with data assimilation.

5.4.4 Numerical experiments adjusting the wind field

In order to assess the combined effect of adjusting the wind field factor and the assimilation of data, further assimilation experiments are considered, still using wave height only. These experiments use the same assimilation gain matrix of test C-6b. In test C-6 g, data assimilation is switched on and the wind speed factor is set to 1.0. For comparison, test C-6h is run with wind factor 1.0 and data assimilation is switched off. Results of these tests are shown in table 5.4.

From table 5.4 it can be seen that wave conditions improve in test C-6g compared to the normal run and compared to test C-6b. Apparently the positive effect of the assimilation is evident form these results. However, results of test C-6h vanish to some extent this conclusion. It can be seen that the results of test C-6h are actually better that those of C-6g and better than those of the normal run. Therefore, comparing C-6g and C-6h it can be seen that the assimilation of data is not improving modelling results. Only significant wave height estimates at WHI are slightly better in the normal run compared to C-6h.

It is interesting to see in the scatter plots of these four runs (figures 5.13 to 5.18) that the general trend of the tests C-6g and C-6h is very similar, both for H_{m0} and $T_{m-1,0}$, and both at BVH and at ONS. That is, there is not a considerable difference in the

trends of the modelling system using data assimilation. That is not the case between the normal run and test C-6b.

Table 5.4. Statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder, Bol van Heist, and Oostende, for the normal run, test C-6b, C-6g and C-6h RMSE and Bias are given in meters for H_{m0} and in seconds for $T_{m-1,0}$, SI is given as fraction.

H_{m0} [m]

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6g	0.0000	0.0000	0.0000	0.2283	0.0008	0.1984	0.2587	0.1394	0.2342
C6h	0.4747	-0.3527	0.2370	0.2482	-0.0519	0.2157	0.2548	0.0993	0.2307

$T_{m-1,0}$ [s]

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6g	0.4245	-0.1656	0.0752	0.7445	0.5725	0.1747	0.5086	0.2641	0.1003
C6h	0.4992	-0.2301	0.0884	0.6233	0.4262	0.1462	0.4875	0.1890	0.0962

Comparing the scatter plots of C-6h with the normal run (figures 5.13 to 5.18) it is remarkable that the normal run shows more scatter of both H_{m0} and $T_{m-1,0}$ (but specially of $T_{m-1,0}$) at low values. Although an overestimation tendency would be expected from the normal run (due to the factor 1.1) with respect to C-6h, it is seen that at low values, H_{m0} and $T_{m-1,0}$ tend to be underestimated. The run C-6h on the other hand show a tendency to overestimate wave conditions similar to the trend of C-6b and C-6g. For instance there are several values of $T_{m-1,0}$ calculated at Oostende for the normal run that are under 3.5 and even 4.0s. Those low values are less often present in C-6h. It is interesting to investigate why these values are lower in the normal run when the wind speeds used are higher.

In addition, it can be seen that the assimilation system is able to correct specially those large deviations from the normal run. Those deviations are identified in the time series of the Oostende buoy. In figures 5.19 time series of H_{m0} and $T_{m-1,0}$ are shown for the buoy measurements, the normal run and test C-6h. Points from the

normal run that diverge from the measurements in more than 1.0s in $T_{m-1,0}$ are also indicated.

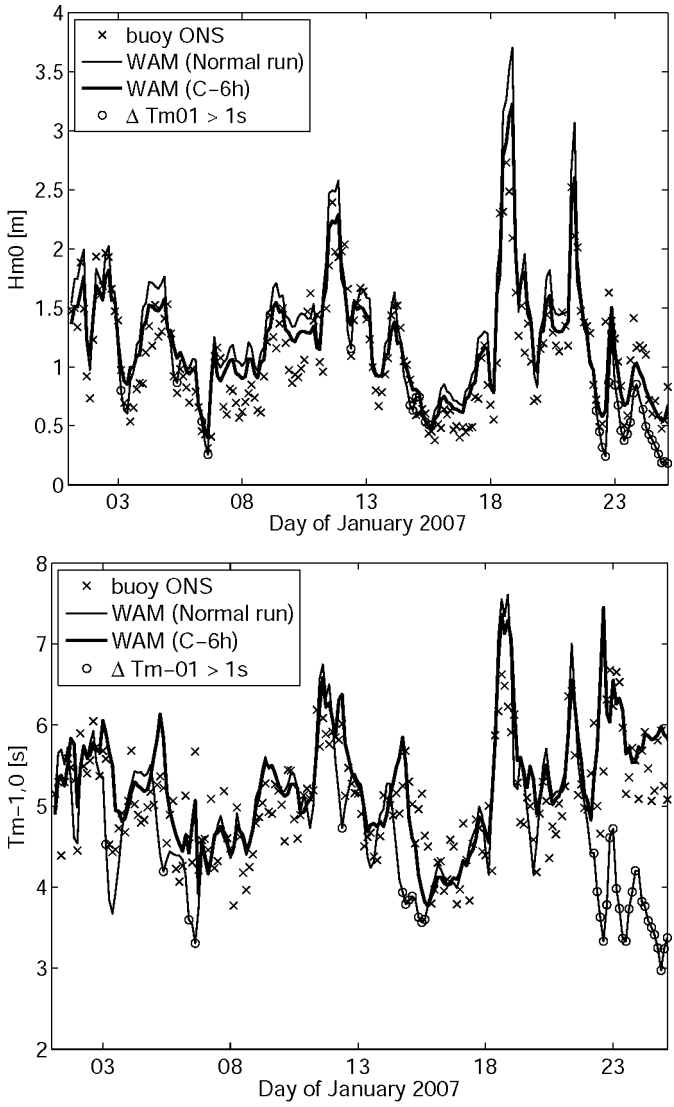

Figure 5.19. Time series of significant wave height (H_{m0}) and mean wave period ($T_{m-1,0}$) at Oostende for the simulation period of January 2007, for the observations, the normal run, test C-6h, and points from the normal run that diverge in more that 1s in $T_{m-1,0}$ from the measurements.

It can be seen from figures 5.19 that indeed the significant wave height from the normal run is systematically higher than that of C-6h, given the higher wind speed.

Other wind related parameters, not shown here (i.e., drag coefficient, wave stress, friction velocities), are also systematically higher in the normal run compared to C-6h. Consequently, the mean wave period of the normal run is lower than that of C-6h (there is more energy in the higher frequencies, wind-sea waves) especially in energy decaying periods. This indicates that dissipation at low frequencies is possibly too high.

The most remarkable differences appear in the period after the 22-Jan-2007 00:00:00. In terms of H_{m0}, the normal run presents a much sharper decay than the C-6h test. In terms of $T_{m-1,0}$, according to the buoy measurements and the results of C-6h, this period corresponds likely to a swell dominated period. While in the normal run it appears as a low energy wind-sea dominated period, characterized by the lower $T_{m-1,0}$ values. The special features of wind conditions at that moment are a fast change in wind direction (from west to north-east), and a considerable decay of wind speed. Actually, when the westerly storm has ceased, a northeasterly moderated wind takes over. This situation is largely affected in terms of modelling performance by the increase in wind speed.

5.4.5 The effect of assimilating wave height in the wave spectrum

It is interesting to analyze what is occurring in the equilibrium of the spectrum in the referred conditions. In order to show the differences among these four tests, some spectra from the period after the 22-Jan-2007 00:00:00 are shown in figures from 5.20 to 5.24, where some interesting features can be appreciated.

In general, spectra for the two tests using data assimilation (C-6b and C-6g) are similar to those of the run without data assimilation and with a wind factor of 1.0 (C-6h). The three of them (C-6b, C-6g, and C-6h) are also in better agreement with the measured spectra than the normal run test. That is specially the case for the peak frequency, in the normal run spectra are always at higher frequency. A positive aspect of the assimilation system is that in any case, the assimilation is able to correct the output, and allows the model to recover the swell system in the spectrum.

It is also interesting to look at the one-dimensional spectra and to the relative magnitudes of the source terms. In figures from 5.25 to 5.30, one-dimensional spectra from WHI are plotted at some particular time instants for: i) buoy measurements, ii) the normal run, iii) for test C-6h, and iv) test C-6b. For the model runs, also the source functions wind input, non-linear interactions, and white-capping dissipation, as calculated by the model, are indicated.

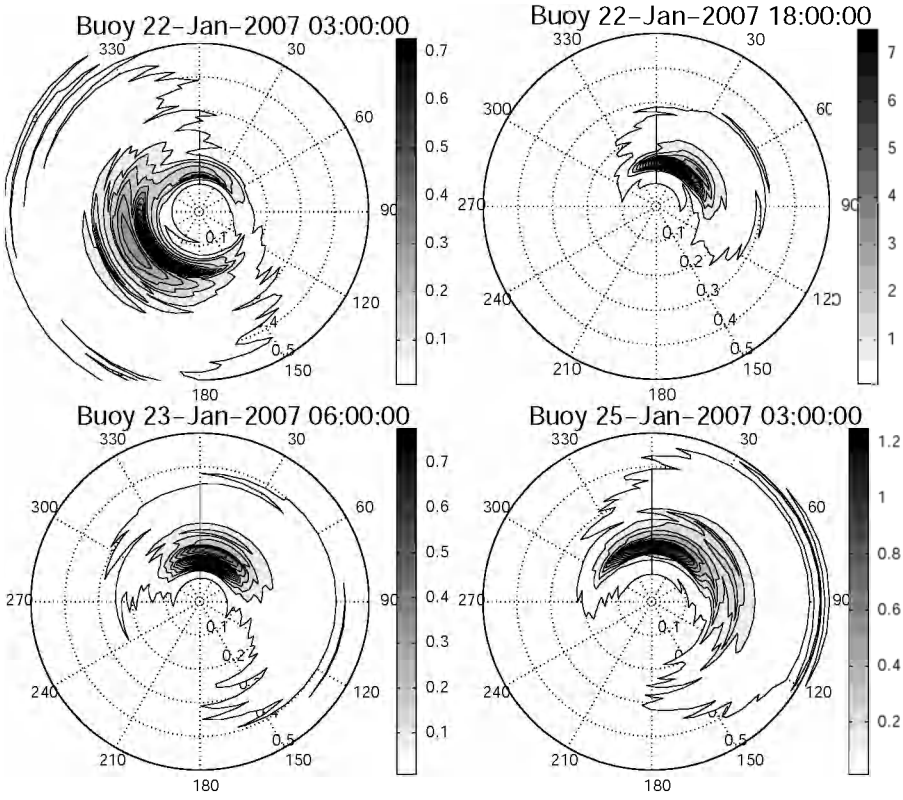

Figure 5.20. Wave spectra on the Jan22-03:00:00, Jan22-18:00:00, Jan23-06:00:00, and Jan25-03:00:00 at WHI for buoy measurements.

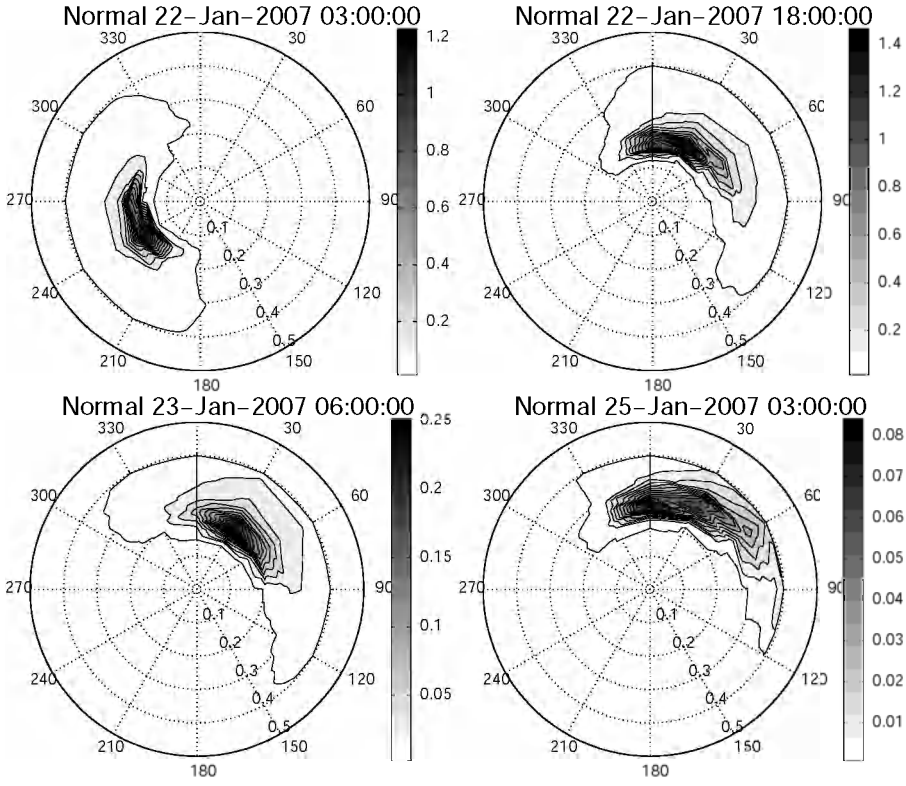

Figure 5.21. Wave spectra on the Jan22-03:00:00, Jan22-18:00:00, Jan23-06:00:00, and Jan25-03:00:00 at WHI for the normal run.

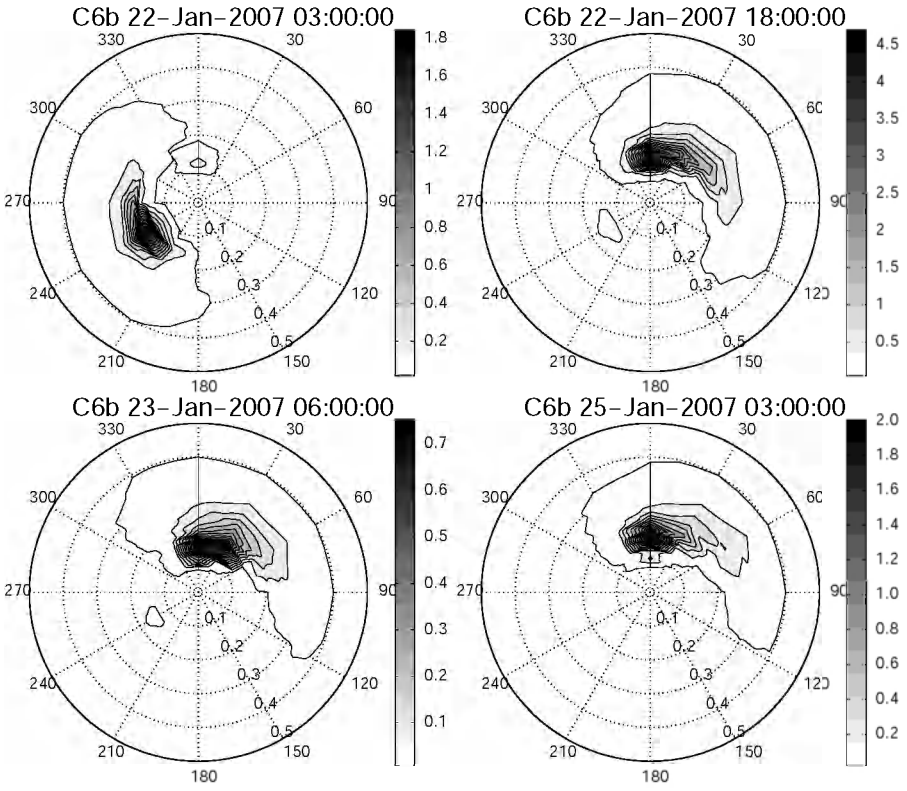

Figure 5.22. Wave spectra on the Jan22-03:00:00, Jan22-18:00:00, Jan23-06:00:00, and Jan25-03:00:00 at WHI for test C-6b.

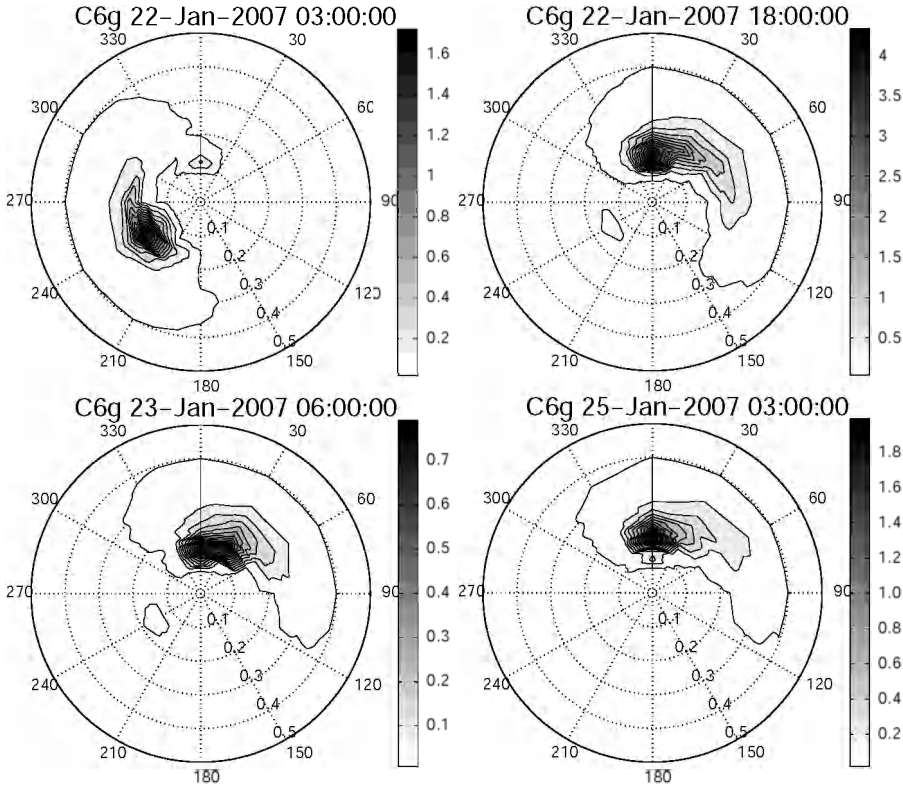

Figure 5.23. Wave spectra on the Jan22-03:00:00, Jan22-18:00:00, Jan23-06:00:00, and Jan25-03:00:00 at WHI for test C-6g.

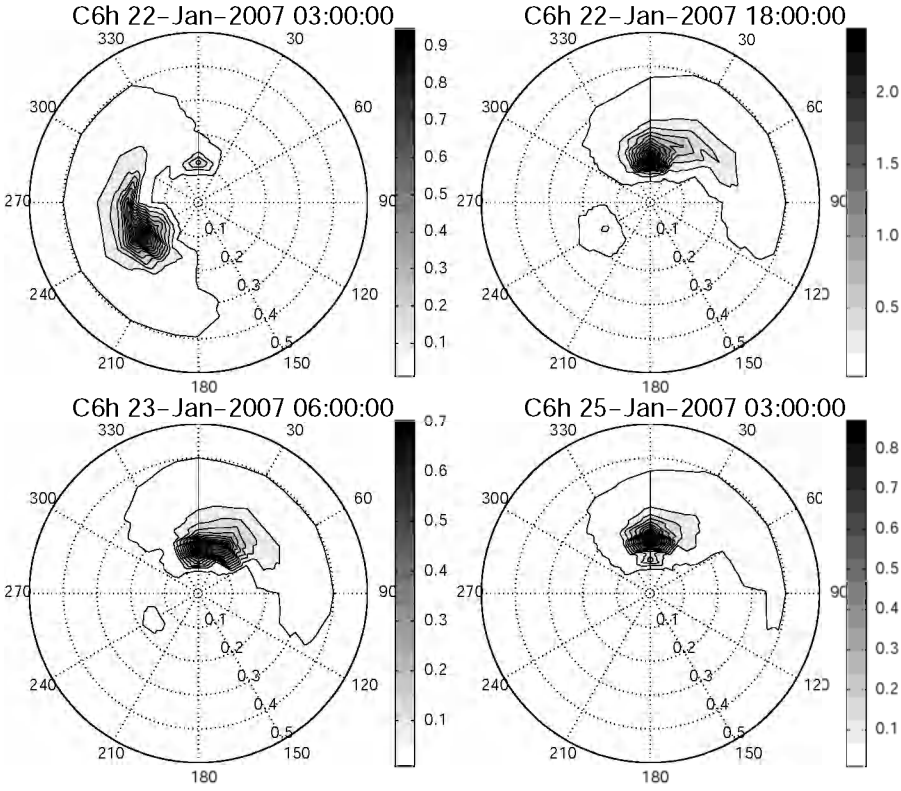

Figure 5.24. Wave spectra on the Jan22-03:00:00, Jan22-18:00:00, Jan23-06:00:00, and Jan25-03:00:00 at WHI for test C-6h.

