Real-time reconstruction and prediction of ocean wave fields from remote optical measurements

Nicolas Desmars

CENTRALE NANTES

THESE DE DOCTORAT DE

L'ÉCOLE CENTRALE DE NANTES

ÉCOLE DOCTORALE N° 602
Sciences pour l'Ingénieur
Spécialité : Mécanique des Milieux Fluides

Par
Nicolas DESMARS

Reconstruction et Prédiction en Temps Réel de Champs de Vagues par Télédétection Optique

Thèse présentée et soutenue à Nantes, le 11/12/2020
Unité de recherche : UMR 6598, Laboratoire de recherche en Hydrodynamique, Énergétique et Environnement Atmosphérique (LHEEA)

Rapporteurs avant soutenance :

Vincent REY	Professeur des universités, Université de Toulon
Karsten TRULSEN	Professeur, University of Oslo (Norvège)

Composition du Jury :

Présidente :	Agnès MAUREL	Directrice de recherche CNRS (HDR), E.S.P.C.I. Paris
Examinateurs :	Stéphan GRILLI	Professeur, University of Rhode Island (États-Unis)
	Olivier KIMMOUN	Maître de Conférence (HDR), École Centrale de Marseille
	Yves PERIGNON	Ingénieur de recherche CNRS, École Centrale de Nantes
Dir. de thèse :	Pierre FERRANT	Professeur des universités, École Centrale de Nantes
Co-dir. de thèse :	Guillaume DUCROZET	Maître de Conférence (HDR), École Centrale de Nantes

Real-Time Reconstruction and Prediction of Ocean Wave Fields from Remote Optical Measurements

.

Reconstruction et Prédiction en Temps Réel de Champs de Vagues par Télédétection Optique

Ph.D. thesis

Nicolas Desmars

ÉCOLE CENTRALE DE NANTES

Laboratoire de recherche en Hydrodynamique, Énergétique et Environnement Atmosphérique

Last updated on 23/12/2020

Abstract

Researches conducted in this thesis address the problem of deterministic prediction of ocean wave fields around a marine structure, a key parameter for the analysis and control of a vast range of offshore operations, on the basis of datasets acquired remotely by an optical sensor. Efforts focus on the inclusion, at low computational cost, of the modeling of nonlinear hydrodynamic phenomena, preserving the reliability the surface representation in case of severe sea state.

A weakly nonlinear Lagrangian approach (ICWM), whose hydrodynamic properties are evaluated by inter-comparison with reference wave models, is selected for the description of the free surface. The prediction problem is then formulated as an inverse problem that aims at fitting the solution described by the wave model to observations, here composed of free surface elevation datasets generated by a synthetic, yet realistic, lidar sensor scanning the ocean surface at grazing angle. Predictions are then issued through the propagation in time and space of the parameterized wave model.

The applicability of the methodology is validated using observations of both unidirectional and directional wave fields, obtained at different instants to compensate for their strong spatial non-uniformity. The relative performance comparison between ICWM and lower-order wave models highlights the improvements due to the modeling of wave nonlinearities, especially those pertaining to the correction of the dispersion relation. A demonstration of the usefulness of ICWM is then provided by means of a procedure that is fully validated experimentally in a wave tank.

Keywords: Ocean waves, Gravity waves, Deterministic prediction, Nonlinear waves, Hydrodynamics, Inverse problem

Résumé

Les recherches présentées dans le cadre de cette thèse portent sur le traitement de mesures par télédétection optique de la surface océanique en vue de la prédiction de champs de vagues au voisinage d'une structure marine, information cruciale pour l'analyse et le contrôle d'une large gamme d'opérations en mer. Elles visent notamment à inclure, à moindre coût calcul, la modélisation de phénomènes non-linéaires, conservant une représentation réaliste de la surface en cas d'état de mer sévère.

Une approche Lagrangienne faiblement non-linéaire (ICWM), dont les propriétés hydrodynamiques sont étudiées par inter-comparaison avec des modèles de référence, est sélectionnée pour la description de la surface libre. Le problème de prédiction est formulé comme un problème inverse dont le but est de faire correspondre la solution décrite par le modèle de vagues à des observations, composées ici d'élévations de surface mesurées par un capteur lidar synthétique balayant la surface en incidence rasante. Les prédictions résultent de la propagation en temps/espace du modèle ainsi paramétré.

L'applicabilité de la méthodologie est validée à l'aide d'observations de champs de vagues unidirectionnels et directionnels, acquises à des instants différents pour pallier à leur forte non-uniformité spatiale. La comparaison relative des performances de ICWM et de modèles d'ordre inférieur met en évidence les améliorations dues à la modélisation des non-linéarités, notamment celles issues de la correction de la relation de dispersion. Une démonstration de la pertinence de l'utilisation de ICWM est ensuite proposée au moyen d'une procédure entièrement validée expérimentalement en bassin de houle.

Mots clés : Vagues, Houle, Ondes de gravité, Prédiction déterministe, Ondes non-linéaires, Hydrodynamique

Related Publications

Journal articles

Desmars, N., Bonnefoy, F., Grilli, S., Ducrozet, G., Perignon, Y., Guérin, C.-A. & Ferrant, P. 2020 Experimental and numerical assessment of deterministic nonlinear ocean waves prediction algorithms using non-uniformly sampled wave gauges. *Ocean Engineering* **212**, 107659

Guérin, C.-A., Desmars, N., Grilli, S. T., Ducrozet, G., Perignon, Y. & Ferrant, P. 2019 An improved lagrangian model for the time evolution of nonlinear surface waves. *Journal of Fluid Mechanics* **876**, 527–552

Conference proceedings

Desmars, N., Perignon, Y., Ducrozet, G., Guérin, C.-A., Grilli, S. T. & Ferrant, P. 2018*b* Phase-Resolved Reconstruction Algorithm and Deterministic Prediction of Nonlinear Ocean Waves From Spatio-Temporal Optical Measurements. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 7B: Ocean Engineering, pp. 1–12. V07BT06A054

Desmars, N., Perignon, Y., Ducrozet, G., Guérin, C.-A., Grilli, S. T. & Ferrant, P. 2018*a* Phase-Resolved Prediction of Nonlinear Ocean Wave Fields from Remote Optical Measurements. In *16èmes Journées de l'Hydrodynamique*. Marseille, France

Remerciements

Tout d'abord, je tiens à exprimer ma reconnaissance à toutes les personnes ayant participé à initier le travail effectué dans cette thèse. Je remercie les membres du laboratoire commun JLMT (Joint Laboratory for Marine Technology) de l'École Centrale de Nantes et Naval Group de m'avoir donné la possibilité de travailler sur ce sujet dans le cadre du projet PREDEMO-Nav (PREdiction DEterministe des MOuvements de Navire sur houle, financements ANR-15-ASTR-0006). C'est dans ce contexte que j'ai eu grand plaisir à travailler avec Charles-Antoine Guérin de l'Institut Méditerranéen d'Océanologie de Toulon, Stéphan Grilli de l'Université du Rhode Island, et les équipes hydrodynamique/traitement du signal de Sirehna, avec qui les collaborations ont assurément enrichi le contenu de cette thèse.

Je remercie bien-sûr mes encadrants Yves Perignon, Guillaume Ducrozet et Pierre Ferrant de m'avoir soutenu dans cette aventure et prodigué de nombreux conseils d'ordre scientifique et organisationnel. Nos discussions ont été pour moi une source de motivation dans le travail, stimulant mes ambitions de rigueur et pertinence scientifiques. J'espère avoir emporté avec moi une partie de leurs enseignements ! Merci également à Félicien Bonnefoy pour son précieux support sur la partie expérimentale de cette thèse et pour son suivi d'une manière générale.

Aussi, je tiens à remercier les membres de mon jury d'avoir accepté d'en faire partie, ce fut un honneur d'avoir pu échanger sur mon travail avec chacun d'entre eux.