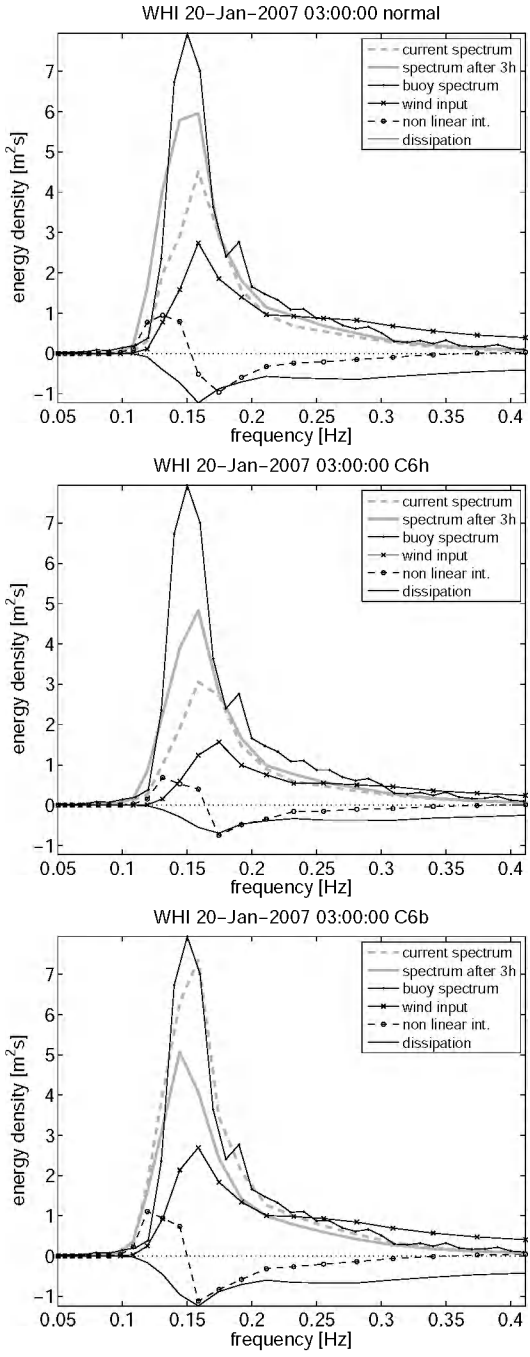

Figure 5.25. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 20-Jan-03:00:00.

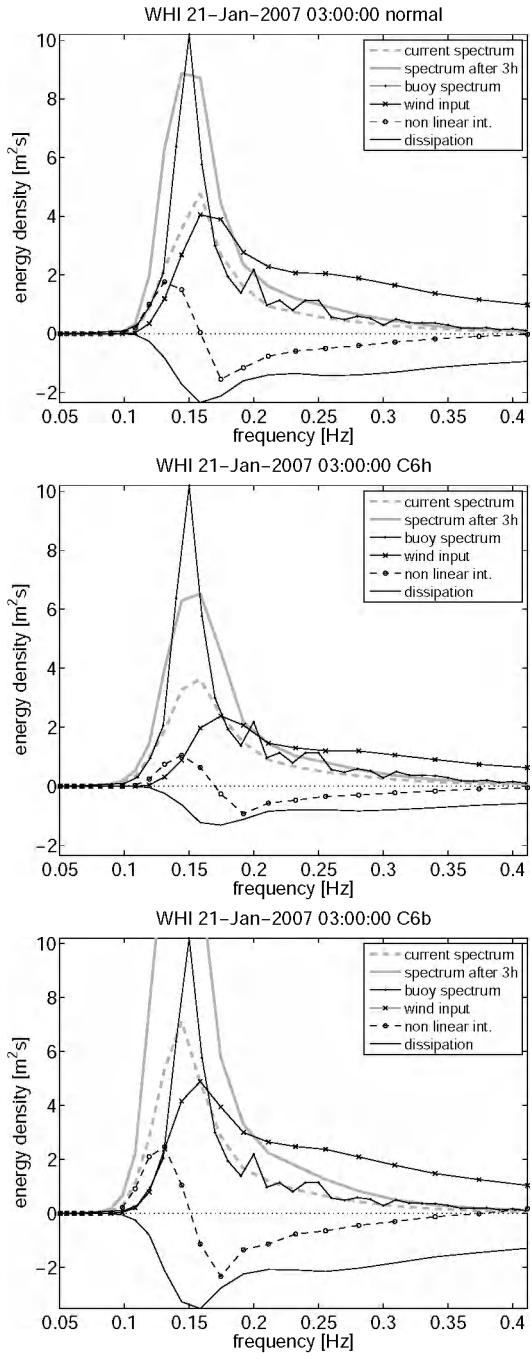

Figure 5.26. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 21-Jan-03:00:00.

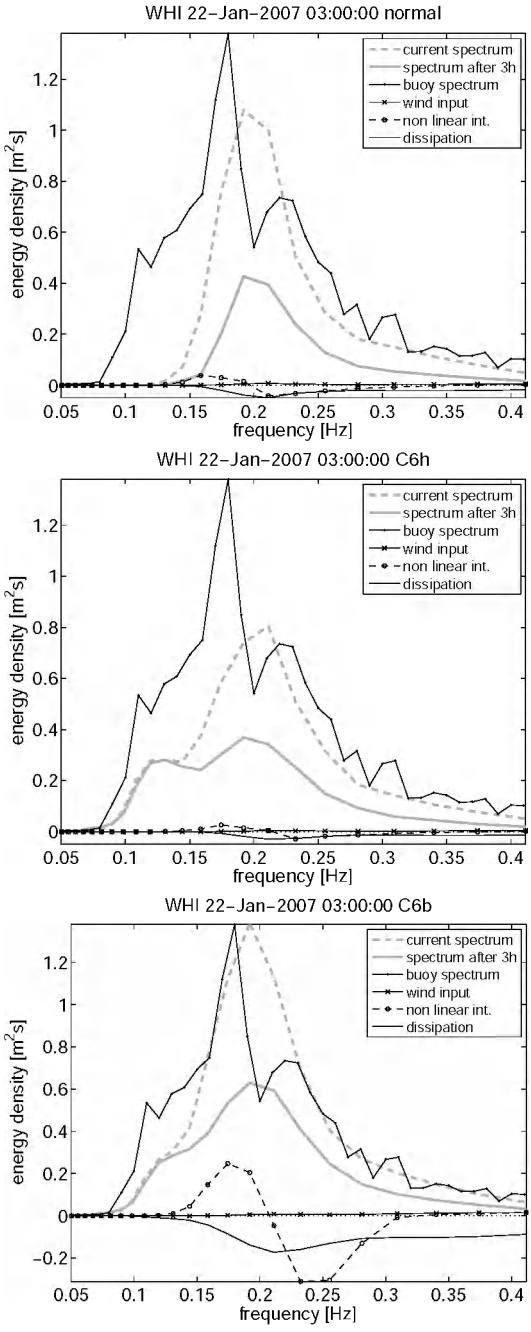

Figure 5.27. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 22-Jan-03:00:00.

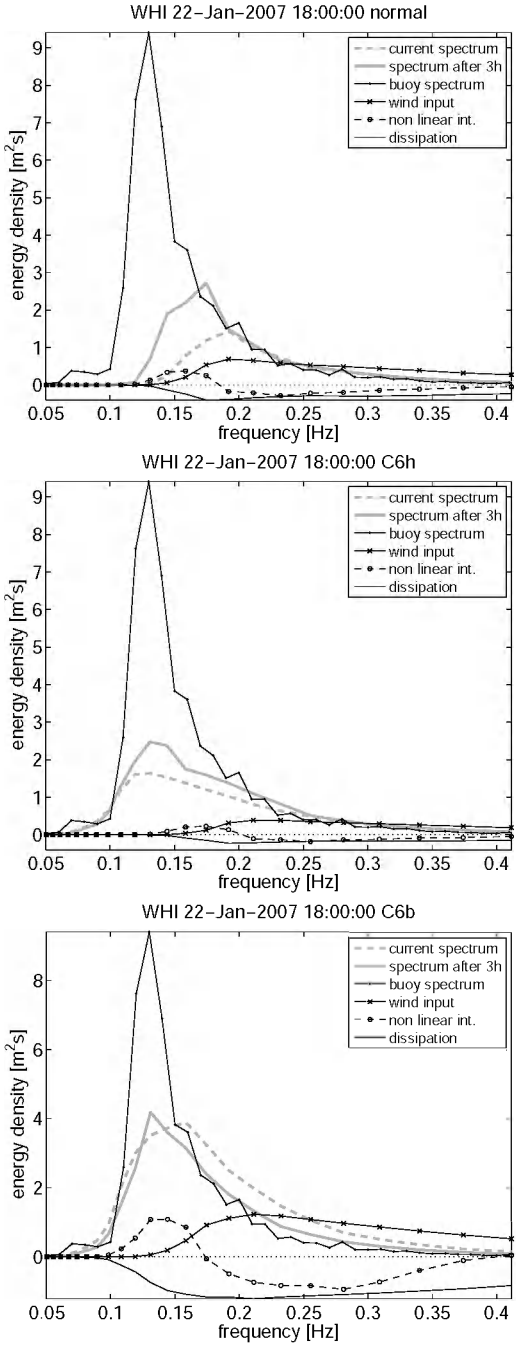

Figure 5.28. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 22-Jan-18:00:00.

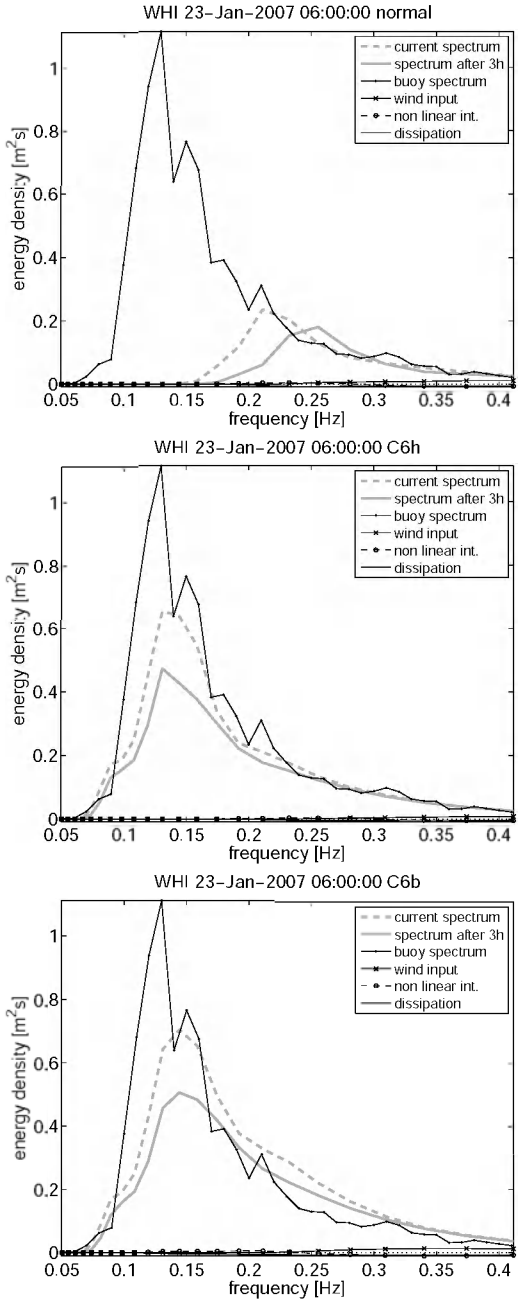

Figure 5.29. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 23-Jan-06:00:00.

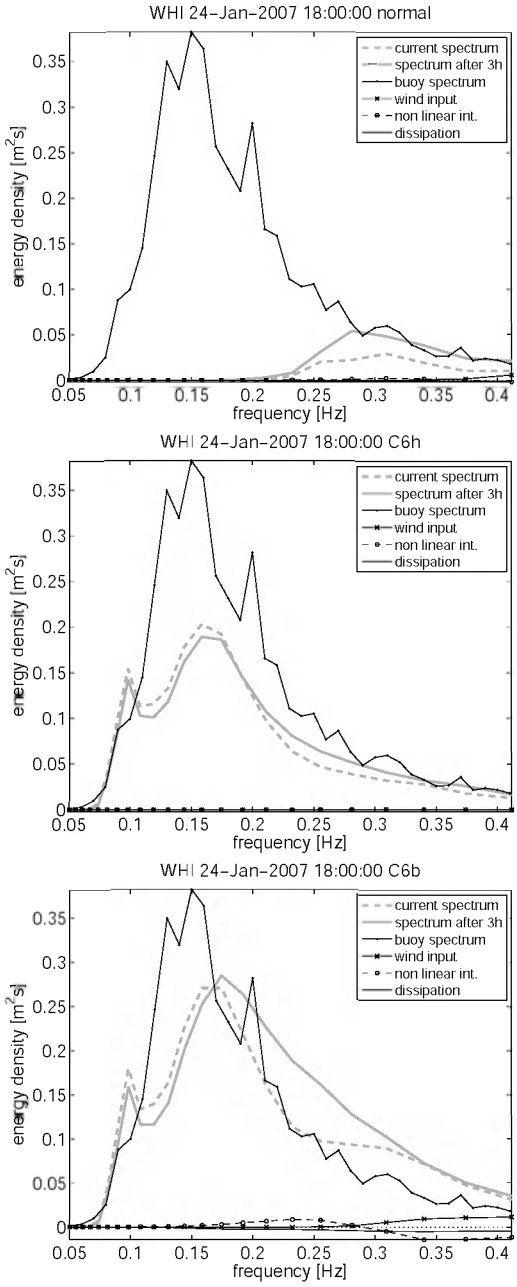

Figure 5.30. One-dimensional spectra and source terms at WHI for the normal run, test C-6h, and test C-6b, at 24-Jan-18:00:00.

Remarkably, in conditions where there is a considerable amount of swell energy and low wind speeds, a relatively small increase in the wind speed produces large energy dissipation due to enhanced white-capping dissipation. Moreover, this large energy dissipation is very detrimental for the energy balance and hinders the development of the swell system after the storm. The deficiency of the white-capping formulation in conditions where wind sea and swell are present in the system has been previously discussed in the literature (e.g., Van Vledder and Hurdle 2002, Rogers et al. 2003, Van der Westhuysen 2007).

From figures from 5.25 to 5.30 it is seen that there is indeed a large dissipation during the storm and decay periods (Jan20-03:00:00 and Jan21-03:00:00) that withdraws energy from the low frequencies and avoid the formation of the swell system (Jan22-03:00:00). Consequently, the energy of the wave system is shifted towards the higher frequencies. In the normal run when the wind starts blowing again it supports the existing wave system but also produces a proportional (large) dissipation that consumes energy from around the peak of the spectrum (Jan22-18:00:00). In test C-6h, the swell system developed during the storm and early decay period is located at much lower frequencies. When wind blows again, this affects only a segment of the tail of the spectrum. Consequently, white-capping dissipation is not able to affect the peak energy and the lower frequency components. Obviously, the assimilation is returning that energy to the system as an extra input term without the price of a large dissipation. In the proceeding period something similar occurs in the normal run. Since the main swell system is not present, an increase in the wind input produces proportional energy dissipation and the model is not able to recover from that deficient initial state.

5.4.6 The effect of data assimilation in time

Other important aspect to assess is the ability of the system to keep the assimilated information in time. Wave conditions are expected to return to the normal condition shortly after assimilation has taken place (within a period of 6 to 12h, Voorrips et al. 1997). In the present assimilation tests the assimilation of data stops at 25-Jan-2007 03:30:00, WHI measurements are not available after that date. The simulation continues until 31-Jan-2007 15:00:00. Nevertheless, data from BVH are available until the end of the model run. These data are used to look at the system after the assimilation period. Time series of H_{m0} and $T_{m-1,0}$ from BVH for the period from 21-Jan-2007 00:00:00 to 31-Jan-2007 15:00:00 are shown in figures 5.31.

It can be seen from figures 5.31 that the results from the tests C-6b, C-6g, and C-6h are very similar. The most remarkable results come from test C-6b. Although assimilation stops at 25-Jan-2007 03:30:00 the system does not return to the normal

condition, but keeps the memory of the assimilated condition until the end of the simulation period (that is for more than 5 days). However, it can not be claimed that the assimilation as such have an influence in the system for more than 5 days. Actually, those conditions have been better simulated not only by the two assimilation tests but also by the test with lower wind factor without assimilation (C-6h).

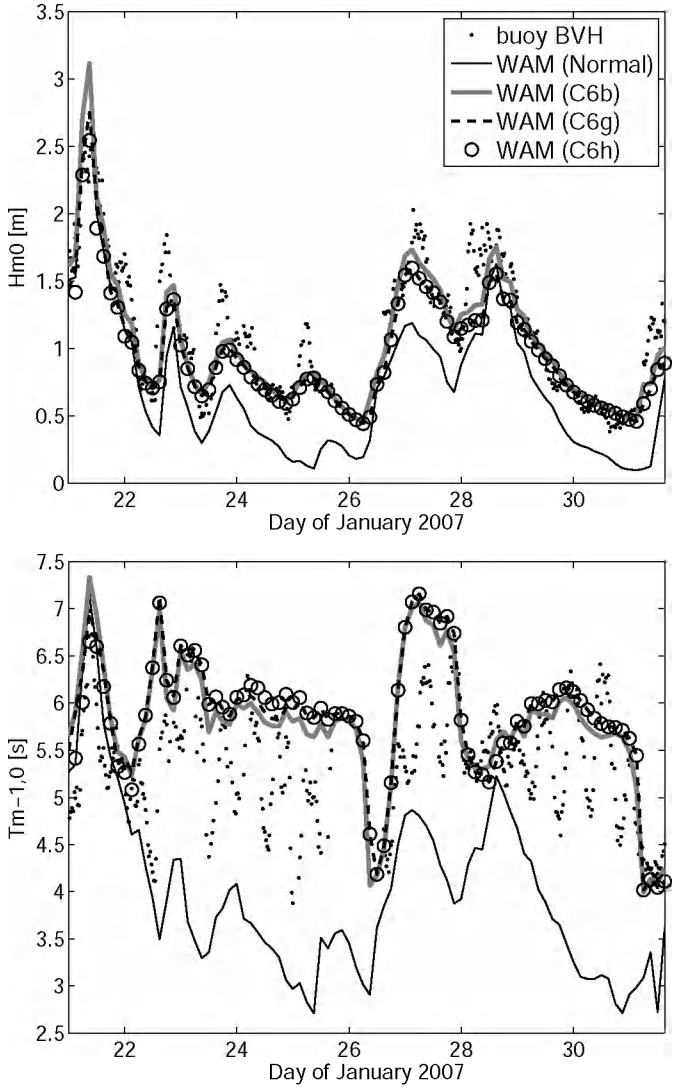

Figure 5.31. Time series of H_{m0} and $T_{m-1,0}$ at BVH for the period from 21-Jan-2007 00:00:00 to 31-Jan-2007 15:00:00 for buoy measurements, the normal WAM run and tests C-6b, C-6g, and C-6h.

The process by which assimilated data is retained in time can be visualized from figures from 5.25 to 5.30. On the 20-Jan-2007 03:00:00 a large wind input is observed in the normal run. The disagreement between the normal run and the observations is not very large, and therefore the dissipation rates between the normal run and test C-6b are comparable. The dissipation in the test C-6h is not only lower, but also it is predominant at higher frequencies. Something similar occurs on the 21-Jan-2007 03:00:00, in this case the disagreement between the normal model run and observations is larger. The input by assimilation is larger and so is the dissipation rate. However, some of the energy is transferred to lower frequencies in the spectrum. During the decay stage, on the 22-Jan-2007 03:00:00, wind input and dissipation are less active, but the dissipation from previous time steps in the normal run hindered the development of low frequency energy. In the assimilation run C-6b, although the dissipation is still larger, there is energy input from assimilation. Consequently, energy at lower frequencies is present and the spectrum resembles more the observed spectrum. The run with low winds C-6h, agrees also well with the observations. Wind increases slightly on the 22-Jan-2007 18:00:00 and although the wind input is lower in C-6h compared to C-6B the total energy is higher and the peak frequency agrees better with the observed spectrum. The input of energy from assimilation in run C-6b results in better agreement with observations, both in terms of energy and frequency. Additionally, in the normal run, the increase of wind causes an increase of dissipation acting again at lower frequencies. Consequently, on the 23-Jan-2007 06:00:00 energy at low frequencies is drastically diminished in the normal run, while there is a better spectral agreement between both C-6b, C-6h, and the observation. As these conditions continue (24-Jan-2007 18:00:00), the normal model run is not able to recover because the initial state of the spectrum is wrong and the moderated increase in wind input is accompanied by dissipation in the front part of the spectrum. This effect will be overridden when the wind input is sufficiently large to reach the equilibrium for which the wind input and dissipation functions have been calibrated.

The initial condition in the sea state in situations where wind sea and swell are present in the system are therefore important for the future evolution of the spectrum. Part of the energy present in the lower frequencies of the spectrum will travel (in space to other locations) due to the advection process. The part that remains is able to receive energy from lower frequencies due to the quadruplets interactions. By this mechanism, the energy transfer from wind is enhanced. Consequently, higher model wave heights occur when wind speed is decreased. The memory effect of the assimilation is the result of improving the initial conditions. This behaviour of the model is less visible in wind sea situations when one storm is proceeding by another. In storm conditions, the balance between wind input and white-capping dissipation is more consistent, while in the swell dominated period

not only the atmospheric boundary condition is important but the initial wave field conditions are also very important.

Finally, it is not evident that results from test C-6g are better than those from C-6h during this period, but it is evident that results from C-6b are better than those from the normal run. Therefore, the assimilation of data has been able to correct the effect of the large dissipation triggered by the large wind input and the consequent large white-capping dissipation.

5.4.7 Spatial patterns

Another important feature to be pointed out appears when looking at the scatter plots of figures 5.13 to 5.18. At the assimilation location WHI, the system systematically underestimates H_{m0} (Bias= -0.6342 Normal run, Bias= -0.2301 C-6h) while that underestimation is much less noticeable at BVH (Bias= -0.0189 Normal run, Bias= -0.0519 C-6h) and contrarily, at ONS wave height is overestimated (Bias= 0.1395 Normal run, Bias= 0.0993 C-6h).

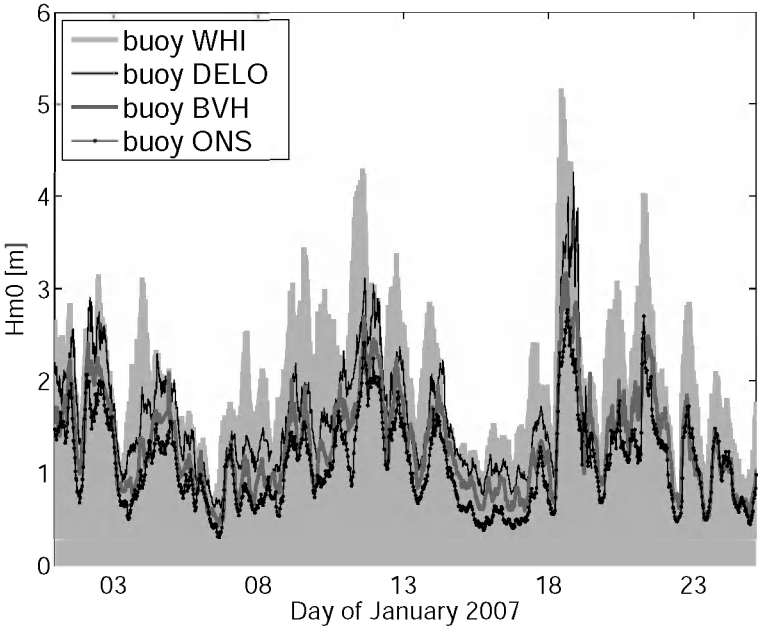

Figure 5.32 Time series of significant wave height (H_{m0}) for the locations of WHI, DELO, BVH, and ONS, for the study period of January 2007. Data from DELO is available only until the 19th of January 2007.

Looking the time series at WHI, it is evident that the introduction of data at WHI will have the effect of an energy input into the system. Because of this particular situation, it is not possible in most conditions to improve wave conditions at the coastal locations since they require a small energy input at BVH and energy reduction in the system at ONS. There are conflicting assimilation requirements in the system regarding the spatial domain. In that sense, it is interesting to see what the wave conditions at other offshore locations are. For instance, wave data (H_{m0}) from the location of DELO is shown in the time series of figure 5.32 together with data from the other buoys WHI, BVH, and ONS.

It can be seen from figure 5.32 that the wave conditions at BVH and ONS are more related to the wave conditions at the Dutch location of DELO. This is not very surprising since DELO is on the one hand closer to BVH and ONS than WHI. On the other, DELO is also closer to the coast (see figure 2.5), therefore, the water depth conditions are also more similar. Regarding the modelling system, it can be seen from figures 2.13 and 2.17 that wave conditions are well represented by the model normal run at DELO. Moreover, the trend is similar to that at BVH. There is a slight underestimation of H_{m0}.

It is apparent that the assimilation of data from a location from the other side of the domain would be very beneficial for the assimilation system. In that sense, the locations of Deurlo (DELO) or Schouwenbank (SCHB) seem particularly attractive. The buoy at Schouwenbank has the advantage that is positioned further offshore and at deeper waters, this allows for a larger assimilation domain. It is also evident that the assimilation system should be able to handle more than one source of buoy data with the same standard procedure. In that sense, the system defined for the previous tests is not very flexible because of the ad-hoc parameterization of the gain matrix. The present system helped on identifying important aspects of the assimilation system, like the appropriate dimensions of the domain and the impact of the assimilation. However, the gain matrix structure as such cannot be comfortably transferred to other locations. In order to assimilate data from other locations, other strategies must be studied to approximate the gain matrix.

5.5 Summary and Conclusions

In this chapter, the implementation of the OI approach nearshore was studied. The implementation of the OI approach requires knowledge of the spatial structure of the errors of the system. Without this information the analysis can not be done objectively according to the statistics of the system but rather in a empirical way. A preliminary test was run using a parameterization described in the literature (Voorrips et al. 1997). This test ended shortly after initialization due to run-time

computing errors. This type of inconvenient is not uncommon in sequential schemes since the assimilation mechanism does not take into account the model dynamics. The analysis suggested that the length scale of the assimilation domain was not appropriate for the area and therefore further tests were set up to tackle this problem. By reducing the size of the assimilation domain, the run-time errors were avoided and the appropriate dimensions of the assimilation domain could be assessed.

The implementation of the OI scheme in the nearshore scenario poses some practical and conceptual issues. In deep water, statistical conditions as homogeneity and isotropy can be assumed for the error correlation structure. However, in the nearshore scenario the relevant processes involved in wave evolution differ over a very short spatial range since the background physical processes within the domain are different. Therefore, homogeneity and isotropy conditions do not hold. This is evidenced not only by poor assimilation results but also by model instability.

It was also noted that the parameterization of the different components of the gain matrix doesn't render the method more objective. Therefore, the gain matrix was parameterized directly according to two different functions, namely a homogeneous-isotropic (concentric) function, and an anisotropic (two-dimensional Gaussian) function fitted in the geographical area. This second function proved to be more appropriate from both the theoretical and the practical point of view. It is observed that the statistical parameters of the assimilation runs are better when the assimilation domain is more restricted.

It is recognized that the performance of the assimilation scheme is actually better than expected. On theoretical bases, the OI scheme is only able to operate in the model output while in practice there is evidence of a constructive interaction between the wave model physics and the OI scheme. The assimilation step behaves as an extra source term and updates the spectrum. In turn, the model processes the new information according to its dynamics.

The present numerical experiments show that it is feasible to improve wave estimates nearshore by data assimilation. Specially, statistical parameters like scatter index and root mean square error show the ability of assimilation to perform corrections. The apparent deterioration of bias, the other parameter observed, is not discouraging because the low bias from the normal model run occurs at the price of a large scatter. In the assimilation runs, a tendency towards overestimation of wave height and wave period is observed but the modelling trend for the different locations is consistent. This is a positive aspect, for instance, for model calibration.

The most important corrections from data assimilation occur in situations of low model performance, namely in moderated wind conditions with the presence of swell. It has been observed that large (white-capping) energy dissipation is triggered by a relatively small increase in wind speed. This large dissipation prevents the formation of swell after storm periods. Accordingly, the formation of wind-sea (from subsequent wind activity events) is also hindered. This energy dissipation could be compensated with data assimilation. Consequently, not only a better overall energy agreement is observed but also energy at the low frequencies is recovered. This model effect is less visible in strong wind conditions since the exponential wind energy growth quickly overrides this effect.

When moderated wind conditions persist, the effect of the assimilated data last for several days. Although this long duration might seem inconsistent considering the travel time of waves in the domain, which is of the order of hours, this effect can be explained because the improved wave conditions. Especially the presence of low frequency wave energy in the spectrum acts as improved initial conditions for the development of waves in subsequent periods. When low frequency waves are present in the system the energy transfer from higher to lower frequencies (due to quadruplet interactions), is enhanced. Consequently, the energy transfer from wind to waves is also enhanced. It should be noted that this effect has interesting implications for forecasting applications.

In wind-sea conditions, the impact of data assimilation is less visible. First of all, because in those cases the model estimates are better. Additionally, the model performance statistics behave differently at offshore and at coastal locations. Significant wave height and mean wave period are underestimated at offshore locations, while those variables are either less underestimated or overestimated at the coastal locations. Under these conditions, wave estimates cannot be improved by single point data assimilation because the assimilation requirements at different locations are conflicting.

The assumption that the main sources of wave modelling errors come from errors in the forcing wind field is misleading for wave data assimilation since wave-modelling errors can play a crucial role as well. Underestimation of wave heights for instance cannot always be associated to wind underestimation. A situation of coexisting wind-sea and swell systems was observed in which increasing wind speeds causes larger dissipation rates and consequently lowers wave heights. Moreover, the ability of the model to represent the spectrum is affected.

Chapter 6

Complementary aspects of the assimilation system

6.1 Introduction

In the previous chapter, some characteristics of the OI scheme in the nearshore scenario were identified. A positive effect of the assimilation was observed not only in some of the statistical parameters but also when looking at specific modelling features. However, in order to cope with implementation limitations, some simplifications to the scheme were applied, mainly related to the parameterization of the errors of the system.

In the present chapter, further experiments are set up in order to overcome some of these simplifications. First of all, the parameterization of the assimilation gain matrix deserves attention. Although assimilation results using a two-dimensional Gaussian function for the gain matrix were positive, such a function cannot be transferred easily to assimilate data from other buoy locations. In order to define a more transferable assimilation function, two other strategies, based on model results, are studied. In the first, the correlation of the model variables is considered, using long-term model estimates. In the second, short-term wave parameters are compared in order to calculate the assimilation gain matrix.

In addition, numerical tests are carried out to study the assimilation of mean wave period. Numerical experiments are also carried out in order to assess the impact of accounting for observations uncertainties. In the approximations of the previous chapter, observations were assumed to correspond to the true values. However, observations do also contain errors which introduced into the model can produce deficient results.

6.2 Parameterization of the gain matrix using long-term model estimates

In this section, the gain matrix structure is calculated from model integral parameters. In this approach, the underlying assumption is that the long-term spatial distribution of model parameters is correct. The correlation coefficients of wave parameters between every model grid point and the assimilation point (WHI) are used to define the gain matrix.

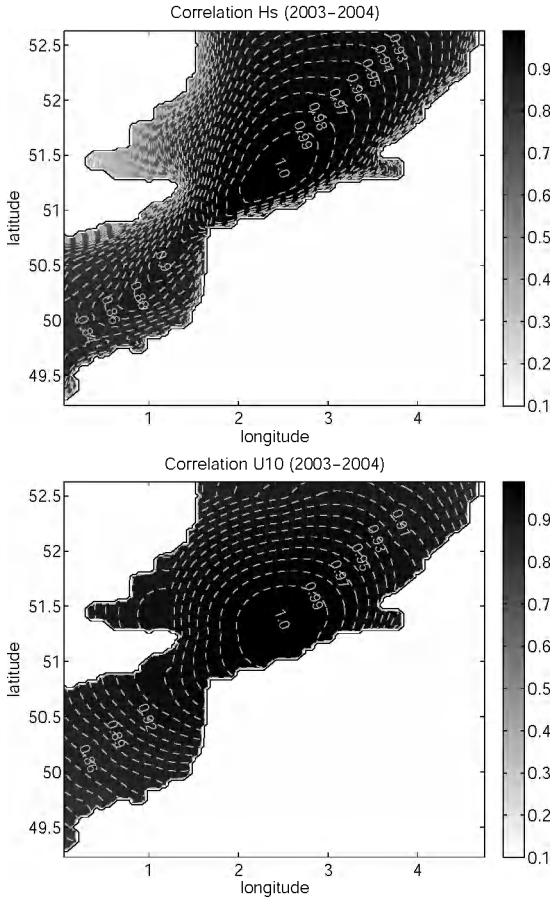

Figure 6.1. Correlation coefficients of significant wave height (H_{m0}), and wind speed (U_{10}), between points of the Local 2 grid and the assimilation point (WHI) for the two years simulation period of 2003 and 2004.

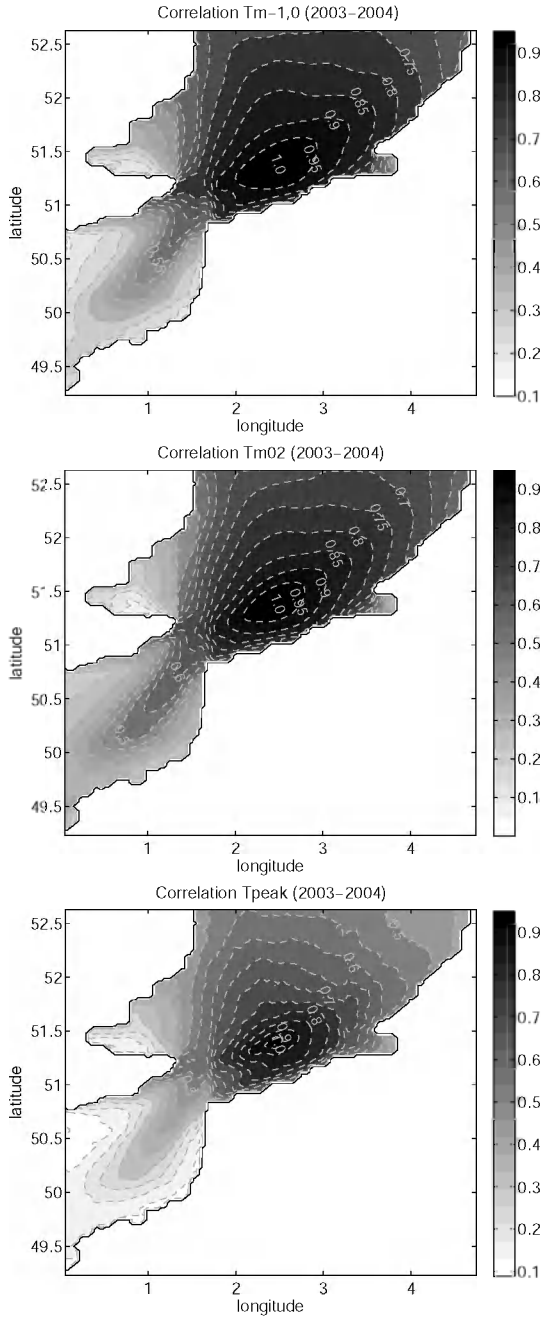

Figure 6.2. Correlation coefficients of mean wave period ($T_{m-1,0}$), second moment wave period (T_{m02}), and peak period (T_{peak}), between points of the Local 2 grid and the assimilation point (WHI) for the two years simulation period of 2003 and 2004.

In order to calculate the correlation coefficients between different grid points, the WAM model was run for a period of 2 years, 2003 and 2004, for which continuous wind field data are available. Correlation coefficients are calculated according to equation 2.25 for different output variables, namely significant wave height, wind speed, first and second moment wave period, and peak period, between each grid point and the assimilation point (WHI). The resulting (gridded) correlation values are presented in figures 6.1 and 6.2.

It can be seen from figures 6.1 that the correlation length from H_{m0} is relatively large. Actually, the dimensions of this function are larger than those given by Voorrips et al. (1997), figure 5.1. In the case of U_{10}, it can be seen that the spatial structure agrees more with the isotropic model. This is largely the effect of the relatively low spatial resolution of the wind field (1.25°). The correlation structures of the other variables have more elongated shapes and correspond better with the geographical configuration. The magnitudes of the correlation values for the different variables are very different (especially between H_{m0} and the wave periods).

6.2.1 Correlation factor

From the experience of the previous assimilations tests, it is clear that none of these spatial structures can be used to represent the gain matrix since the large dimensions are likely to cause model run-time errors due to the influence of the assimilation scheme at coastal locations. Moreover, the tests of the previous section suggested that better assimilation results are obtained when the dimensions of the domain are more restricted. For the present approximation, the gain matrix is calculated as the product of the different factors given by the correlation coefficients, equation 6.1.

$$k_{i,j} = \rho_{i,j}^{Hm0} \rho_{i,j}^{Tm-1,0} \rho_{i,j}^{Tm02} \rho_{i,j}^{Tp} \rho_{i,j}^{U_{10}} \qquad [6.1]$$

where k_{ij} represents the element i, j of the gain matrix **K**, where i is the index of a generic grid point, and j is the grid point at the assimilation location. $\rho_{i,j}^{Hm0}$, $\rho_{i,j}^{Tm-1,0}$, $\rho_{i,j}^{Tm02}$, $\rho_{i,j}^{Tp}$, $\rho_{i,j}^{U10}$, are the correlation coefficients between the points i and j for the indicated variables. The gain matrix structure calculated with this equation is shown in figure 6.3.

It can be seen from figure 6.3 that length scale of this structure is still large and is likely to cause problems in the assimilation system. It has been seen in section 5.2.2 that even a direct influence of WHI at BVH of 0.16% has a negative effect on the model stability. In order to cope with this limitation, two extra factors are introduced for decoupling the coastal zone from the assimilation domain. These factors are

obtained based on the bathymetric characteristics of the basin and the distance from the assimilation point. The introduction of these factors is given in the next section.

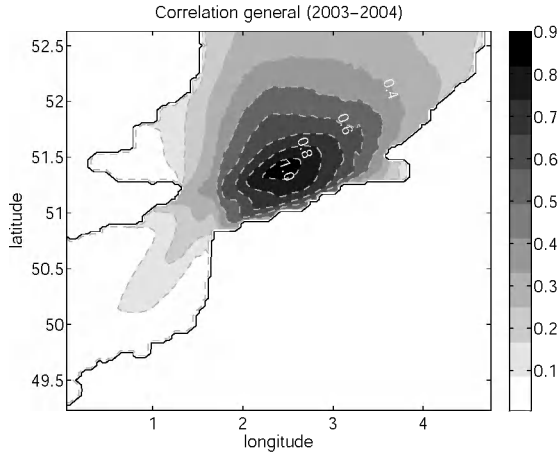

Figure 6.3. Gain matrix structure obtained from the product of the correlation of significant wave height (H_{m0}), wind speed (U_{10}), mean wave period ($T_{m-1,0}$), second moment wave period (T_{m02}), and peak period (T_{peak}), equation 6.1, between each point of the Local 2 grid and the assimilation point (WHI) for the simulation period of 2003 and 2004.

6.2.2 Bathymetry factor

For the bathymetry factor, grid points with water depths above a certain threshold value are removed from the assimilation domain (bathymetry factor is zero). The reason is to avoid interference of the assimilation with very shallow water areas. Grid points with water depths between the threshold value and the water depth at the assimilation point (Westhinder, 30m) are rescaled linearly with a factor from 0 to 1, according to the water depth. The reason is to have a smooth transition between the decoupled coastal zone and the assimilation point. Grid points in deeper waters that the assimilation point are not affected (bathymetry factor is one). The function for the bathymetry factor is given by equation 6.2 and illustrated in figure 6.4 (for the threshold value of 5m).

$$f_{depth} = \begin{cases} 0, & h < h_{th} \\ \dfrac{h}{h_{th}} - 1, & h_{th} < h < h_j ; \quad h > 0 \\ 1, & h > h_j \end{cases} \qquad [6.2]$$

where f_{depth} is the depth (bathymetry) factor, h is the water depth at any model grid point (positive in the downwards direction), h_j is the water depth of the model grid point at the observation/assimilation location (WHI), and h_{th} is a threshold water depth value for full decoupling. The threshold values used in the experiments are 5 and 15m

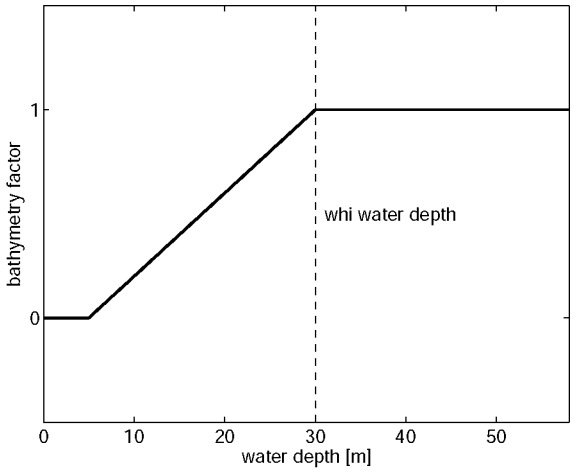

Figure 6.4. Depth (bathymetry) factor function of equation 6.2 used to decouple the coastal zone from the assimilation domain of figure 6.3.