De même, j'exprime ici ma reconnaissance à tous ceux qui ont fait en sorte de rendre l'atmosphère du LHEEA conviviale. Il est bien difficile de rendre compte en quelques lignes de tous les bons moments passés au labo... Je tiens néanmoins à remercier : la team tupperware, dont la liste des membres a bien changée au cours des années (je laisse les personnes se reconnaitre par peur d'en oublier certaines), mais avec qui les midis ont toujours été un régal ; le bureau D233, toujours prêt à improviser une pause café au rythme des innombrables histoires de Marco, racontées avec un enthousiasme immodéré, nous permettant ainsi de profiter d'un certain folklore italien ; les motivés des sessions footing du midi au bord de l'Erdre, menées tant bien que mal par coach Vincent. Merci aussi à Marion pour les nombreuses discussions techniques sur la prédiction déterministe de houle !

Pour finir, je remercie mes amis et ma famille pour l'intérêt qu'ils ont porté à mon projet lorsque j'avais besoin d'en parler. Mention spéciale à Janine et Alicia pour avoir partagé mes derniers mois d'avant soutenance et m'avoir permis, malgré la pression, de profiter de ma nouvelle vie Hambourgeoise.

Résumé Substantiel

Dans l'environnement marin, les chargements induits par les ondes de gravité à la surface de l'océan (vagues) constituent l'un des principaux phénomènes naturels déterminant le mouvement des structures marines et leur fatigue structurelle. Ainsi, pour une large gamme d'applications marines, la disponibilité d'une prédiction vague-à-vague (ou déterministe) en temps réel du champ de vagues incident est cruciale pour leur optimisation. La prédiction et le contrôle du mouvement induit par les vagues jouent un rôle essentiel dans l'accroissement de l'enveloppe opérationnelle des opérations marines telles que la manœuvrabilité optimale des navires de surface, l'appontage d'aéronefs, le transfert de biens entre navires ou l'installation et la maintenance de structures offshores. Pour les systèmes de récupération d'énergies marines renouvelables (par exemple, un système houlomoteur, une éolienne flottante), la connaissance des vagues incidentes conditionne les performances des stratégies de contrôle. De telles prédictions déterministes des mouvements de la surface océanique reposent sur l'extraction des caractéristiques du champ de vagues à partir d'observations, c'est-à-dire des mesures contenant des informations relatives à l'évolution spatiale et/ou temporelle de la surface océanique. Un modèle de vagues, paramétrés en conséquence, est alors utilisé pour la propagation du champ vagues observé jusqu'à l'endroit d'intérêt, qui peut être mobile (comme le voisinage d'un navire).

Cette thèse prend part à un projet dont le but est de prédire, de manière déterministe et en temps réel, le mouvement d'un navire sur houle. La méthode généralement utilisée aujourd'hui pour effectuer de telles prédictions est basée sur l'analyse des mouvements passés sur navire, et permet à l'équipage d'anticiper les futurs mouvements à un horizon de quelques secondes. De leur côté, les prédictions reposant sur l'observation à distance de l'évolution de la surface permettent d'augmenter significativement l'horizon de prédiction pour atteindre quelques dizaines de secondes, soit quelques périodes caractéristiques de l'état de mer. Cependant, cela implique d'adresser trois difficultés majeures. La première est la mesure des grandeurs déterministes du champ de vagues (comme les élévations de surface, variations de pression, profiles de vitesse, etc.) permettant de procéder à une prédiction à un instant ultérieur et autour du navire qui peut être mobile. La deuxième difficulté est l'extraction d'informations relatives aux vagues (par exemple, amplitudes, phases) des mesures, dans le but d'initialiser un modèle physique et ainsi pouvoir propager dans l'espace et dans le temps le champ de vagues sous-jacent. Les choix du modèle de vagues et de la méthode d'extraction des informations relatives aux vagues pour son initialisation sont à effectuer en considérant la contrainte du temps réel : les vagues incidentes doivent être connues avant leur arrivées, avec un certain temps d'avance permettant à une stratégie de contrôle d'être exécutée ou à l'équipage d'utiliser un support de décision en cas d'opération délicate. La troisième difficulté est l'estimation de la réponse du navire à l'excitation du champ de vagues, elle aussi soumise à la contrainte du temps réel.

Les systèmes opérationnels fournissant des prédictions déterministes de champ de vagues en temps réel (comme ceux proposés par Hilmer & Thornhill (2015), Kusters *et al.* (2016) ou Naaijen *et al.* (2018)) utilisent des images de radars à bande X et des modèles physiques linéarisés. Ces choix proviennent de la capacités des radars nautiques à mesurer certaines propriétés locales de la surface océanique sur un large domaine (jusqu'à $\sim$ 3–4 km^2) autour de la structure sur laquelle ils sont montés, et de la contrainte du temps réel restreignant fortement l'utilisation

d'une modélisation complexe pour les processus physiques liés à la propagation vagues et à leur interaction avec la structure. Les prédictions ainsi faites sont donc limitées aux conditions pour lesquelles les processus physiques non-linéaires sont négligeables (comme en cas de mer peu agitée ou d'horizon de prédiction restreint), affectant leur fiabilité lorsque ces conditions ne sont pas vérifiées. Même si ces limitations peuvent être levées via l'utilisation de modèles de vagues plus fidèles (voir, par exemple, Blondel *et al.*, 2010; Klein *et al.*, 2020; Qi *et al.*, 2018*a*; Simanesew *et al.*, 2017), l'initialisation de tels modèles non-linéaires reste une difficulté majeure (voir, par exemple, Fujimoto & Waseda, 2020; Köllisch *et al.*, 2018; Yoon *et al.*, 2016) empêchant aujourd'hui leur utilisation pour la prédiction de champs de vagues en temps réel.

Cette thèse porte sur le problème de la prédiction de champs de vagues non-linéaires autour d'un navire en mouvement, à un coût calcul modéré, en vue d'atteindre un horizon de prédiction de quelques périodes caractéristiques de vagues. Comme présenté dans le premier chapitre, seules les mesures s'appuyant sur des capteurs permettant de mesurer à distance sont capables de fournir les informations nécessaires à la prédiction déterministe de champ de vagues autour d'une structure mobile. Quelques études portant sur les solutions technologiques permettant l'acquisition de mesures à distance se sont penchées sur les capacités des capteurs optiques, tels que des caméras lidar, à fournir des mesures d'élévation de surface (Belmont *et al.*, 2007; Kabel *et al.*, 2019). Grilli *et al.* (2011) et Nouguier *et al.* (2014) ont démontré l'exploitabilité de mesures telles que fournies par un lidar pour la prédiction déterministe de vagues non-linéaires. Dans la continuité de ces travaux, nous considérons dans cette thèse des mesures optiques similaires pour la paramétrisation du modèle physique décrivant l'évolution spatiale et temporelle des champs de vagues observés.

Les études de Grilli *et al.* (2011) et Nouguier *et al.* (2014) proposent également l'utilisation d'un modèle de vagues faiblement non-linéaire, basé sur une description Lagrangienne de l'évolution de la surface océanique afin de modéliser certaines propriétés non-linéaires avec une efficacité numérique proche de celle associée à la théorie linéaire. Cette approche est davantage développée et qualifiée dans cette thèse, et les performances d'une version améliorée sont étudiées dans le contexte de la prédiction déterministe.

Résumé des chapitres

Le premier chapitre de cette thèse commence par décrire la méthodologie classiquement utilisée pour la caractérisation des propriétés statistiques des vagues, puis les différentes méthodes de mesure envisageables pour la prédiction déterministe de champs de vagues sont présentées. Parmi elles, les caméras lidar ont plusieurs avantages – le principe de fonctionnement qui permet une mesure directe et à distance d'une grandeur d'intérêt (élévation de surface), l'échantillonnage spatial et temporel élevé, et les récents développements en vue de systèmes opérationnels – qui font d'elles une solution intéressante pour l'élaboration d'un nouvel algorithme de prédiction. Les méthodes de prédiction déterministes existantes sont ensuite présentées, en insistant sur les performances hydrodynamiques et computationnelles des modèles de vague utilisés et des méthodologies d'inversion de ces modèles à partir des mesures. Nous remarquons que la théorie linéaire – Linear Wave Theory (LWT) en anglais – permet une inversion aisée à partir de n'importe quel jeu de mesure de vagues, même si des précautions doivent être prises concernant le caractère non-uniforme et non-périodique des mesures. Cependant, les hypothèses relatives à l'approche linéaire restreignent son application à la prédiction de champs de vagues en cas de mer calme et pour un horizon de prédiction relativement court. Les modèles de vague faiblement non-linéaires permettent, jusqu'à un certain degré, d'améliorer la précision des prédictions si des corrections non-linéaires des vitesses de vague sont modélisées, ce qui nécessite un développement à l'ordre trois au moins pour les modèles Eulériens. Les approches complètement non-linéaires, comme la méthode High-Order Spectral (HOS), requièrent des procédures d'inversion complexes