6.2.3 Distance factor

The distance factor is calculated using an exponential decay function (similar to the correlation model of equation 4.33).

$$f_{dist} = \exp\left[-\left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{L} \right)^{\lambda} \right] \qquad [6.3]$$

where $|\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the model grid point i and the assimilation point j, L is the correlation length, and λ= 1.5 is the shape parameter. The lengths used in the present experiments are 50 and 100Km respectively.

The elements $k_{i,j}$ of the gain matrix **K** obtained with these assumptions are calculated according to equation 6.4. A graphical representation of the resulting gain matrix is shown in the spatial model grid (figures 6.5) for decoupling lengths of 50 and 100Km respectively.

$$k_{i,j} = f_{depth}\, f_{dist}\, k_{i,j} \qquad [6.4]$$

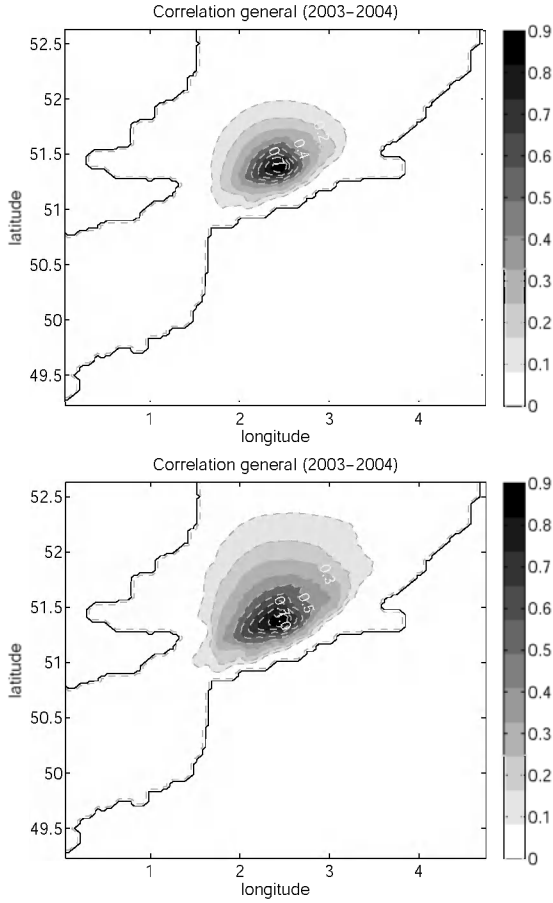

Figure 6.5. Gain matrix structure obtained from the product of the correlation of significant wave height (H_{m0}), wind speed (U_{10}), mean wave period ($T_{m-1,0}$), second moment wave period (T_{m02}), and peak period (T_{peak}), and bathymetry and distance factors, equation 6.4. With decoupling lengths of 50km and 100Km. Correlations are calculated between each point of the Local 2 grid and the assimilation point (WHI) for the two years simulation period of 2003 and 2004.

6.2.4 Numerical Experiments (assimilation of wave height)

In table 6.1 a description is given of the numerical tests performed. The results of the present tests are summarized in table 6.2. Results of the normal run and test C-6b are also included for comparison.

Table 6.1. Description of the numerical tests using a parameterized gain matrix computed from the product of the correlation of the model output variables and including bathymetry and distance factors. The wind factor is 1.1.

Test	Variables used	f_{depth}	f_{dist}	Result
C-7a	H_{m0}, $T_{m-1,0}$, T_p, U_{10}, T_{m02}	-	-	rte*
C-7b	H_{m0}, $T_{m-1,0}$, T_p, U_{10}, T_{m02}	5m	100Km	rte*
C-7c	H_{m0}, $T_{m-1,0}$, T_p, U_{10}, T_{m02}	5m	50Km	rte*
C-7d	H_{m0}, $T_{m-1,0}$, T_p, U_{10}, T_{m02}	15m	100Km	OK
C-7e	H_{m0}, $T_{m-1,0}$, T_p, U_{10}, T_{m02}	15m	50Km	OK
rte*	*Run-time error*			

Table 6.2 Statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder, Bol van Heist, and Oostende, for the normal run, test C-6b and the assimilation tests C-7d and C-7e. RMSE and Bias are given in meters for H_{m0}, and in seconds for $T_{m-1,0}$, SI is given as fraction.

H_{m0} **[m]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C-6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C-7d	0.0000	0.0000	0.0000	0.2704	0.1344	0.2350	0.3824	0.3129	0.3463
C-7e	0.0000	0.0000	0.0000	0.2738	0.1280	0.2379	0.3802	0.3058	0.3443

$T_{m-1,0}$ **[s]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C-6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C-7d	0.3736	0.0009	0.0662	0.8608	0.6802	0.2020	0.6490	0.4602	0.1281
C-7e	0.3655	-0.0159	0.0647	0.8443	0.6633	0.1981	0.6340	0.4387	0.1251

As it can be expected by looking at the dimensions of the assimilation domain calculated from the correlation coefficients only, the simulation C-7a ended with a run-time error. Problems were experienced also from the runs using a bathymetry threshold of 5m. The runs using a bathymetry threshold of 15m were completed.

It can be seen from table 6.2, that in terms of H_{m0}, the results of these two tests are very similar to the results of test C-6b. In terms of $T_{m-1,0}$, the results of tests C-7d and C-7e are slightly worse that those of C6b. That is not surprising since it has been previously noticed (within tests C-6), that the wider the structure of the gain matrix the worse the assimilation results. The assimilation domain in the two present cases is larger than that of test C-6b. There are two positive aspects of the current implementation. The first is that results are similar to those from test C-6b, which shows a consistency of the system. The second is that with this approach the gain matrix structure can be computed in a more flexible way for assimilating data from other buoy locations.

6.3 Parameterization of the gain matrix using short-term model estimates

In the previous approaches, a constant structure was used to represent the errors for all wave conditions. Although long-term statistical realizations provide a good estimate of the correlation structure, the current wave state (and therefore the current representation of the errors) is blurred in the long-term estimates. In this regard, the previous approach to calculate the gain matrix is worked out in order to calculate the gain matrix from the short-term model output. The purpose is to identify grid points with similar spectral characteristics to the assimilation point. The magnitude of the assimilation gain factors is assumed to be proportional to the degree of similarity in wave conditions.

There are two main practical advantages of this approach compared to the previous. The first is that long-term model results are not needed. The second is that this criterion can be programmed inside the model code itself.

A dynamic gain matrix is thus calculated considering the short-term wave conditions. In this approach, the following wave parameters: wave height, mean frequency, second moment wave period and mean direction (H_{m0}, $f_{m-1,0}$, T_{m02}, θ_{m}) are compared between each grid point and the assimilation point (WHI). That information is then used to determine the assimilation gain factors. In order to build an initial criterion for this scheme, wave spectra from several locations around the assimilation point (WHI) are obtained from the normal run. The output points are indicated in the model grid domain in figure 6.6.

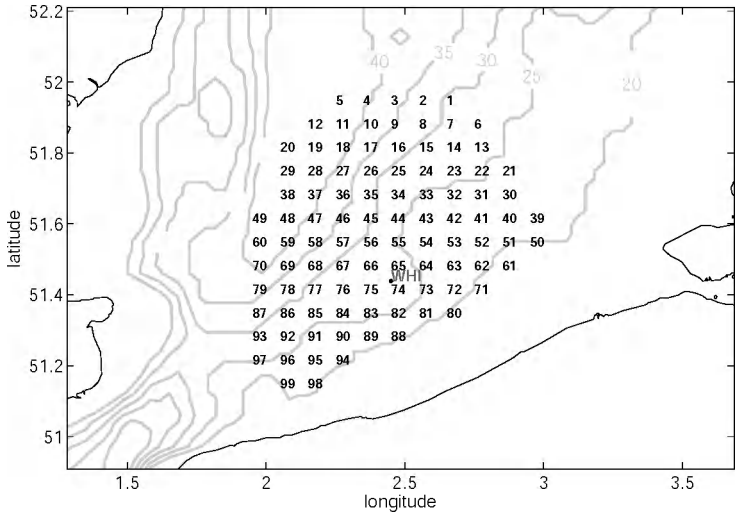

Figure 6.6. Local 2 model grid domain and output points for the present analysis (99 points in the vicinity of WHI).

One-dimensional spectra for the output points are given in figures 6.7, for different times. The spectrum at WHI is plotted with black line. The highest and the lowest energy spectra at every time are also indicated in each plot.

From figures 6.6 (and figure 2.9) it can be seen that wave spectra are more similar to each other in the spatial domain during storm conditions than in low wind conditions. For instance at 18-Jan-2007 12:00:00 and at 21-Jan-2007 00:00:00, the differences in wave parameters are expected to be very low for all these spectra. In those conditions, the dimensions of the assimilation domain should be larger. On the other hand, at periods like 24-Jan-2007 12:00:00 and 25-Jan-2007 00:00:00 wave conditions in the spatial domain differ in larger magnitudes. In those cases, the dimensions of the assimilation domain must be restricted.

This approach is in agreement with the OI approach in the sense that for higher waves, the model and observation variances are higher and the assimilation effect takes place in a larger area. An advantage of the present approach is that several integral parameters (instead of H_{m0} only) are considered.

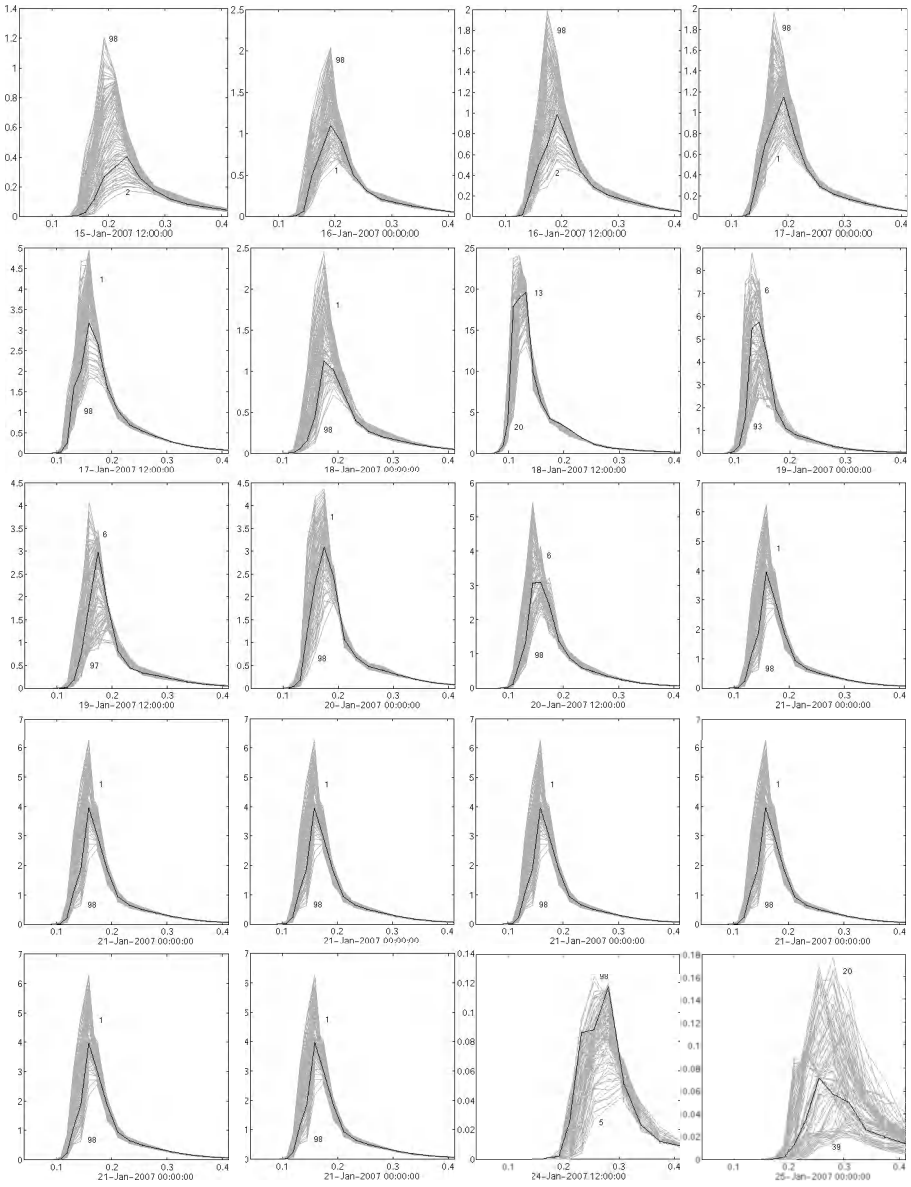

Figure 6.7. Frequency spectra of 99 points in the vicinity of Westhinder (see figure 6.6) at different time steps of the simulation period. The spectrum at WHI is indicated with black line. The two numbers correspond to the point ID of the highest and lowest energy spectra.

6.3.1 Calculation of the assimilation gain matrix

In order built this criterion, the mean parameters: H_{m0}, $f_{m-1,0}$, T_{m02}, and θ_m are considered. The variable $f_{m-1,0}$ has been preferred with respect to $T_{m-1,0}$ because it is calculated directly with an internal model subroutine. The absolute differences of integral parameters (between each grid point and the assimilation point) are stored in temporary matrices. A factor of (no influence of assimilation) is set by defining threshold values for the absolute differences in the integral parameters considered. A value of one is set at the assimilation point (observations correspond to the truth). The factors for the grid points with a difference in integral parameter lower than the threshold value are calculated using a linear scaling of that absolute difference between zero and one. The calculations for this criterion are expressed by equation 6.5.

$$f_X = \max\left\{0, -\frac{1}{X^{th}}\left|X^i - X^j\right| + 1\right\} \qquad [6.5]$$

where f_X is the weighting factor corresponding to the generic integral parameter X, X^i is the value of X at the model grid point i, X^j is the value at the assimilation point, and X^{th} is the threshold value of the variable that specifies a factor of zero. This equation is used to calculate a factor for each of the integral parameters considered. It should be noted that for the mean direction, the absolute difference is found as $\Delta\theta = 180 - |180 - |\theta_1 - \theta_2||$ (Van Vledder, 1993).

Finally, the elements of the gain matrix are found as the product of the different variables factors, equation 6.6.

$$k_{i,j} = f_{H_{m0}} f_{f_{m-1}} f_{T_{m02}} f_{\theta_m} \qquad [6.6]$$

6.3.2 Numerical Experiments (assimilation of wave height)

The threshold values for the numerical experiments using this approach are given in table 6.3.

Note that the threshold values of $f_{m-1,0}$ in the first tests have been taken rather large (1.0Hz) as a consequence $f_{m-1,0}$ does not have influence in the structure of the gain matrix. This is done in order to assess the influence of the other parameters first. The results of these numerical experiments are presented in table 6.4, also results from the previous experiments C-6b, C-6g, and C-6h are shown for reference.

Table 6.3. Numerical tests and threshold values for the computation of the gain matrix according to equations 6.5 and 6.6.

	H_{m0} (WHI)th	$f_{m-1,0}$(WHI)th	T_{m02} (WHI)th	θ_m (WHI)th
S8a	20% H_{m0}	1.0 Hz	2.0s	90°
S8b	20% H_{m0}	1.0 Hz	2.0s	60°
S8c	20% H_{m0}	1.0 Hz	2.0s	30°
S8d	20% H_{m0}	1.0 Hz	2.0s	10°
S8e	20% H_{m0}	0.5 Hz	2.0s	10°
S8f	20% H_{m0}	0.25 Hz	1.0s	10°
S8g	20% H_{m0}	0.12 Hz	0.5s	10°
S8h	20% H_{m0}	0.06 Hz	0.25s	10°
S8i	20% H_{m0}	0.03 Hz	0.10s	10°
S8j	10% H_{m0}	0.03 Hz	0.10s	10°
S8k	5% H_{m0}	0.03 Hz	0.10s	10°

Table 6.4 Statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder, Bol van Heist, and Oostende, for the normal run, tests C-6b, C-6g, C-6h and the present assimilation tests S-8 from a to k. RMSE and Bias are given in meters for H_{m0} and in seconds for $T_{m-1,0}$, SI is given as fraction.

H_{m0} **[m]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6g	0.0000	0.0000	0.0000	0.2283	0.0008	0.1984	0.2587	0.1394	0.2342
C6h	0.4747	-0.3527	0.2370	0.2482	-0.0519	0.2157	0.2548	0.0993	0.2307
S8a	0.0000	0.0000	0.0000	0.2913	0.1563	0.2531	0.3858	0.3067	0.3494
S8b	0.0000	0.0000	0.0000	0.2893	0.1542	0.2514	0.3858	0.3062	0.3494
S8c	0.0000	0.0000	0.0000	0.2815	0.1427	0.2446	0.3846	0.3022	0.3483
S8d	0.0000	0.0000	0.0000	0.2758	0.1172	0.2396	0.3773	0.2852	0.3417
S8e	0.0000	0.0000	0.0000	0.2758	0.1171	0.2396	0.3773	0.2851	0.3417
S8f	0.0000	0.0000	0.0000	0.2745	0.1131	0.2385	0.3757	0.2818	0.3402
S8g	0.0000	0.0000	0.0000	0.2723	0.1049	0.2366	0.3711	0.2753	0.3361
S8h	0.0000	0.0000	0.0000	0.2739	0.0989	0.2380	0.3672	0.2668	0.3326
S8i	0.0000	0.0000	0.0000	0.2764	0.0944	0.2402	0.3677	0.2598	0.3330
S8j	0.0000	0.0000	0.0000	0.2778	0.0931	0.2414	0.3693	0.2581	0.3345
S8k	0.0000	0.0000	0.0000	0.2798	0.0930	0.2432	0.3697	0.2570	0.3348

$T_{m-1,0}$ [s]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6g	0.4245	-0.1656	0.0752	0.7445	0.5725	0.1747	0.5086	0.2641	0.1003
C6h	0.4992	-0.2301	0.0884	0.6233	0.4262	0.1462	0.4875	0.1890	0.0962
S8a	0.4154	0.0085	0.0736	0.8504	0.6800	0.1995	0.6350	0.4354	0.1253
S8b	0.4240	0.0034	0.0751	0.8480	0.6783	0.1990	0.6309	0.4336	0.1245
S8c	0.4270	-0.0065	0.0756	0.8361	0.6690	0.1962	0.6181	0.4221	0.1220
S8d	0.4193	-0.0301	0.0743	0.7965	0.6331	0.1869	0.6049	0.3992	0.1194
S8e	0.4190	-0.0302	0.0742	0.7962	0.6329	0.1868	0.6049	0.3991	0.1194
S8f	0.4163	-0.0339	0.0737	0.7897	0.6251	0.1853	0.6048	0.3922	0.1193
S8g	0.4146	-0.0477	0.0734	0.7796	0.6139	0.1829	0.5988	0.3822	0.1181
S8h	0.4090	-0.0652	0.0724	0.7677	0.6000	0.1801	0.5903	0.3683	0.1165
S8i	0.4175	-0.0925	0.0740	0.7544	0.5855	0.1770	0.5840	0.3532	0.1152
S8j	0.4223	-0.1028	0.0748	0.7460	0.5781	0.1750	0.5821	0.3473	0.1149
S8k	0.4263	-0.1089	0.0755	0.7425	0.5755	0.1742	0.5811	0.3443	0.1147

There are some positive aspects in this approach. The first is that even though the bathymetry was not taken into account to determine the gain matrix structure, shallow water areas are automatically left out of the assimilation domain in nearly all time steps. Consequently, the scheme is able to run without experiencing run-time errors even with relatively large values of the threshold parameters, which allow a bigger assimilation domain.

Regarding the statistical parameters, they are similar to those from the previous experiments. In addition, the swell corrections (not shown) do also take place. These facts confirm to some extent the consistency of this approach.

6.4 Assimilation of mean wave period

An advantage of disposing of spectral wave measurements for data assimilation is that other wave parameters can be assimilated in addition to significant wave height. The assimilation of mean wave period ($T_{m-1,0}$) is specially an interesting parameter since by correcting mean wave period, not only the wave energy is made to agree

better with observations but also the distribution of energy density in the frequency domain can be improved.

It should be taken into account that the assimilation of mean wave period demands the identification of distinct wave systems wind-sea and swell, since the relationship between wave energy and wave period is different for these two wave systems.