qui ne sont pas facilement applicable à des données réelles et dans des temps compatibles avec le temps réel. De part son formalisme, l'approche Lagrangienne (Choppy Wave Model) permet de modéliser des propriétés non-linéaires avec un faible coût en calcul. N'ayant pas été étudiée dans le contexte de prédiction déterministe, la solution Lagrangienne au second ordre de perturbation, qui comprend une correction (partielle) de la vitesse des vagues, apparaît donc comme un bon compromis complexité/efficacité et sera étudiée par la suite.

Le deuxième chapitre présente l'approche Lagrangienne pour la modélisation des ondes de gravité à la surface de l'océan, ainsi que les propriétés intrinsèques à cette approches. La dérivation de la solution au premier ordre de perturbation est exposée, puis nous comparons, jusqu'au second ordre, les descriptions Lagrangienne et Eulérienne de la surface libre pour différentes configurations d'état de mer. La solution Lagrangienne inclut au second ordre des effets non-linéaires qui affectent à la fois la forme de la surface libre et la relation de dispersion (reliant fréquence et longueur d'onde des vagues). Le premier effet crée une asymétrie (c'est-à-dire des crêtes plus prononcées et creux plus plats), alors que le second engendre une modification de la vitesse de propagation des vagues. Nous introduisons ensuite une version améliorée de cette solution au second ordre, appelée Improved Choppy Wave Model (ICWM), qui corrige la relation de dispersion non-linéaire afin de mieux représenter les phases lors de la propagation du champ de vagues. Nous montrons que ICWM permet de retrouver la solution Eulérienne au troisième ordre pour une vague régulière. Dans le cas de champs de vagues irréguliers, ICWM corrige convenablement les composantes énergétiques du champ de vagues, mais manque les effets de modulation des vagues plus courtes. Afin d'utiliser ICWM lors de la reconstruction de champ de vagues à partir des mesures, nous en proposons une formulation Eulérienne pour l'élévation de surface. Nous montrons qu'une simple approximation dans un cadre Eulérien, ne nécessitant pas de développements mathématiques complexe ni de procédure itérative, permet d'obtenir des résultats consistants. Dans le but de valider l'intérêt de la formulation améliorée de la solution Lagrangienne, nous comparons divers résultats d'élévation de surface à une solution HOS complètement non-linéaire (c'est-à-dire convergée en ordre de non-linéarité). Il est montré que la correction non-linéaire de la vitesse de phase des vagues est cruciale pour une représentation correcte du champ de vagues après seulement quelques périodes de propagation. Ce travail a été publié dans Guérin *et al.* (2019). Pour finir, une caractérisation de quelques propriétés du champ de pression issu de ICWM est présentée.

Nous nous concentrons dans le troisième chapitre sur la procédure d'assimilation de données, consistant à accéder aux informations pertinentes contenues dans les mesures de vagues (appelées observations) afin de définir les conditions initiales du modèle physique servant à propager le champ de vagues en vue de sa prédiction. La méthode d'assimilation développée ici se base sur des mesures optiques telles que générées par une caméra lidar. La distribution de telles mesures est d'abord caractérisée au travers d'un approche géométrique, mettant en avant leur forte non-uniformité spatiale due à l'incidence rasante des rayons du capteur à la surface. La reconstruction du champ de vagues (c'est-à-dire l'inversion du modèle physique) est formulée pour le modèle non-linéaire ICWM étudié dans le chapitre précédent, puis nous détaillons le processus itératif de résolution. La formulation analytique de ICWM permet d'écrire le problème inverse sous une forme matricielle explicite, et la solution est trouvée en passant par une procédure de régularisation Tikhonov et une décomposition en valeurs singulières. Une fois le champ de vagues reconstruit, la région spatio-temporelle dans laquelle les informations assimilées restent valides pendant la propagation du modèle de vagues est bornée. À partir de l'étendue et de l'échantillonnage des observations, nous décrivons l'évolution de la zone de prédiction théoriquement accessible. La distribution spatio-temporelle d'une erreur théorique de prédiction est décrite à partir de la quantité d'énergie tronquée lors du processus de mesure, et montre que la taille de la zone de prédiction est fortement dépendante de la précision attendue de la prédiction. Des

critères sont ensuite présentés pour la sélection des fréquences et directions de coupure pour le champ de vagues modélisé ainsi que pour l'évolution de la zone de prédiction, nous permettant de fermer le problème inverse et définir clairement une région à l'intérieur de laquelle les performances de prédiction de notre méthode seront évaluées. Enfin, les questionnements relatifs à la détermination de la zone optimale de mesure (menant à l'horizon de prédiction souhaité) sont exposés et en partie traités. Depuis la caractérisation des mesures optiques d'élévation de surface jusqu'à la formulation du problème d'inversion de ICWM et la description de la zone de prédiction théoriquement accessible, ce chapitre termine l'explication des bases théoriques de la méthode de prédiction employée dans cette thèse.

Dans le quatrième chapitre, les propriétés et performances du modèle non-linéaire et de la méthode de prédiction développés sont étudiés à partir de jeux de données réalistes – bien que synthétiques et supposés sans erreur de mesure – de mesures optiques. Des simulations haute-fidélité de champs de vagues sont utilisés pour fournir des surfaces océaniques de références, à partir desquelles les observations sont générées. Après une description de la méthode de génération des mesures lidar synthétiques, nous étudions la dépendance de la précision des reconstructions/prédictions à divers paramètres d'assimilation, à savoir le nombre de composantes dans le modèle pour la représentation du champ de vagues, la fréquence de coupure haute de l'intervalle utilisé, la non-uniformité de la distribution des observations, la durée d'assimilation et l'étendue directionnelle des observations dans le cas d'un champ de vagues directionnel. Nous montrons que la précision de la reconstruction converge rapidement avec le nombre de composantes dans le modèle de vagues (quelques dizaines suffisent pour un champ de vagues unidirectionnel), ainsi qu'avec la fréquence de coupure haute. La caractéristique principale des mesures optiques, qui est de produire des données distribuées spatialement de manière très irrégulière, rend la reconstruction spatiale de champ de vagues particulièrement difficile. En revanche, inclure plusieurs jeux de mesures acquis à des instants différents dans l'inversion du modèle de vagues permet de contrer cette difficulté. Avec une procédure d'assimilation correctement paramétrée, l'effet de la prise en compte des propriétés non-linéaires de ICWM sur la qualité de la prédiction est analysé pour des champs de vagues de différentes cambrures caractéristiques. Il est montré que, malgré sa formulation analytique relativement simple, ICWM est capable de modéliser des effets non-linéaires menant à une amélioration non-négligeable de la prédiction. Comparée à celle de LWT, l'utilisation de ICWM dans l'algorithme de prédiction permet une réduction relative de l'erreur de prédiction de l'ordre de 25% pour la cambrure étudiée la plus élevée, soit $H_s/\lambda_p \approx 3.8\%$. Une large partie des résultats de ce chapitre est également présentée dans Desmars *et al.* (2018*a*,*b*).