6.4.1 Retrieval of analyzed wave height and wave period

In order to retrieve analyzed wave height and period, not only a rescaling strategy for updating energy is necessary, but also the spectrum has to be shifted in the frequency domain. Equation 4.34, recalled as equation 6.7.

$$F(f,\theta)_a = \alpha\, F(\beta f,\theta)_b \qquad [6.7]$$

This operation must be constrained by physical considerations of the dynamical system, typically wind-sea and swell systems are considered. To this end, Lionello et al. (1992) proposed the following approach. In the case of wind-sea, fetch laws for time-limited conditions are assumed. Accordingly, the energy rescaling factor and the frequency shift can be computed according to equations 4.35, recalled as equation 6.8.

wind-sea:
$$\alpha = \frac{E_a}{E_b}\beta, \qquad \beta = \frac{f_{mb}}{f_{ma}} \qquad [6.8]$$

In the case of swell the wave steepness is assumed to remain constant in the decay period, equation 4.36, 6.9.

swell:
$$\alpha = \frac{E_a}{E_b}\beta, \qquad \beta = \Delta\left(\frac{E_a}{E_b}\right)^{1/4} \qquad [6.9]$$

Although consistent algorithms for identifying wind-sea and swell were developed in Chapter 3, the implementation of these algorithms could not be integrated in the assimilation scheme. The main difficulty is that not only wave systems need to be identified but also wave systems from different points and at different time steps need to be cross-assigned. Preliminary implementations (see for instance Portilla and Monbaliu 2006) showed that this cross-assignment step is not trivial, since wave fields of wind-sea and swell need to be smooth in space and time in order to avoid model instability. Therefore, further investigation is needed in order to obtain a more robust cross-assignment algorithm.

Nonetheless, numerical experiments for assimilating wave period are carried out using a simpler approach proposed by Lionello et al. (1992). In this approach, wind-

sea is defined as the wave state in which energy is growing and swell as the state when energy decays. Note that using this approach spectral partitioning is not carried out. The whole spectrum is considered either wind-sea or swell.

6.4.2 Numerical experiments (wave height and wave period)

The gain matrix for the present experiments corresponds to the two-dimensional Gaussian function (equation 5.2) used for the experiments C6b and C6g. The following tests have been performed in order to assess the impact of assimilating both H_{m0} and $T_{m-1,0}$.

Table 6.5. Numerical experiments for the assimilation of wave height and wave period and parameters of the two-dimensional Gaussian function for the gain matrix.

	σ_x[Km]	σ_y[Km]	θ [º]	U_{10} *factor*	β *factor**
C9a	50	10	-	1.1	Wind-sea only
C9b	50	10	20	1.0	Wind-sea only
C9c	50	10	20	1.1	Wind-sea – swell
C9d	50	10	20	1.0	Wind-sea – swell
* strategy for the computation of the shift frequency factor					

At the assimilation point (WHI), wave height and wave period do not completely agree with the observations as in the previous experiments where only wave height was assimilated. This becomes evident when looking at equations 6.8 and 6.9, where the energy scaling factor depends on the frequency shift factor. Since β is in general different than 1, α does not impose the model energy to match the observed value.

Results from test C9a are very similar to those from C6b. Only a slight improvement is apparent from test C9a. That is the case also for test C9c. The effect of introducing two update strategies (wind-sea and swell) does not show a clear positive impact. The same is concluded from the runs using a factor of 1.0 for the wind speeds. The trends are in general the same, the results of tests C9b are very similar to those from test C6g, but a clear improvement is not evident.

From tests C9a and C9c, compared to the normal run and to test C6b, and from tests C9b and C9d, compared to C6h (normal run with 1.0 wind factor) and C6g, it can be said that better assimilation results are obtained when the model results themselves are worse. Actually, the results of the run C6h are remarkably good and little improvement can be achieved from data assimilation.

Table 6.6 Statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder, Bol van Heist, and Oostende, for the normal run, tests C-6b, C-6g, C-6h and the present assimilation tests using H_{m0} and $T_{m-1,0}$. The tests with wind factor 1.0 have been highlighted. RMSE and Bias are given in meters for H_{m0} and in seconds for $T_{m-1,0}$. SI is given as fraction.

H_{m0} **[m]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6g	0.0000	0.0000	0.0000	0.2283	0.0008	0.1984	0.2587	0.1394	0.2342
C6h	0.4747	-0.3527	0.2370	0.2482	-0.0519	0.2157	0.2548	0.0993	0.2307
C9a	0.0624	-0.0321	0.0311	0.2795	0.0988	0.2429	0.3710	0.2675	0.3360
C9b	0.0755	-0.0449	0.0377	0.2329	-0.0162	0.2024	0.2560	0.1317	0.2318
C9c	0.0538	-0.0208	0.0269	0.2811	0.1001	0.2443	0.3718	0.2675	0.3367
C9d	0.0627	-0.0223	0.0313	0.2321	-0.0115	0.2017	0.2560	0.1329	0.2319

$T_{m-1,0}$ **[s]**

	WHI			BVH			ONS		
	RMSE	**Bias**	**SI**	**RMSE**	**Bias**	**SI**	**RMSE**	**Bias**	**SI**
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6g	0.4245	-0.1656	0.0752	0.7445	0.5725	0.1747	0.5086	0.2641	0.1003
C6h	0.4992	-0.2301	0.0884	0.6233	0.4262	0.1462	0.4875	0.1890	0.0962
C9a	0.0198	-0.0151	0.0035	0.7619	0.5988	0.1788	0.5855	0.3647	0.1155
C9b	0.0206	-0.0152	0.0036	0.7081	0.5476	0.1661	0.5027	0.2656	0.0992
C9c	0.3560	-0.1217	0.0631	0.7618	0.5951	0.1787	0.5854	0.3551	0.1155
C9d	0.4742	-0.2561	0.0840	0.7026	0.5349	0.1648	0.4948	0.2430	0.0976

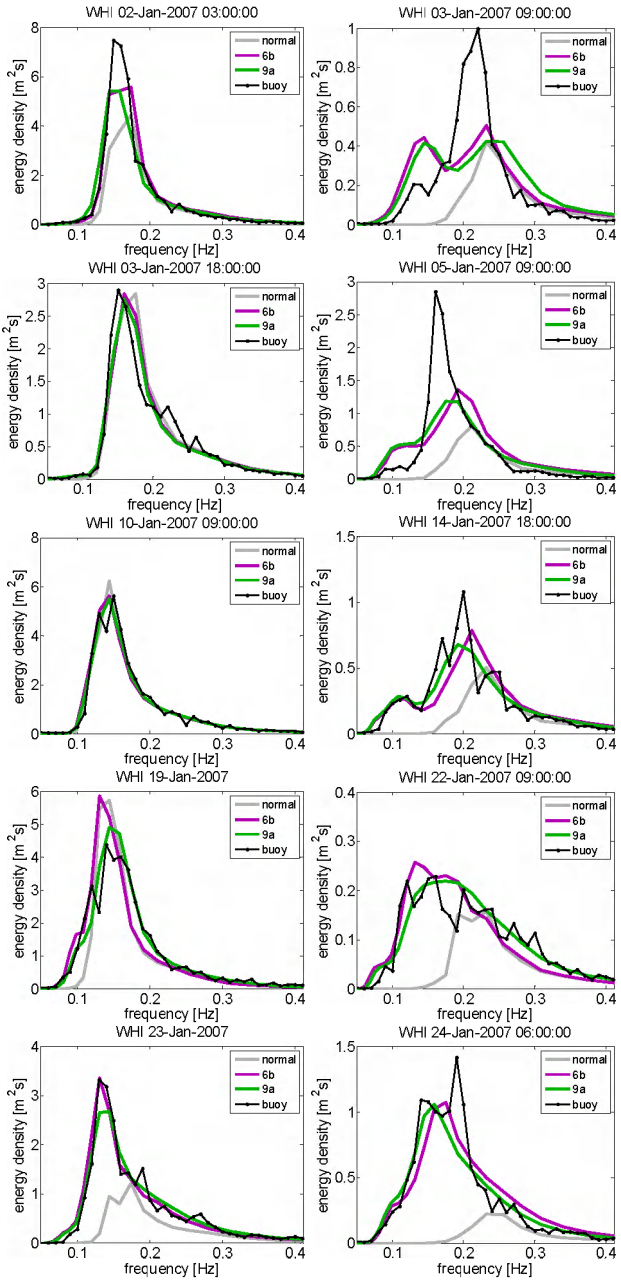

Figure 6.8. Wave spectra for the normal run, the test C6b assimilating wave height only, and test C9a assimilating wave height and wave period.

It can be seen that accounting for wind-sea and swell in the computation of the shift frequency, at the assimilation location (WHI), a lower agreement is obtained in $T_{m-1,0}$ between model and buoy compared to the experiments where only the wind-sea strategy is used. In this last case, the model is forced to match to a larger extent the observed conditions, both in H_{m0} and $T_{m-1,0}$. It can be said in general that the assimilation of $T_{m-1,0}$ did not improve considerably the estimates compared to the normal runs using the simple approach to account for different wave systems.

Additionally, a similar modelling behaviour as with the assimilation of wave height only occurs also in this case. $T_{m-1,0}$ is underestimated at WHI (assimilation location) while at the coastal locations $T_{m-1,0}$ is overestimated. Therefore, it is very difficult by the assimilation of data at WHI to correct this effect.

In the experiments where only significant wave height was assimilated, it could be seen that due to model dynamics, the energy density is considerably redistributed in the frequency domain and consequently the modelling of low frequency waves was improved. Due to this redistribution of energy, other mean wave parameters were also affected, mean period for instance was improved during the periods where wave height was improved.

A similar conclusion can be derived by looking at the spectra (figures 6.8). In situations where the model performance is good, the spectral corrections carried out by the assimilation of wave height only (C6b) are minor (see for instance spectra at 3-Jan-2007 18:00:00). In situations where normal model results are deficient, the assimilation of wave height only produced a corrected spectral shape and consequently the mean wave period is also corrected (see for instance spectra at 24-Jan-2007 06:00:00). Further corrections to the wave period (by assimilating wave period) are therefore limited and do not show a large impact.

6.5 Experimental assessment of the observations uncertainty

One of the simplifications made to the assimilation system in the previous sections was to consider measurements to correspond to the true value of the variables. In reality, buoy data do also contain errors. The introduction of those errors into the system might deteriorate the results, especially in situations where model estimates are good. In this section, numerical experiments are carried out in order to evaluate the effect of accounting for uncertainties in the observations.

The account for observations uncertainties can be done by introducing a factor (lower than one) into the assimilation gain matrix. That is, the higher the uncertainty assumed for the observations the lower the factor.

6.5.1 Numerical experiments

In the present experiments the assimilation gain matrix corresponds to the two-dimensional Gaussian function used in the previous experiments C6b and C6g (equation 5.2). The assimilation parameter is wave height only. In order to simulate the observations uncertainty, the gain matrix is multiplied by a factor A, lower than 1.0. The following tests are considered.

Table 6.7. Numerical experiments for the assimilation of wave height accounting for observations uncertainties and parameters of the two-dimensional Gaussian function (equation 5.2) used for the gain matrix.

	σ_x[Km]	σ_y[Km]	θ [º]	U_{10} *factor*	*A**
C10a0	50	10	20	1.1	0.9
C10b0	50	10	20	1.1	0.8
C10c0	50	10	20	1.1	0.7
C10d0	50	10	20	1.1	0.6
C10e0	50	10	20	1.1	0.5
C10a	50	10	20	1.0	0.9
C10b	50	10	20	1.0	0.8
C10c	50	10	20	1.0	0.7
C10d	50	10	20	1.0	0.6
C10e	50	10	20	1.0	0.5
* factor accounting for observations uncertainties					

Experiments C6b and C6g serve as reference here because the assimilation matrix is the same and the observation uncertainty factor is 1. In the case of the experiments with wind factor of 1.0, there is a general trend of the system to return to the situation where no assimilation occurs (C6h), more or less proportionally as the weight of observations is lowered. It should be noted that the reduction of the wind factor produced considerable improvement of the wave estimates and consequently the impact of assimilation in those conditions was negative (i.e., experiment C6g with respect to C6h).

In the case of experiment C6b with respect to the normal run, some of the statistical parameters were improved (for instance RMSE and SI at BVH for H_{m0}, and RMSE and SI at ONS for $T_{m-1,0}$). Other parameters deteriorated. However, a similar trend of the system, to return to the state where no assimilation occurs, is also observed. See for instance H_{m0} at ONS, the three statistical parameters.

Table 6.8 Statistical parameters for WHI, BVH, and ONS, for the normal run, tests C-6b, C-6g, C-6h and tests evaluating observations uncertainties. RMSE and Bias are in meters for H_{m0} and in seconds for $T_{m-1,0}$, SI is given as fraction.

H_{m0} [m]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6g	0.0000	0.0000	0.0000	0.2283	0.0008	0.1984	0.2587	0.1394	0.2342
C6h	0.4747	-0.3527	0.2370	0.2482	-0.0519	0.2157	0.2548	0.0993	0.2307
C10a0	0.0185	-0.0043	0.0092	0.2821	0.1023	0.2451	0.3734	0.2690	0.3381
C10b0	0.0206	-0.0077	0.0103	0.2820	0.1016	0.2450	0.3732	0.2683	0.3380
C10c0	0.0289	-0.0119	0.0144	0.2819	0.1008	0.2450	0.3730	0.2676	0.3378
C10d0	0.0425	-0.0172	0.0212	0.2819	0.0998	0.2449	0.3728	0.2667	0.3376
C10e0	0.0611	-0.0240	0.0305	0.2818	0.0985	0.2449	0.3726	0.2656	0.3374
C10a	0.0143	-0.0020	0.0072	0.2328	-0.0020	0.2023	0.2583	0.1359	0.2339
C10b	0.0222	-0.0113	0.0111	0.2329	-0.0041	0.2024	0.2578	0.1344	0.2335
C10c	0.0363	-0.0227	0.0181	0.2331	-0.0065	0.2025	0.2572	0.1327	0.2329
C10d	0.0555	-0.0369	0.0277	0.2333	-0.0095	0.2028	0.2566	0.1307	0.2324
C10e	0.0805	-0.0552	0.0402	0.2338	-0.0132	0.2031	0.2559	0.1281	0.2318

$T_{m-1,0}$ [s]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6g	0.4245	-0.1656	0.0752	0.7445	0.5725	0.1747	0.5086	0.2641	0.1003
C6h	0.4992	-0.2301	0.0884	0.6233	0.4262	0.1462	0.4875	0.1890	0.0962
C10a0	0.4546	-0.1557	0.0805	0.7696	0.5989	0.1806	0.5902	0.3559	0.1165
C10b0	0.4516	-0.1541	0.0800	0.7677	0.5973	0.1801	0.5899	0.3552	0.1164
C10c0	0.4483	-0.1524	0.0794	0.7655	0.5955	0.1796	0.5896	0.3543	0.1163
C10d0	0.4442	-0.1500	0.0787	0.7630	0.5934	0.1790	0.5892	0.3532	0.1162
C10e0	0.4389	-0.1470	0.0777	0.7599	0.5907	0.1783	0.5886	0.3519	0.1161
C10a	0.5724	-0.3487	0.1014	0.7294	0.5517	0.1712	0.5012	0.2432	0.0989
C10b	0.5697	-0.3469	0.1009	0.7245	0.5470	0.1700	0.5005	0.2412	0.0988
C10c	0.5662	-0.3445	0.1003	0.7185	0.5414	0.1686	0.4997	0.2388	0.0986
C10d	0.5621	-0.3417	0.0996	0.7111	0.5344	0.1669	0.4987	0.2358	0.0984
C10e	0.5564	-0.3377	0.0986	0.7023	0.5259	0.1648	0.4975	0.2321	0.0982

That trend is not evident for other parameters like for instance H_{m0} and $T_{m-1,0}$ at BVH and $T_{m-1,0}$ at ONS. When the uncertainty factor is lowered from 1.0 to 0.9, parameters like RMSE and SI increase but as the amplitude gets lower than 0.9 these parameters do not increase proportional to the value of the normal run. Contrarily they decrease smoothly.

The main purpose of these experiments was to evaluate empirically the relative contribution of observation errors into the assimilation system. However, that assessment cannot be clearly done from the present results since the trends in the different experiments do not point in a particular direction. However, the fact that lowering the weight of observations produces a positive impact in some parameters is an indication that in some situations the assimilation system benefits from this reduction. Further analysis is needed in order to determine the situations in which improvements occur.

6.6 Summary and Conclusions

In this chapter, numerical experiments were carried out in order to evaluate complementary aspects of the assimilation scheme tested in Chapter 5. One of the limitations of the two-dimensional Gaussian function implemented before is the low transferability to assimilate data from other buoy locations. Two strategies were tested in order to overcome this limitation. In the first, the spatial correlation of the model variables was calculated from long-term model results. In the second strategy, short-term differences of integral parameters between the grid points and the assimilation point were used to parameterize the gain matrix. Additionally, numerical experiments were carried out to evaluate the impact of including wave period in the assimilation. Finally, the introduction of observation uncertainties was evaluated experimentally by introducing a factor in the assimilation gain matrix that reduces the weight of observations.

The spatial structure of the correlations for the different wave variables suggests that anisotropic conditions predominate in the area. Moreover, the shapes of those structures as such agree well with the two-dimensional Gaussian shape assumed in the experiments of the previous chapter. This result adds confidence to the consistency of those previous results. The relative values of the correlations are however different and according to the previous tests they are rather high. Especially correlation values of significant wave height are the largest amongst all the wave variables. This is remarkable since typically, the study of the errors of the system is based on this variable only. It became evident that using the correlation values as such to represent the gain matrix would not be possible due to the large domain

dimensions. Therefore, a parameterized gain matrix was calculated using those correlations. In this matrix, however the coastal zone was still included in the assimilation domain so other factors accounting for bathymetric characteristics and distance had to be introduced to force a decoupling of shallow water areas from the assimilation domain.

In the second strategy, this idea was extended in order to calculate a gain matrix based on short-term model estimates, considering the differences of integral parameters between the grid points and the assimilation point. Advantages of this approach are that the subroutines for calculating the gain matrix can be programmed inside the model code itself. Besides long-term model simulations are not needed. In addition, the dimensions of the assimilation domain can be adjusted dynamically depending on the similarity of wave conditions in the area.

The results obtained using these two strategies are similar to those obtained in the previous chapter using a two-dimensional Gaussian function. The advantage of the last two, is the more flexible adaptation of the assimilation algorithm (and therefore transferability of the system) for assimilating data from other buoy locations.

The assimilation of wave period did not show a clear impact in the wave conditions compared to the runs where only significant wave height was assimilated. One of the main reasons refers to the positive interaction between the assimilation scheme and the wave model pointed out before. Due to model dynamics, the wave spectrum is changed when energy update occurs and other integral parameters are therefore affected. Particularly the influence in the wave period was positive by assimilating wave height and further improvement was difficult to obtain. In addition, a similar limitation as in the case of wave height takes place. By assimilating wave period at WHI waves are brought from higher to lower frequencies, however the requirements for correction at the coastal locations (BVH and ONS) are the opposite. The simplifications of the retrieval scheme can play an important role as well. Further investigation is necessary to optimize the assimilation of the two parameters.

Regarding the evaluation of the observations uncertainties, consistency in the results was observed in the sense that by lowering the weight of the observations the system tends to return to the situation where no assimilation occurs. However, a quantitative assessment can not be clearly done from the present results since the trends from the different experiments do not point in a particular direction. Some parameters are positively affected by the reduced influence of the observations, which is an indication that in particular situations the system benefits from this reduction. Further analysis is needed in order to determine the situations in which improvements occur.

Chapter 7

Summary, Conclusions, and Further work

7.1 Summary

An Optimal Interpolation scheme was implemented in the WAM model in order to assimilate buoy wave data in a nearshore scenario, namely at the Belgian coast in the southern North Sea. Using this methodology, buoy data from a single offshore buoy position (location Westhinder) was introduced into the model grid computations. The impact of data assimilation was evaluated at nearshore locations (locations Bol Van Heist and Oostende). The aim of the assimilation exercise is to improve wave estimates in hindcast mode along the coast. The target application of this methodology is improved estimation of long-term statistics for wave climate.

Emphasis was given to the separation of wind sea and swell in the hope of building up a more robust algorithm for the retrieval of analyzed variables into the wave spectrum. The advantage of identifying wave systems is that corrections can be applied according to the different physical characteristics. Several spectral partitioning and wind-sea / swell identification schemes were revisited. The algorithms introduced here aim at obtaining more consistent results for determining spectral partitions and identifying wind-sea and swell. Having consistent wave systems is important in data assimilation as the benefit of the separation can only be realized when it does not imply a reduction of accuracy. However, the integration of these procedures into the assimilation scheme still requires the development of a specialized cross-assignment algorithm to match buoy and model partitions. Nonetheless, partitioning and identification procedures can prove useful in a number of other applications for data analysis and model evaluation (e.g., Portilla and Monbaliu 2007).

Data assimilation experiments were carried out in order to study the behaviour of the assimilation scheme and the impact of assimilating buoy data. Namely, the

assimilation of both significant wave height and mean wave period were investigated. In view of the limited knowledge of the errors characteristics in the study area, the assimilation gain matrix has been approximated using different functions. The results of the numerical experiments are promising and show several benefits of data assimilation nearshore. The effect is especially interesting in complicated modelling situations where wind-sea and swell systems coexist. The analysis of model results also benefited from the assimilation exercise as some modelling shortcomings could be better understood in the light of the assimilation results. In this sense, modelling development and data assimilation have to be regarded as complementary activities.

7.2 Conclusions

From the investigation of the different spectral partitioning techniques it is emphasized that the different existing methods differ mainly in the way they assess whether partitions are significant or not. It was found that the reported mechanisms are not very robust and give inconsistent results. An improvement for the 2D partitioning scheme is proposed using a 2D low-pass filter, which reduces spectral noise. For 1D spectra a combining criterion has been proposed in order to remove the most obvious spurious peaks.

Results from different methods to identify wind-sea and swell deviate largely. It is concluded that the identification of wind-sea and swell both in the 2D and 1D spectra is more consistent in combination with partitioning. For 2D spectra, the combination of a wind wave generation mechanism and 2D partitioning was found to provide the most consistent estimates. For 1D spectra, a method has been proposed looking at the ratio (γ^*) between the energy at the spectral peak of a partition and the energy at the peak of a PM-spectrum with the same peak frequency.

From the overview of data assimilation schemes, it is noted that different methodologies have been applied in previous studies, each with their own inherent advantages and disadvantages. It is also clear that further efforts are needed in order to cope with algorithm simplifications, excessive computational demands, and other practical limitations. From a statistical point of view, Kalman filtering techniques are favoured for the ability to determine explicitly the error covariance matrix, which is the core of the assimilation problem. On the other hand, variational approaches are very attractive as they operate in the model control variables and not in the model output. In that way, variational methods are able to produce results consistent with the model dynamics. However, the practical application of these two methodologies does still require further development. At present, the application of

OI techniques is very attractive, in view of its low computational cost and robust conceptualization.

In data assimilation applications, knowledge of errors from the modelling system and their spatial structure is very important. However, it is noted that in general that knowledge is very poor. For the implementation of OI schemes, that information is particularly important since the analysis relies on its a-priori specification. In variational schemes, the system errors constitute the weight factors in the construction of the cost function. In Kalman filtering the initialization is based on the specification of the errors as well. Moreover, knowledge of the errors is important for evaluating assimilation systems. Therefore, this aspect deserves more attention for further developing and implementing assimilation techniques.

The implementation of the OI scheme in the nearshore scenario poses some practical and conceptual issues. In deep water, statistical conditions as homogeneity and isotropy can be assumed for the error correlation structure. However, in the nearshore scenario the relevant processes involved in wave evolution differ over a very short spatial range since the background physical processes within the domain are different. Therefore, homogeneity and isotropy conditions do not hold. This is evidenced not only by poor assimilation results but also by model instability.

Different approaches were tested to confront this limitation. In the first approach a 2D Gaussian function was used to approximate the gain matrix. In the second, long-term model estimates were used. The third approach is an adaptation of the second, using short-term model estimates. Results from the three approaches are similar. The advantage of the last two, is the more flexible adaption of the assimilation algorithm (and therefore transferability of the system) to other buoy locations.

The numerical experiments carried out here show that it is feasible to improve wave estimates nearshore by data assimilation. Specially, statistical parameters like scatter index and root mean square error show the ability of assimilation to perform corrections. The apparent deterioration of bias, the other parameter observed, is not discouraging because the low bias from the normal model run occurs at the price of a large scatter. In the assimilation runs, a tendency towards overestimation of wave height and period is observed but the modelling trend for the different locations is consistent. This is a positive aspect, for instance, for model calibration.

The most important corrections from data assimilation occur in situations of low model performance, namely in moderated wind conditions with the presence of swell. It has been observed that large (white-capping) energy dissipation is triggered by a relatively small increase in wind speed. This large dissipation prevents the

formation of swell after storm periods. Accordingly, the formation of wind-sea (from subsequent wind activity events) is also hindered. This energy dissipation could be compensated with data assimilation. As a consequence, not only a better overall energy agreement is observed but energy at the low frequencies is recovered. This model effect is less visible in strong wind conditions since the exponential energy growth under wind quickly overrides the large dissipation effect.

When moderated wind conditions persist, the effect of the assimilated data last for several days. Although this long duration might seem inconsistent considering the travel time of waves in the domain, which is of the order of hours, this effect can be explained because the improved wave conditions, and specially the presence of low frequency wave energy in the spectrum, act as improved initial conditions for the development of waves in subsequent periods. When low frequency waves are present in the system the energy transfer from higher to lower frequencies (due to quadruplet interactions), is enhanced. Consequently, the energy transfer from wind to waves is also enhanced. It should be noted that this effect has interesting implications for forecasting applications.

In wind-sea conditions, the impact of assimilating data is less visible. First of all, in those cases the model estimates are better. Additionally, the model performance statistics behave differently at offshore and at coastal locations. Significant wave height and mean wave period are underestimated at offshore locations, while those variables are either less underestimated or overestimated at the coastal locations. Under these conditions, wave estimates cannot be improved by single point data assimilation because the assimilation requirements at different locations are conflicting.

The assimilation of both wave height and wave period did not show an advantage over the assimilation of wave height only. This is partially because the assimilation of wave height has the ability to correct the wave spectrum and therefore it affects other mean parameters as well. However, the limitations of the algorithm should also be taken into account. For the case of wind-sea, the frequency shift is rather straightforward since wave height and wave period can be related to each other using wave growth curves. For the case of swell, correlation between height and period are less clear and the relations proposed in the literature are rather hypothetical. Therefore, the swell update is made on more empirical bases.

A similar remark applies for the assimilation of different wave systems (wind-sea and swell). A complicating factor is that more variables have to be imposed into the wave spectrum (i.e., wave height and period for both wind-sea and swell). Moreover, the two wave systems require different specifications for the errors.

Attention should be given to a number of modelling aspects. Several authors, for instance, have pointed out the deficiency of the white-capping formulation in conditions of coexisting wind-sea and swell, and alternative formulations have been proposed (e.g., Van Vledder and Hurdle 2002, Rogers et al. 2003, Van der Westhuysen 2007). The inconsistent behaviour of the model output statistics, points at deficiencies in the model system to transfer wave energy from offshore to nearshore.

7.3 Further work

For the implementation and evaluation of the assimilation system the characteristics of both the background and observation errors of the system need to be studied. In this regard, the present study area is favoured by the abundance of observations available.

The present modelling setup should be reconsidered in order to overcome modelling limitations that can be corrected by model improvements apart from the data assimilation framework. Attention should be given to the model white-capping formulation, to model parameters (e.g., wind factor), to taking into account shallow water processes that can play an important role (e.g., depth-induced breaking, water depth variations and currents, …), and to the use of a higher spatial and spectral resolution model grid.

The impact of data assimilation can be assessed using a coastal model implementation (e.g. SWAN) nested in the analyzed Local 2 grid. The advantage is that the boundary conditions passed from the Local 2 grid are corrected by data assimilation. The coastal model then allows more flexibility to account for other physical mechanisms and modelling options.

Data assimilation experiments should be carried out to evaluate the effect of assimilating data from more that one source point. Under certain meteorological conditions, waves at coastal locations (e.g., Bol Van Heist, Oostende) are more related to wave conditions in the east flank of the domain (e.g., Deurloo, Schouwenbank). The assimilation of data from two (or more) off-shore buoys is expected to have a positive impact on the wave parameter estimates nearshore.

Further experiments are necessary to assimilate several wave parameters. Within the present experiments, the effect of assimilating mean wave period was not large. Analyzed wave fields were retrieved into the spectrum by rescaling all frequency

components. Other retrieval functions need to be tested in order to optimize the impact of assimilation.

For the assimilation of distinct wave systems (wind-sea and swell) a cross-assignment algorithm needs to be developed. The two main tasks of this algorithm are: i) to match buoy and model partitions at a particular instant in time, and ii) to track model and buoy partitions at consecutive time steps. A basic requirement is that the wave systems have to be smooth in space and time.

Long-term assimilation experiments are necessary to evaluate the impact of the assimilation system. The assimilation scheme tested has a bigger influence in conditions where wind-sea and swell waves coexist. It is important to assess how these improvements affect the overall statistics.

Appendix A: Summary of numerical experiments

Table A.1. Summary of numerical experiments.

	Wind factor	Amplitude factor	Assimilation parameter	Gain matrix parameterization
Normal	1.1	-	No assimilation	-
C5b	1.1	1.0	H_s only	Homogeneous-isotropic
C6a	1.1	1.0	H_s only	2D Gaussian: σ_x= 50, σ_y= 25, θ= 20°
C6b	1.1	1.0	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, θ= 20°
C6c	1.1	1.0	H_s only	2D Gaussian: σ_x= 100, σ_y= 25, θ= 20°
C6d	1.1	1.0	H_s only	2D Gaussian: σ_x= 100, σ_y= 10, θ= 20°
C6e	1.1	1.0	H_s only	2D Gaussian: σ_x= 150, σ_y= 25, θ= 20°
C6f	1.1	1.0	H_s only	2D Gaussian: σ_x= 150, σ_y= 10, θ= 20°
C6g	1.0	1.0	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, θ= 20°
C6h	1.0	-	No assimilation	-
C7d	1.1	1.0	H_s only	Correlation of H_s, T_{m01}, Tp, $U_{10,}$ and T_{m02}, f_{depth}= 15m, f_{dist}= 100Km
C7e	1.1	1.0	H_s only	Correlation of H_s, T_{m01}, T_p, $U_{10,}$ and T_{m02}, f_{depth}= 15m, f_{dist}= 50Km
S8a		1.0	H_s only	Similarity of H_s= 20%, Δf= 1.0Hz, ΔT_{m02}= 2.0s, $\Delta\theta$= 90°
S8b		1.0	H_s only	Similarity of H_s= 20%, Δf= 1.0Hz, Δ_{Tm02}= 2.0s, $\Delta\theta$= 60°
S8c		1.0	H_s only	Similarity of H_s= 20%, Δf= 1.0Hz, ΔT_{m02}= 2.0s, $\Delta\theta$= 30°
S8d		1.0	H_s only	Similarity of H_s= 20%, Δf= 1.0Hz, ΔT_{m02}= 2.0s, $\Delta\theta$= 10°
S8e		1.0	H_s only	Similarity of H_s= 20%, Δf= 0.5Hz, ΔT_{m02}= 2.0s, $\Delta\theta$= 10°
S8f		1.0	H_s only	Similarity of H_s= 20%, Δf= 0.25Hz, ΔT_{m02}= 1.0s, $\Delta\theta$= 10°
S8g		1.0	H_s only	Similarity of H_s= 20%, Δf= 0.12Hz, ΔT_{m02}= 0.5s, $\Delta\theta$= 10°
S8h		1.0	H_s only	Similarity of H_s= 20%, Δf= 0.06Hz, ΔT_{m02}= 0.25s, $\Delta\theta$= 10°

S8i		1.0	H_s only	Similarity of H_s= 20%, Δf= 0.03Hz, ΔT_{m02}= 0.1s, $\Delta\theta$= 10°
S8j		1.0	H_s only	Similarity of H_s= 10%, Δf= 0.03Hz, ΔT_{m02}= 0.1s, $\Delta\theta$= 10°
S8k		1.0	H_s only	Similarity of H_s= 5%, Δf= 0.03Hz, ΔT_{m02}= 0.1s, $\Delta\theta$= 10°
C9a	1.1	1.0	H_s & $T_{m-1,0}$ Wind-sea only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C9b	1.0	1.0	H_s & $T_{m-1,0}$ Wind-sea only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C9c	1.1	1.0	H_s & $T_{m-1,0}$ Wind-sea and swell	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C9d	1.0	1.0	H_s & $T_{m-1,0}$ Wind-sea and swell	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10a0	1.0	0.9	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10b0	1.0	0.8	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10c0	1.0	0.7	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10d0	1.0	0.6	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10e0	1.0	0.5	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10a	1.1	0.9	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10b	1.1	0.8	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10c	1.1	0.7	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10d	1.1	0.6	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°
C10e	1.1	0.5	H_s only	2D Gaussian: σ_x= 50, σ_y= 10, □= 20°

Table A.2. Summary of statistical parameters for H_{m0} and $T_{m-1,0}$ for the locations of Westhinder, Bol van Heist, and Oostende, for the data assimilation experiments. RMSE and bias are given for H_{m0} in meters and for T_{m0-1} in seconds, and SI is given as fraction.

H_{m0} [m]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.4633	-0.3048	0.2313	0.3211	-0.0189	0.2790	0.3582	0.1395	0.3244
C5b	0.0000	0.0000	0.0000	0.2818	0.1264	0.2449	0.3947	0.3209	0.3575
C6a	0.0000	0.0000	0.0000	0.2795	0.1163	0.2429	0.3779	0.2917	0.3423
C6b	0.0000	0.0000	0.0000	0.2797	0.1046	0.2430	0.3737	0.2714	0.3385
C6c	0.0000	0.0000	0.0000	0.2691	0.1271	0.2338	0.3801	0.3070	0.3442
C6d	0.0000	0.0000	0.0000	0.2742	0.1077	0.2383	0.3747	0.2831	0.3394
C6e	0.0000	0.0000	0.0000	0.2650	0.1339	0.2303	0.3814	0.3119	0.3454
C6f	0.0000	0.0000	0.0000	0.2716	0.1108	0.2360	0.3754	0.2869	0.3399
C6g	0.0000	0.0000	0.0000	0.2283	0.0008	0.1984	0.2587	0.1394	0.2342
C6h	0.4747	-0.3527	0.2370	0.2482	-0.0519	0.2157	0.2548	0.0993	0.2307
C7d	0.0000	0.0000	0.0000	0.2704	0.1344	0.2350	0.3824	0.3129	0.3463
C7e	0.0000	0.0000	0.0000	0.2738	0.1280	0.2379	0.3802	0.3058	0.3443
S8a	0.0000	0.0000	0.0000	0.2913	0.1563	0.2531	0.3858	0.3067	0.3494
S8b	0.0000	0.0000	0.0000	0.2893	0.1542	0.2514	0.3858	0.3062	0.3494
S8c	0.0000	0.0000	0.0000	0.2815	0.1427	0.2446	0.3846	0.3022	0.3483
S8d	0.0000	0.0000	0.0000	0.2758	0.1172	0.2396	0.3773	0.2852	0.3417
S8e	0.0000	0.0000	0.0000	0.2758	0.1171	0.2396	0.3773	0.2851	0.3417
S8f	0.0000	0.0000	0.0000	0.2745	0.1131	0.2385	0.3757	0.2818	0.3402
S8g	0.0000	0.0000	0.0000	0.2723	0.1049	0.2366	0.3711	0.2753	0.3361
S8h	0.0000	0.0000	0.0000	0.2739	0.0989	0.2380	0.3672	0.2668	0.3326
S8i	0.0000	0.0000	0.0000	0.2764	0.0944	0.2402	0.3677	0.2598	0.3330
S8j	0.0000	0.0000	0.0000	0.2778	0.0931	0.2414	0.3693	0.2581	0.3345
S8k	0.0000	0.0000	0.0000	0.2798	0.0930	0.2432	0.3697	0.2570	0.3348
C9a	0.0624	-0.0321	0.0311	0.2795	0.0988	0.2429	0.3710	0.2675	0.3360
C9b	0.0755	-0.0449	0.0377	0.2329	-0.0162	0.2024	0.2560	0.1317	0.2318
C9c	0.0538	-0.0208	0.0269	0.2811	0.1001	0.2443	0.3718	0.2675	0.3367
C9d	0.0627	-0.0223	0.0313	0.2321	-0.0115	0.2017	0.2560	0.1329	0.2319
C10a0	0.0185	-0.0043	0.0092	0.2821	0.1023	0.2451	0.3734	0.2690	0.3381
C10b0	0.0206	-0.0077	0.0103	0.2820	0.1016	0.2450	0.3732	0.2683	0.3380
C10c0	0.0289	-0.0119	0.0144	0.2819	0.1008	0.2450	0.3730	0.2676	0.3378
C10d0	0.0425	-0.0172	0.0212	0.2819	0.0998	0.2449	0.3728	0.2667	0.3376
C10e0	0.0611	-0.0240	0.0305	0.2818	0.0985	0.2449	0.3726	0.2656	0.3374
C10a	0.0143	-0.0020	0.0072	0.2328	-0.0020	0.2023	0.2583	0.1359	0.2339
C10b	0.0222	-0.0113	0.0111	0.2329	-0.0041	0.2024	0.2578	0.1344	0.2335
C10c	0.0363	-0.0227	0.0181	0.2331	-0.0065	0.2025	0.2572	0.1327	0.2329
C10d	0.0555	-0.0369	0.0277	0.2333	-0.0095	0.2028	0.2566	0.1307	0.2324
C10e	0.0805	-0.0552	0.0402	0.2338	-0.0132	0.2031	0.2559	0.1281	0.2318

$T_{m-1,0}$ [s]

	WHI			BVH			ONS		
	RMSE	Bias	SI	RMSE	Bias	SI	RMSE	Bias	SI
Normal	0.8819	-0.6342	0.1562	0.7621	0.0058	0.1788	0.8953	-0.2699	0.1767
C5b	0.3625	-0.0275	0.0642	0.8482	0.6593	0.1990	0.6527	0.4316	0.1288
C6a	0.3669	-0.0456	0.0650	0.8148	0.6371	0.1912	0.6138	0.4033	0.1211
C6b	0.3784	-0.0687	0.0670	0.7737	0.6049	0.1815	0.5934	0.3666	0.1171
C6c	0.3704	-0.0299	0.0656	0.8298	0.6522	0.1947	0.6285	0.4336	0.1240
C6d	0.3745	-0.0528	0.0663	0.7791	0.6111	0.1828	0.6011	0.3924	0.1186
C6e	0.3766	-0.0208	0.0667	0.8368	0.6606	0.1963	0.6362	0.4445	0.1255
C6f	0.3761	-0.0428	0.0666	0.7837	0.6164	0.1839	0.6051	0.4019	0.1194
C6g	0.4245	-0.1656	0.0752	0.7445	0.5725	0.1747	0.5086	0.2641	0.1003
C6h	0.4992	-0.2301	0.0884	0.6233	0.4262	0.1462	0.4875	0.1890	0.0962
C7d	0.3736	0.0009	0.0662	0.8608	0.6802	0.2020	0.6490	0.4602	0.1281
C7e	0.3655	-0.0159	0.0647	0.8443	0.6633	0.1981	0.6340	0.4387	0.1251
S8a	0.4154	0.0085	0.0736	0.8504	0.6800	0.1995	0.6350	0.4354	0.1253
S8b	0.4240	0.0034	0.0751	0.8480	0.6783	0.1990	0.6309	0.4336	0.1245
S8c	0.4270	-0.0065	0.0756	0.8361	0.6690	0.1962	0.6181	0.4221	0.1220
S8d	0.4193	-0.0301	0.0743	0.7965	0.6331	0.1869	0.6049	0.3992	0.1194
S8e	0.4190	-0.0302	0.0742	0.7962	0.6329	0.1868	0.6049	0.3991	0.1194
S8f	0.4163	-0.0339	0.0737	0.7897	0.6251	0.1853	0.6048	0.3922	0.1193
S8g	0.4146	-0.0477	0.0734	0.7796	0.6139	0.1829	0.5988	0.3822	0.1181
S8h	0.4090	-0.0652	0.0724	0.7677	0.6000	0.1801	0.5903	0.3683	0.1165
S8i	0.4175	-0.0925	0.0740	0.7544	0.5855	0.1770	0.5840	0.3532	0.1152
S8j	0.4223	-0.1028	0.0748	0.7460	0.5781	0.1750	0.5821	0.3473	0.1149
S8k	0.4263	-0.1089	0.0755	0.7425	0.5755	0.1742	0.5811	0.3443	0.1147
C9a	0.0198	-0.0151	0.0035	0.7619	0.5988	0.1788	0.5855	0.3647	0.1155
C9b	0.0206	-0.0152	0.0036	0.7081	0.5476	0.1661	0.5027	0.2656	0.0992
C9c	0.3560	-0.1217	0.0631	0.7618	0.5951	0.1787	0.5854	0.3551	0.1155
C9d	0.4742	-0.2561	0.0840	0.7026	0.5349	0.1648	0.4948	0.2430	0.0976
C10a0	0.4546	-0.1557	0.0805	0.7696	0.5989	0.1806	0.5902	0.3559	0.1165
C10b0	0.4516	-0.1541	0.0800	0.7677	0.5973	0.1801	0.5899	0.3552	0.1164
C10c0	0.4483	-0.1524	0.0794	0.7655	0.5955	0.1796	0.5896	0.3543	0.1163
C10d0	0.4442	-0.1500	0.0787	0.7630	0.5934	0.1790	0.5892	0.3532	0.1162
C10e0	0.4389	-0.1470	0.0777	0.7599	0.5907	0.1783	0.5886	0.3519	0.1161
C10a	0.5724	-0.3487	0.1014	0.7294	0.5517	0.1712	0.5012	0.2432	0.0989
C10b	0.5697	-0.3469	0.1009	0.7245	0.5470	0.1700	0.5005	0.2412	0.0988
C10c	0.5662	-0.3445	0.1003	0.7185	0.5414	0.1686	0.4997	0.2388	0.0986
C10d	0.5621	-0.3417	0.0996	0.7111	0.5344	0.1669	0.4987	0.2358	0.0984
C10e	0.5564	-0.3377	0.0986	0.7023	0.5259	0.1648	0.4975	0.2321	0.0982

Appendix B: Examples of statistical hypothesis tests

F-Test

The F-test (Snedecor and Cochran, 1983) is used to test if the standard deviations of two populations are equal.

$$\frac{s_{X1}^2}{\left(a^2\right)s_{X2}^2} \neg F\left(\nu_1 - 1, \nu_2 - 1\right) \qquad \text{[B.1]}$$

where s_X^2 is the sample variance, ν is the number of degrees of freedom, a is a factor (usually equal to 1), F is the F-distribution, and the subscripts 1 and 2 correspond to the two samples.

t-test

The two-sample t-test is used to determine if two population means are equal.

$$\frac{\overline{X}_1 - (a)\overline{X}_2}{\sqrt{\frac{s_{X1}^2}{\nu_1} + \frac{(a^2)s_{X2}^2}{\nu_2}}} \neg t\left(\nu_1 + \nu_2 - 2\right) \qquad \text{[B.2]}$$

where $\overline{X}$ is the sample mean, s_X^2 is the sample variance, ν is the number of degrees of freedom, a is a factor (usually equal to 1), t is the t-distribution, and the subscripts 1 and 2 correspond to the two samples.

Table B.1. Upper critical values of the F distribution for v_1 numerator degrees of freedom and v_2 denominator degrees of freedom, and for the t distribution for v degrees of freedom with 5% and 10% significance level.

$v_1 = v_2 = v$	$F_5(v_1, v_2)$	$F_{10}(v_1, v_2)$	$t_5(v)$	$t_{10}(v)$
10	2.9769	2.4098	2.1265	1.6740
20	2.4638	2.1408	2.0845	1.6704
30	2.2048	1.9743	2.0696	1.6685
40	2.0451	1.8634	2.0619	1.6674
50	1.9349	1.7834	2.0572	1.6667
60	1.8532	1.7223	2.0541	1.6662
70	1.7896	1.6738	2.0518	1.6658
80	1.7383	1.6341	2.0501	1.6656
90	1.6959	1.6008	2.0488	1.6654
100	1.6600	1.5723	2.0477	1.6652
110	1.6292	1.5477	2.0469	1.6650
120	1.6024	1.5260	2.0461	1.6649
130	1.5788	1.5068	2.0455	1.6648
140	1.5577	1.4897	2.0450	1.6647
150	1.5388	1.4742	2.0445	1.6647
160	1.5218	1.4601	2.0441	1.6646
170	1.5063	1.4473	2.0438	1.6645
180	1.4921	1.4355	2.0435	1.6645
190	1.4790	1.4246	2.0432	1.6644
200	1.4670	1.4146	2.0429	1.6644

List of symbols

Latin symbols

A, B	calibration parameters for spectral low energy threshold removal
A_k,B_k,G_k	coefficients of linear transformation in the Kalman filter
C_{bf}	bottom friction coefficient
C_k	observational operator, obtains state values at observation locations
C_{nl}	constant for nonlinear interactions, DIA approximation
C_{wc}	white-capping coefficient
C_x,C_y	propagation velocities in geographical space
C_θ,C_σ	propagation velocities in spectral space
c, c_p	wave phase speed [m/s]
d	water depth [m]
$E(f)$	1D wave energy density spectrum [m^2 s]
E	total mean wave energy [J]
e	mean energy of a wave component [J]
$e_{k\|\|k}$, $e_{k\|\|k-1}$	state estimator error, prior state estimator error
$F(f)$	1D wave energy spectrum [m^2 s]
F	wave energy spectrum [m^2 s rad^{-1}]
$\hat{F}$	filtered wave energy spectrum [m^2 s rad^{-1}]
F_5,F_{10}	F-distribution for 5% and 10% significance level
f	wave frequency [Hz]
f_{Hm0}	weighting factor corresponding to the H_{m0} variable
f_m	mean wave frequency [Hz]
$f_{m-1,0}$	mean wave frequency based on the moments -1 and 0 [Hz]
f_M	peak frequency of the steepness function [Hz]
$f_{\max}$	upper-frequency limit of the wave energy spectrum [Hz]
f_{PM}	peak frequency of the Pierson-Moskowitz spectrum [Hz]
f_p	peak frequency [Hz]
f_s	wind-sea – swell separation frequency [Hz]
G	total derivative of the cost function respect to the control variable (∇J)
G	coupling coefficient in the Boltzmann integral
g	gravity acceleration [ms^{-1}]
H	observational operator, projects state values at observation locations
H_{m0}	zero moment significant wave height [m]
H_s	Significant wave height [m]
h	water depth (positive in the downwards direction) [m]
I	identity matrix

I_n	identity matrix or order n
J	cost or objective function
Jo	initial value of the cost function
$\mathbf{K}$	assimilation gain matrix in the OI approach
K_k	Kalman gain matrix
k	wave number
$\tilde{k}$	mean wave number
L	unconstrained cost function
L	mean height above the water for wind observation [m]
L	correlation length
M_{ij}	weighting matrix for the cost function
N	wave action density spectrum [m^2 s^2 rad^{-1}]
n	white-capping coefficient (equal to 0.5)
$P_{k\|k-1}$	prior error covariance matrix
$P_{k\|k}$	error covariance matrix
p	white-capping coefficient (equal to 4)
$p(x\|x_b)$	background probability distribution
$p(x\|x_o)$	observation probability distribution
$p(x\|x_b,x_o)$	joint probability distribution of x given $p(x\|x_b)$, and $p(x\|x_o)$
$\mathbf{Q}$	system, model, or background error covariance matrix
Q_k	system, model, or background error covariance matrix
$\mathbf{R}$	observations error covariance matrix
R_k	observations error covariance matrix
r_{xy}	correlation coefficient
$\|\mathbf{r}_i - \mathbf{r}_j\|$	distance between points i and j
S_{bf}	bottom friction dissipation term [m^2 rad^{-1}]
S_{db}	depth induced wave braking term [m^2 rad^{-1}]
S_{in}	wind input term [m^2 rad^{-1}]
S_{nl}	non linear wave-wave interactions term [m^2 rad^{-1}]
S_{nl3}	triad wave-wave interactions term [m^2 rad^{-1}]
S_{TOT}	total source function [m^2 rad^{-1}]
S_{wc}	white-capping dissipation term [m^2 rad^{-1}]
s^2, $s_X{}^2$	sample variance
$\tilde{s}$	mean wave steepness
$\tilde{s}_{PM}$	mean wave steepness of the Pierson and Moskowitz spectrum
T_m	mean wave period [s]
T_{m-01}	mean wave period (-1 moment) [s]
T_{m02}	second moment mean wave period [s]
T_p,T_{peak}	peak period [s]
t	time [s]
t_5,t_{10}	t-distribution with 5% (10%) significance level

U	wind speed [m/s]
$U_{\mathbf{k}}$	mean square of the bottom velocity associated to the wave number **k**
U_L, U_z	wind speed at height L, z [m/s]
U_{10}, U_{25}	wind velocity at 10m, 25m height [m/s]
u_*	wind friction velocity [m/s]
u_k	input vector for the Kalman filter approach
$u_{rms,bottom}$	root mean square orbital velocity at the bottom [m/s]
v_k	measurements stochastic noise
w_i, w_j, w_k	state stochastic noise
$\overline{X}$	sample mean
x	state variable
x, y	space coordinates [m]
x_a	mean value of the analyzed state variable
x_b	mean value of the background state variable
x_k	system sate vector
x_o	mean value of the observed state variable
$\hat{x}_{k\|k-1}$	system sate estimator
$\hat{x}$	mean value of the state variable
y_k	observations vector
$\hat{y}_{k\|k-1}$	observation state estimator
z_0	effective roughness length [m]

Greek symbols

α	initial linear growth rate of wind input [m^2 rad^{-1}]
α	roughness length constant
α	Phillips' equilibrium range parameter
α	spectral energy rescaling factor
$\alpha(f_*)$	wave mean steepness function
α_{PM}	Phillips' constant for the Pierson-Moskowitz spectrum (0.0081)
β	spectral frequency shift factor
β	exponential growth rate of wind input [s^{-1}]
β_m	wind input constant
β	calibration factor for wind-sea – swell determination
Γ	constant for bottom friction coefficient [$m^2 s^{-3}$]
γ	JONSWAP peak enhancement factor
γ^*	ratio of peak energies between a given spectrum and a PM spectrum
Δf^2	squared distance between spectral peaks [Hz^2]
δ	Dirac delta function
$\overline{\delta f^2}$	spectral spread [Hz^2]
ε	ratio of air to water density
θ_m	mean wave direction [rad]

θ	wave direction [rad]
θ_w	mean wave direction [rad]
κ	von Karman constant (equal to 0.41)
κ	convolution kernel
λ	parameter for nonlinear interactions, DIA approximation (0.25)
λ	Lagrange multiplier
λ	fitting parameter for the correlation coefficient
μ	statistical parameter for white-capping dissipation
ν	number of degrees of freedom in a sample
ρ_{ij}	correlation coefficient
σ	intrinsic radiant frequency [Hz]
$\overline{\sigma}$	mean radiant frequency [Hz]
σ	standard deviation
σ^2	variance of the state variable
σ_a, σ_b	JONSWAP spectral width factors
σ_x,σ_y	axis lengths for 2D Gaussian function
τ_w	wave-induced stress
τ	total stress
$\tau_{turbulent}$	turbulent stress
ϕ	angular parameter for the two-dimensional Gaussian function
ψ	wind direction
ψ,ψ_1,ψ_2	control or tuning variables of the cost function
Ω	wind profile parameter

Mathematical operators

| |	absolute value
‹ ›	expected value operator
$\otimes$	convolution operator

List of acronyms

A2B	A2 Buoy (buoy location)
AKZ	Akkaert Zuid (buoy location)
BVH	Bol van Heist (buoy location)
CICESE	Centro de Investigación Científica y de Educación Superior de Ensenada, B.C. México
DELO	Deurlo (buoy location)
DOLPHIN-B	Second generation wave model
ECMWF	European Centre for Medium-Range Weather Forecasts
EKF	Extended Kalman filter
EnKF	Ensemble Kalman Filter
ERS	European Remote Sensing satellite
HFR	High Frequency Radar
JONSWAP	Joint North Sea Wave Project (spectrum)
MDK	Maritieme Dienstverlening en Kust
MP7	Meteo Paal 7 (Meteorological station)
MUMM	Management Unit of the North Sea Mathematical Models
NDBC	National Data Buoy Center
NMC	National Meteorological Center
ONEMOD	One dimensional version of the WAM model
ONS	Oostende (buoy location)
PM	Pierson-Moskowitz (spectrum)
SAR	Synthetic Aperture Radar
SCHB	Schouwenbank (buoy location)
SWAN	Simulating Waves Nearshore
UKMO	United Kingdom Meteorological Office
UTC	Coordinated Universal Time
VLIZ	Vlaams Instituut voor de Zee
WAM	Wave Model
WHI	Westhinder (buoy location)
ZEGE	Monitoring network Zeeuwse tidal waters

References

Aarnes, Jorg E. and Harald E. Krogstad, Partitioning sequences for the dissection of directional ocean wave spectra: A review, Part of work package 4 (Wp4) of the EnviWave (EVG-2001-00017) research program under the EU Energy, Environment and Sustainable Development program, 23 pp., 2001

Abbott, E. and D. Powell, Land-vehicle navigation using GPS, *Proceedings of the IEEE, 87*(1), 145-162, 1999.

Afdeling Kust (MDK) and VLIZ, Vlaamse Overheid, Afdeling Kust, Maritieme Dienstverlening en Kust (MDK), Mobiliteit en Openbare Werken, Vlaams Instituut voor de zee, VLIZ (on line), *http://www. vliz. be/,* 2008.

Anonymous, Golfvoortplantingsmodel voor de zuidelijke noordzee, Rapport Deelopdracht 2, Leveren numeriek noordzeemodel voor stroming en golven. Veiligheidsniveau Vlaanderen Kustverdediging, Bestek nr. 16EB/2000/08. Administratie Waterwegen en Zeewezen Afdeling Waterbouwkundig Laboratorium en Hydrologisch Onderzoek, Faculteit Toegepastewetenschappen - Departement Burgerlijke Bouwkunde - Laboratorium voor Hydraulica, 63 pp, 2002.

Apte, A., M. Hairer, A. M. Stuart and J. Voss, Sampling the posterior: An approach to non-Gaussian data assimilation, *Physica D: Nonlinear Phenomena,* **230**(1-2), 50-64, 2007.

Babovic, V., S. A. Sannasiraj and E. S. Chan, Error correction of a predictive ocean wave model using local model approximation, *Journal of Marine Systems,* **53**, 1-17, 2005.

Barrero Mendoza, Oscar, Data assimilation in magnetohydrodynamics systems using Kalman filtering, *PhD Thesis* of the Katholieke Universiteit Leuven (Belgium), 186 pp., 2005.

Barzel, G. and R. B. Long, Wave model fitting using the adjoint technique, in *Dynamics and Modeling of Ocean Waves*, edited by G. J. Komen, L. Cavaleri, M. A. Donelan, K. Hasselmann and S. Hasselmann, (**VI.**7), 447-453, Cambridge University Press, 1994.

Battjes, J. A. and P. A. E. M. Janssen, Energy loss and set-up due to breaking of random waves, *Proceedings of the 16th International Conference Coastal Engineering*, 569-587, 1978.

Battjes, J. A., Shallow water wave modelling, proceedings of the *Int.Symp.Waves - Physical and Numerical Modelling*, Vancouver, (**I**), 1-23, 1994.

Bauer, E., S. Hasselmann and K. Hasselmann, Validation and assimilation of Seasat altimeter wave heights using the WAM wave model, *Journal of Geophysical Research,* **97**(C8), 12671-12682, 1992.

Bauer, E., K. Hasselmann, I. R. Young and S. Hasselmann, Assimilation of wave data into the wave model WAM using an impulse response function method, *Journal of Geophysical Research,* **101**, 3801-3816, 1996.

Bidlot, J. R., F. Ovidio, and D. Van den Eynde, Validation and improvement of the quality of the operational wave model MU-WAVE by the use of ERS-I satellite data, MUMM/T3/AR04, 53 pp., 1995.

Bidlot, J. R., ECMWF wave-model products, Newsletter No. 91, ECMWF, 9-15, 2001

Bidlot, J. R., P. A. E. M. Janssen, and Saleh Abdalla, A revised formulation of ocean wave dissipation and its model impact, Technical Memorandum 509, ECMWF, 27 pp., 2007

Bouttier, F. and P. Courtier, Data assimilation concepts and methods. Meteorological Training Course Lecture Series, ECMWF, 59 pp., 1999.

Breivik, L. A. and M. Reistad, Assimilation of ERS-1 altimeter waveheights in an operational numerical wave model, *Weather and Forecasting,* **9**, 440-551, 1994.

Breivik, L. A., M. Reistad, H. Schyberg, J. Sunde, H. E. Krogstad and H. Johnsen, Assimilation of ERS SAR wave spectra in an operational wave model, *Journal of Geophysical Research,* **103(C4)**, 7887-7900, 1998.

Brüning, C., S. Hasselmann, K. Hasselmann, S. Lehner and T. W. Gerling, First evaluation of ERS-1 synthetic aperture radar wave mode data, *The Global, Athmospheric and Ocean System,* **2**, 61-98, 1994.

Cavaleri, L., J. H. G. M. Alves, F. Ardhuin, A. V. Babanin, M. L. Banner, K. Belibassakis, M. Benoit, M. A. Donelan, J. Groeneweg, T. H. C. Herbers, P. A. Hwang, P. A. E. M. Janssen, T. Janssen, I. V. Lavrenov, R. Magne, J. Monbaliu, M. Onorato, V. Polnikov, D. T. Resio, W. E. Rogers, A. Sheremet, J. Kee Smith, H. L. Tolman, G. van Vledder, J. Wolf and I. R. Young, Wave modelling : the state of the art, *Progress In Oceanography,* **75**(4), 603-674, 2007.

Coli, A., J. A. Santos and A. A. Pires-Silva, Inverse methodology for coastal wave data assimilation, proceedings of the *International Conference of Coastal Engineering 2006*, San Diego U.S.A., (**1**), 425-437, 2006.

Collins, S. M. a. o., Hurricane Katrina: A nation still unprepared. Special report of the Committee on Homeland Security and Governmental Affairs, United States Senate together with additional views, 737 pp., 2006.

Cressman, G. P., An operational objective analysis system, *Monthly Weather Review,* 367-374, 1959.

Daley, R., Atmospheric Data Analysis, Cambridge University Press, 471 pp., 1991.

Datawell manual, Datawell - Oceanographic Instruments, Datawell Waverider Reference Manual, 137 pp., Haarlem, 2006.

De Las Heras, M. M., G. Burgers and P. A. E. M. Janssen, Variational Wave Data Assimilation in a Third-Generation Wave Model, *Journal of Atmospheric and Oceanic Technology,* **11**(5), 1350-1369, 1994.

De Las Heras, M. M., G. Burgers and P. A. E. M. Janssen, Wave data assimilation in the WAM wave model, *Journal of Marine Systems,* **6**(1-2), 77-85, 1995.

De Mulder, Tom, Jaak Monbaliu, and Frank Mostaert, Veiligheidsniveau Vlaanderen kustverdediging. Opmaak van een numerieke golfdatabank voor de Vlaamse kust, MOD 644, Waterbouwkundig Laboratorium Borgerhout - Laboratorium voor Hydraulica K.U.Leuven, 31 pp., 2004.

De Valk, C. F. and C. J. Calkoen, Wave data assimilation in a third generation wave prediction model for the North Sea - an optimal control approach, Report X38, Delft Hydraulics, The Netherlands, 123 pp., 1989

De Valk, C. F., A wind and wave data assimilation scheme based on the adjoint technique, in *Dynamics and Modeling of Ocean Waves*, edited by G. J.

Komen, L. Cavaleri, M. A. Donelan, K. Hasselmann and S. Hasselmann, (**VI.9**), 460-468, Cambridge University Press, 1994.

Dee, D. P., On-line Estimation of Error Covariance Parameters for Atmospheric Data Assimilation, *Monthly Weather Review,* **123**, 1128-1145, 1995.

Donelan, M. A., J. Hamilton and W. Hui, Directional spectra of wind-generated waves, *Philosophical transactions of the Royal Society, London,* **315**, 509-562, 1985.

Drennan, W. M., Graber H.C., D. Hauser and C. Quentin, On the wave age dependence of wind stress over pure wind seas, *Journal of Geophysical Research,* **108**(C3 FET10), 1-13, 2003.

Dunlap, E. M., R. B. Olsen, L. Wilson, S. De Margerie and R. Lalbeharry, The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model, *Journal of Geophysical Research,* **103**(C4), 7901-7915, 1998.

Earl, M. D., Development of algorithms for separation of sea and swell, Tech. Rep. MEC-87-1, National Data Buoy Center, 1984.

Eisner, R. K., Planning for tsunami: Reducing future losses through mitigation, *Natural Hazards,* **35**(1), 155-162, 2005.

Esteva, D. C., Evaluation of Preliminary Experiments Assimilating Seasat Significant Wave Height into a Spectral Wave Model, *Journal of Geophysical Research,* **93**(C), 14099-14106, 1988.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research,* **99** (C5), 10143-10162, 1994.

Evensen, G., The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics,* **53**, 343-367, 2003.

Evensen, G., Data Assimilation, The ensemble Kalman filter, Springer, 279 pp., 2006.

Flather, R. A., Results from a model of the North east Atlantic relating to the Norwegian Coastal current, proceedings of the *Norwegian Coastal Current Symposium*, Geilo, (**2**), 427-458, 1981.

Francis, P. E. and R. A. Stratton, Some experiments to investigate the assimilation of Seasat altimeter wave height data into a global wave model, *Quarterly Journal of the Royal Meteorological Society,* **116**, 1225-1251, 1990.

García, Nava Héctor, Evaluación del flujo de momento entre la atmósfera y el océano bajo diferentes condiciones de oleaje, *Master Thesis* of the CICESE (México), pp. 74, 2006.

Gelb, A., J. F. Kasper and R. A. Nash, Applied optimal estimation, Cambridge (Mass.): MIT press, 374 pp., 1994.

Gelci, R., H. Cazalé and J. Vassal, Prévision de la houle. La méthode des densités spectroangulaires, *Bulletin d'information du Comité central d'Océanographie et d'Etude des Cøtes,* **9**, 416-435, 1957.

Gerling, T. W., Partitioning sequences and arrays of directional ocean wave spectra into component wave systems, *Journal of Atmospheric and Oceanic Technology,* **9**(4), 444-458, 1992.

Giering, R. and T. Kaminski, Recipes for Adjoint Code Construction, *ACM Transactions on Mathematical Software,* **24**(4), 437-474, 1998.

Giles, M. and N. Pierce, An introduction to the adjoint approach to design, *European Journal of Flow, Turbulence and Control,* 1-26, 2000.

Gilhousen, D. B. and R. Hervey, Improved Estimates of Swell from Moored Buoys, Proceedings of the *Fourth International Symposium WAVES 2000*, ASCE, Alexandria, 387-393, 2001.

Greenslade, D. J. M., The assimilation of ERS-2 significant wave height data in the Australian region, *Journal of Marine Systems,* **28**, 141-160, 2001.

Greenslade, D. J. M. and I. R. Young, Background errors in a global wave model determined from altimeter data, *Journal of Geophysical Research,* **109**(C09007), 1-24, 2004.

Grewal, M. S. and A. P. Andrews, Kalman filtering: theory and practice using MATLAB (2nd ed.), Wiley-Interscience, 416 pp., 2001.

Günther, Heinz, S. Hasselmann, and P. A. E. M. Janssen, WAM Model Cycle 4. (User Manual), Technical Report No. 4, Max-Planck-Institut für Meteorologie, Hamburg, 1992.

Hanson, J. L. and O. M. Phillips, Automated Analysis of Ocean Surface Directional Wave Spectra, *Journal of Atmospheric and Oceanic Technology,* **18**(2), 277-293, 2001.

Hasselmann, K., On the non-linear energy transfer in a gravity wave spectrum. Part 1: General theory, *Journal of Fluid Mechanics,* **12**, 481-500, 1962.

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

Hasselmann, K., On the spectral dissipation of ocean waves due to white capping, *Boundary-Layer Meteorology,* **6**, 107-127, 1974.

Hasselmann, S. and K. Hasselmann, Computation and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part I: a new method for efficient computations of the exact nonlinear transfer, *Journal of Physical Oceanography,* **15**, 1369-1377, 1985.

Hasselmann, S., K. Hasselmann and C. Brüning, Extraction of wave spectra from SAR image spectra, in *Dynamics and Modelling of Ocean Waves*, edited by G. J. Komen, L. Cavaleri, M. A. Donelan, K. Hasselmann and S. Hasselmann, (**V.4**), 391-401, Cambridge University Press, 1994.

Hasselmann, S., C. Brüning, K. Hasselmann and P. Heimbach, An improved algorithm for retrieval of ocean wave spectra from synthetic aperture radar image spectra, *Journal of Geophysical Research,* **101**(C7), 16615-16629, 1996.

Hasselmann, S., P. Lionello and K. Hasselmann, An optimal interpolation scheme for the assimilation of spectral wave data, *Journal of Geophysical Research,* **102**, 15823-15836, 1997.

Hauser, D., K. Kahma, H. E. Krogstad, S. Lehner, J. Monbaliu and L. Wyatt, "COST action 714 -Measuring and analyzing the directional spectrum of ocean waves" Luxembourg: Office for Official Publications of the European Communities, **XV**, Luxembourg, 2005.

Hersbach, H., The adjoint of the WAM model, Scientific Report WR 97-01, KNMI, 28 pp., 1997.

Hoang, S., R. Baraille, O. Talagrand, X. Carton and P. De Mey, Adaptive filtering: application to satellite data assimilation in oceanography, *Dynamics of Atmospheres and Oceans,* **27**, 257-281, 1997.

Hollingsworth, A., Objective analysis for numerical weather prediction. Short- and medium-range numerical weather prediction, proceedings of the *WMO/IUGG NWP Symp.*, Meteorological Society of Japan, Tokyo, 11-59, 1986.

Hollingsworth, A. and P. Lonnberg, The statistical structure of short-range forecast errors as determined from radiosonde data, Part I: The wind field, *Tellus,* **38A**, 111-136, 1986.

Holthuijsen, L. H., N. Booij, M. van Endt, S. Caires and C. Guedes Soares, Assimilation of buoy and satellite data in wave forecasts with integral control variables, *Journal of Marine Systems,* **13**, 21-31, 1997.

Holthuijsen, L. H., Waves in oceanic and coastal waters, Cambridge University Press, 387 pp. 2007.

Hoteit, I., G. Korres and G. Triantafyllou, Comparison of extended and ensemble based Kalman filters with low and high resolution primitive equation ocean models, *Nonlinear Processes in Geophysics,* **12**(Part of Special Issue "Quantifying predictability"), 755-765, 2005.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie and H. L. Mitchell, A System Simulation Approach to Ensemble Prediction, *Monthly Weather Review,* **124**(6), 1225-1242, 1996.

IMDC and Alkyon, Structureel herstel van de kustverdediging te Oostende en verbetering van de haventoegang naar de haven van Oostende; hydrodynamische randvoorwaarden voor het ontwerp; waterstanden & golfclimaat, in opdracht van de Afdeling Waterwegen Kust (bestek nr. 16/EH/01/27), 26-30070-200/SR002, 262 pp., 2002.

Janssen, P. A. E. M., P. Lionello, M. Reistad and A. Hollingsworth, Hindcast and data assimilation studies in a third generation wave model during the SEASAT period, *Journal of Geophysical Research,* **94**(C), 973-993, 1989.

Janssen, P. A. E. M., Quasi-linear theory of wind wave generation applied to wave forecasting, Report 00-90-13, 22 pp., 1990.

Janssen, P. A. E. M., Quasi-linear theory of wind wave generation applied to wave forecasting, *Journal of Physical Oceanography,* **21**, 1631-1642, 1991.

Janssen, P. A. E. M., The interaction of ocean waves and wind, Cambridge University Press, 308 pp., 2004.

Janssen, P. A. E. M., Progress in ocean wave forecasting, *Journal of Computational Physics,* **227**(7), 3572-3594, 2008.

Jazwinski, A. H., Stochastic Processes and Filtering Theory, Academic Press, New York, 376 pp., 1970.

Julier, S. and J. Uhlmann, A new extension of the Kalman filter to nonlinear systems, proceedings of the *Signal Processing, Sensor Fusion, and Target Recognition VI Conference*, Society of Photo-Optical Instrumentation Engineers (SPIE), (**3068**), 182-193, 1997.

Kahma, K. K. and C. J. Calkoen, Reconciling discrepancies in the observed growth of wind-generated waves, *Journal of Physical Oceanography,* **22**, 1389-1405, 1991.

Kalman, R. E., A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME - Journal of Basic Engineering,* **82**, 35-45, 1960.

Kitaigorodskii, S. A., Application of the theory of similarity to the analysis of wind-generated water waves as a stochastic process, Bull.Acad.Sci.USSR Geophys., Ser. No. 1, 73pp., 1962.

Komen, G. J., S. Hasselmann and K. Hasselmann, On the Existence of a Fully Developed Wind-Sea Spectrum, *Journal of Physical Oceanography,* **14**(8), 1271-1285, 1984.

Komen, G. J., L. Cavaleri, M. A. Donelan, K. Hasselmann, S. Hasselmann and P. A. E. M. Janssen, Dynamics and Modelling of Ocean Waves., Cambridge University Press, 554 pp., 1994.

Kuik, A. J., G. van Vledder and L. H. Holthuijsen, A Method for the Routine Analysis of Pitch-and-Roll Buoy Wave Data, *Journal of Physical Oceanography,* **18**, 1020-1034, 1988.

Langley, A., A. Wright, G. Hurry, J. Hampton, T. Aqorua and L. Rodwel, Slow steps towards management of the world's largest tuna fishery, *Marine Policy,* ***33***(2), 271-279, 2009.

Lefevre, J. M., C. Quentin, D. Hauser and J. Bidlot, Partitioning of wave spectra, in "COST action 714 -Measuring and analysing the directional spectrum of ocean waves", edited by D. Hauser, K. Kahma, H. E. Krogstad, S. Lehner, J. Monbaliu and L. Wyatt, 183-195, 2005.

Lionello, P. and P. A. E. M. Janssen, Assimilation of altimeter measurements to update swell spectra in wave models, proceedings of the *International symposium on assimilation of observations in meteorology and oceanography*, World Meteorological Organization, Geneva, 241-246, 1990.

Lionello, P., H. Günther and P. A. E. M. Janssen, Assimilation of altimeter data in a global third-generation wave model, *Journal of Geophysical Research,* **97**(14), 453-474, 1992.

Lionello, P., H. Günther and B. Hansen, A sequential assimilation scheme applied to global wave analysis and prediction, *Journal of Marine Systems,* **6**, 87-107, 1995.

Longuet-Higgins, M. S., D. E. Cartwright and N. D. Smith, Observations of the Directional Spectrum of Sea Waves Using the Motions of a Floating Buoy, in Ocean Wave Spectra, Englewood Cliffs, N.J., Prentice-Hall, 111-136, 1963.

Luo, Weimin, Wave modelling in shallow water, *PhD Thesis* of the Civil Engineering Department, Katholieke Universiteit Leuven (Belgium), 1995.

Madsen, H. and R. Cañizares, Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling, *International Journal for Numerical Methods in Fluids,* **31**, 961-981, 1999.

Mandal, S., S. Rao and D. H. Raju, Ocean wave parameters estimation using back propagation neural networks, *Marine Structures,* **18**(3), 301-308, 2005.

Mastenbroek, C., V. K. Makin, A. C. Voorrips and G. J. Komen, Validation of ERS-1 altimeter wave height measurements and assimilation in a North Sea wave model, *The Global Atmosphere and Ocean System,* **2**, 143-161, 1994.

Maybeck, P. S., Stochastic Models, Estimation, and Control, Academic Press, Inc., 16 pp., 1979.

Miles, J. W., On the generation of surface waves by shear flows., *Journal of Fluid Mechanics,* **3**, 185-204, 1957.

Monbaliu, Jaak, Wind and waves: Investigation of an optimization approach to parameter estimation, *PhD Thesis* of the Department of Civil Engineering, Katholieke Universiteit Leuven (Belgium), 1992.

Monbaliu, J., J. C. Hargreaves, J. C. Carretero Albiach, H. Gerritsen and R. A. Flather, Wave modelling in the PROMISE project, *Coastal Engineering,* **37**(3-4), 379-407, 1999.

Monbaliu, J., R. Padilla-Hernández, J. C. Hargreaves, J. C. Carretero Albiach, W. Luo, M. Sclavo and H. Günther, The spectral wave model, WAM, adapted for applications with high spatial resolution, *Coastal Engineering,* **41**, 41-62, 2000.

Ngodock, H. E., S. R. Smith and G. A. Jacobs, Cycling the Representer Algorithm for Variational Data Assimilation with the Lorenz Attractor, *Monthly Weather Review, **135***(2), 373-386, 2007.

Osuna, P., On the high-resolution simulation of the dynamic interaction between current and waves in coastal waters: an application to the southern North Sea, *PhD Thesis* of the Katholieke Universiteit Leuven (Belgium), 196 pp., 2002.

Ovidio, F., J. R. Bidlot, D. Van den Eynde, W. Luo and J. Monbaliu, ERS-1 Data Assimilation in a Second-Generation Wave Model for the North Sea, proceedings of the *ERS Applications*, (**383**), 1-425, 1996.

Ozer, J., R. Padilla-Hernandez, J. Monbaliu, E. Alvarez Fanjul, J. C. Carretero Albiach, P. Osuna, J. C. S. Yu and J. Wolf, A coupling module for tides, surges and waves, *Coastal Engineering,* **41**, 95-124, 2000.

Padilla-Hernandez, R., Numerical modelling of wind wave energy dissipation at the bottom including ambient currents, PhD Katholieke Universiteit Leuven, 2002.

Parrish, D. F. and J. C. Derber, The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, *Monthly Weather Review,* **120**(8), 1747-1763, 1992.

Phillips, O. M., On the generation of waves by turbulent wind, *Journal of Fluid Mechanics,* **2**, 415-417, 1957.

Phillips, O. M., The equilibrium range in the spectrum of wind-generated waves, *Journal of Fluid Mechanics,* **4**, 426-433, 1958.

Phillips, O. M., Spectral and statistical properties of the equilibrium range in wind-generated gravity waves, *Journal of Fluid Mechanics,* **156**, 505-531, 1985.

Pierson, W. J., G. Neumann, and R. James, Practical Methods for Observing and Forecasting Ocean Waves, Publication 603, U.S.Navy Hydrographic Office, 284 pp., 1955.

Pierson, W. J. and L. Moskowitz, A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii, *Journal of Geophysical Research,* **69**, 5181-5190, 1964.

Portilla, J. and J. Monbaliu, Data assimilation in nearshore wave modeling, 2nd International Short Course on Coastal Processes and Port Engineering, IAHR - Universita della Calabria, 227-240, 2006.

Portilla, J. and J. Monbaliu, Verkennend onderzoek voor de kwaliteitscontrole op voorspelde golfspectra. Splitsing en statistische analyse van golfsystemen als een mogelijke methodologie, Faculteit ingenieurswetenschappen, Departement Burgerlijke Bouwkunde, Laboratorium voor Hydraulica, 2007.

Portilla, J., F. J. Ocampo-Torres and J. Monbaliu, Spectral Partitioning and Identification of Wind Sea and Swell, *Journal of Atmospheric and Oceanic Technology,* **26**(1), 107-122, 2009.

Quentin, Céline, Etude de la surface océanique, de sa signature radar et de ses interactions avec le flux turbulent de quantité de mouvement dans le cadre de l'expérience FETCH, *PhD Thesis* of the Université de Paris (France), 2002.

Rodríguez, G. and C. Guedes Soares, A criterion for the automatic identification of multimodal sea wave spectra, *Applied Ocean Research,* **21**(6), 329-333, 1999.

Rogers, W. E., P. A. Hwang and D. W. Wang, Investigation of Wave Growth and Decay in the SWAN Model: Three Regional-Scale Applications, *Journal of Physical Oceanography,* **33**, 366-389, 2003.

Sannasiraj, S. A., V. Babovic and E. S. Chan, Wave data assimilation using ensemble error covariances for operational wave forecast, *Ocean Modelling,* **14**, 102-121, 2006.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker and K. E. Taylor, Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *Journal of Geophysical Research,* **105** (D6), 7337–7356, 2000.

Siddons, Lee, Data Assimilation of HF Radar Data into Coastal Wave Models, *PhD Thesis* of the Department of Applied Mathematics, School of Mathematics and Statistics, The University of Sheffield (U.K.), 260 pp., 2007.

Snedecor, G. W. and W. G. Cochran, Statistical Methods (eighth edition), Iowa State University Press, 503 pp., 1989.

Stewart, R. H., Introduction To Physical Oceanography, Department of Oceanography Texas A & M University, 2008.

Sverdrup, H. and W. Munk, Wind sea and swell: theory of relations for forecasting, Publication 601, *U.S.Navy, Hydrographic Office*, 55 pp., 1947

The SWAMP Group, The Sea Wave Modeling Project (SWAMP): Principal results and conclusions, Prenum Press, New York, 256 pp., 1985.

Theocharis, Z., C. Memos and D. Koutsoyiannis, Wave height background errors simulation and forecasting via stochastic methods in deep and intermediate waters, proceedings of the *International Conference of Coastal Engineering 2006*, San Diego, U.S.A., (**1**), 578-589, 2006.

Thomas, J., Retrieval of energy spectra from measured data for assimilation into a wave model, *Quarterly Journal of the Royal Meteorological Society,* **114**(481), 781-800, 1988.

UNCTAD, Review of Maritime Transport, United Nations Conference on Trade and Development, Geneva, 2008.

Van der Westhuysen, A., Advances in the spectral modelling of wind waves in the nearshore, *PhD Thesis* of the Technische Universiteit Delft (The Nederlands), 207 pp., 2007.

Van Lancker, V., S. Deleu, V. Bellec, S. Le Bot, E. Verfaillie, M. Fettweis, D. Van den Eynde, F. Francken, V. Pison, S. Wartel, Jaak Monbaliu, J. Portilla, J. Lanckneus, G. Moerkerke, and S. Degraer, Management, research and budgeting of aggregates in shelf seas related to end-users (Marebasse), Scientific Report Year 2, Belgian Science Policy, 144 pp., 2004.

Van Lancker, V., S. Deleu, V. Bellec, I. Du Four, E. Verfaillie, M. Fettweis, D. Van den Eynde, F. Francken, Jaak Monbaliu, A. Giardino, J. Portilla, J. Lanckneus, G. Moerkerke, and S. Degraer, Management, research and budgeting of aggregates in shelf seas related to end-users (Marebasse), Scientific Report Year 3, Belgian Science Policy, 2005.

Van Vledder, G. Ph., Evaluation of model performance (Methodology, examples, subroutines), Delft Hydraulics, Delft, 1993.

Van Vledder, G. Ph. and D. P. Hurdle, Performance of Formulations for Whitecapping in Wave Prediction Models, OMAE 2002, 21st International Conference on Offshore Mechanics & Arctic Engineering, Oslo, Norway, 2002.

Van Vledder, G., The WRT method for the computation of non-linear four-wave interaction in discrete spectral wave models, *Coastal Engineering,* **53**, 223-242, 2006a.

Van Vledder, G., Towards an optimal computation of non-linear four-wave interactions in operational wave models. JCOMM Technical Report No. 34 / WMO-TD. No. 1368, Proceedings of the *9th International workshop on wave hindcasting and forecasting*, Victoria, B.C., Canada, 1-13, 2006b.

Violante-Carvalho, N., C. E. Parente, I. S. Robinson and L. M. P. Nunes, On the growth of wind generated waves in a swell dominated region in the South Atlantic, *Journal of Offshore Mechanics and Arctic Engineering,* **124**, 14-21, 2002.

Violante-Carvalho, N., F. J. Ocampo-Torres and I. S. Robinson, Buoy observations of the influence of swell on wind waves in the open ocean, *Applied Ocean Research,* ***26*****(1-2)**, 49-60, 2004.

Voorrips, A. C. and C. F. de Valk, A comparison of two operational wave assimilation methods (KNMI Memorandum, OO-97-02.), KNMI Preprint 97-06, Physics - Atmospheric and Oceanic Physics, 46 pp., 1997.

Voorrips, A. C., V. K. Makin and S. Hasselmann, Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model, *Journal of Geophysical Research,* **102**(C3), 5829-5849, 1997.

Voorrips, A. C., A. W. Heemink and G. J. Komen, Wave data assimilation with the Kalman filter, *Journal of Marine Systems,* **19**(4), 267-291, 1999.

Vos, Sander, Variational wave data-assimilation for operational wave forecasting, *PhD Thesis* of the Technische Universiteit Delft (The Nederlands), 2002.

WAMDI Group, The WAM model - a third generation ocean wave prediction model, *Journal of Physical Oceanography,* **18**, 1775-1810, 1988.

Wang, D. W. and P. A. Hwang, An Operational Method for Separating Wind Sea and Swell from Ocean Wave Spectra, *Journal of Atmospheric & Oceanic Technology,* **18**, 2052-2062, 2001.

Welch, G. and G. Bishop, An Introduction to the Kalman Filter (Course 8), Department of Computer Science, University of North Carolina at Chapel Hill, *ACM, Inc., Chapel Hill*, 46 pp., 2001.

Wikle, C. K. and L. M. Berliner, A Bayesian tutorial for data assimilation, *Physica D: Nonlinear Phenomena,* **230**(1-2), 1-16, 2007.

Young, I. R. and T. J. Glowacki, Assimilation of altimeter wave height data into a spectral wave model using statistical interpolation, *Ocean Engineering,* **23**(8), 667-689, 1996.

Young, I. R., Wind Generated Ocean Waves, Elsevier Ocean Engineering Series, 306 pp., 1999.

Young, I. R., An intercomparison of GEOSAT, TOPEX and ERS1 measurements of wind speed and wave height, *Ocean Engineering,* ***26***(1), 67-81, 1999.

Zhang, L. and P. B. Luh, Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method, *IEEE Transactions on Power Systems,* **20**(1), 59-66, 2005.

ZEGE Meetnet, Ministerie van Verkeer en Waterstaat, Hydro Meteo Centrum Zeeland, Meetnet ZEGE (on line), *http://www. hmcz. nl/,* 2008.