Dans le cinquième et dernier chapitre, nous appliquons notre algorithme de prédiction à des données expérimentales issues d'une campagne menée dans le bassin océanique de l'École Centrale de Nantes. Des champs de vagues de différentes cambrures caractéristiques et/ou différents étalements directionnels ont été générés physiquement en bassin de houle, ainsi qu'à l'aide d'un bassin numérique basé sur le formalisme HOS, fournissant un jumeau numérique ("digital twin") du dispositif expérimental. Les sondes (physiques ou numériques, mesurant des séries temporelles d'élévation de surface en un point de l'espace) ont été non-uniformément distribuées en espace, afin de représenter l'échantillonnage typique d'une surface par un capteur lidar observant la surface avec un angle d'incidence rasant. Dans le but de valider la méthodologie de génération de données, l'adéquation des états de mer générés avec leurs caractéristiques théoriques (hauteurs significatives cibles) est d'abord vérifiée. Ensuite, le jeu de données correspondant aux cas de houle unidirectionnelle est analysé plus finement pour la quantification et la qualification physique des perturbations relatives aux conditions expérimentales. Au travers de l'examination des oscillations résiduelles mesurées par les sondes dans le bassin de houle réel, ces perturbations semblent être la conséquence de l'excitation par le batteur de modes propres transverses dans le bassin. Elles sont également identifiées comme la cause principale des différences observées

entre les données expérimentales et numériques. Nous évaluons par la suite, la sensibilité de l'algorithme de prédiction à des paramètres d'assimilation. Notre étude montre que la précision de la prédiction converge avec la durée des séries temporelles d'élévation de surface assimilées, c'est-à-dire la quantité de données utilisées pour effectuer la prédiction. Aussi, plus le pas de temps des données assimilée est faible, plus la convergence est rapide, et plus la cambrure caractéristique est élevée, plus l'erreur de prédiction augmente, sous la forme d'un déphasage non-linéaire des vagues prédites avec la solution de référence et d'une divergence de forme. Ces deux types de désaccord sont explicités au travers de la comparaison entre des élévations de surface pour un groupe de vagues spécifique. Les performances de ICWM relatives à LWT et LWT-CDR (LWT incluant une correction de phase similaire à celle de ICWM) sont ensuite quantifiées pour différentes configurations d'état de mer. Nous montrons que, pour l'estimateur d'erreur définie dans le cadre de cette étude, ICWM améliore les prédictions à l'intérieur de la zone de prédiction accessible, comparé à LWT, avec une réduction maximale de l'erreur de prédiction de $\sim 35\%$ pour une cambrure intermédiaire de $H_s/\lambda_p \approx 3.18\%$ (résultats numériques). Dans la région spatio-temporelle couverte par les observations, la correction non-linéaire de forme incluse dans ICWM améliore la représentation de l'élévation de surface. Pour les vagues qui se sont propagées (en espace ou en temps), le facteur principal de réduction d'erreur de prédiction est la modélisation correcte des vitesses de propagation des vagues. En revanche, en ce qui concerne la prédiction de pente, les propriétés de représentation de forme asymétrique des vagues de ICWM permet d'obtenir une réduction de l'erreur moyenne de 20% sur la zone de prédiction, pour une cambrure de $\sim$ 1–2% et des données numériques, en comparaison à une modélisation linéaire avec correction de la vitesse de propagation (LWT-CDR). Au travers d'une procédure entièrement validée expérimentalement, ces résultats démontrent la pertinence de l'utilisation de ICWM en vue de la prédiction de champs déterministes de la surface océanique. Les résultats présentés dans ce chapitre ont été publiés dans Desmars *et al.* (2020).

Perspectives

Le cœur de la méthode de prédiction proposée est l'inversion d'un modèle physique, consistant, dans notre cas, en la minimisation d'une fonction quadratique. La résolution du problème se fait au travers d'une décomposition en valeurs singulières à laquelle on adjoint une procédure de régularisation, et un processus itératif est utilisé pour l'évaluation les termes non-linéaires. Malgré le faible coût en calcul de la propagation en temps/espace d'un champ de vagues à l'aide de ICWM, la méthode d'inversion (nécessaire à l'initialisation du modèle), choisie pour sa robustesse et adaptation à la non-uniformité des observations, n'est pas numériquement efficace et représente la limitation principale à son utilisation pour des applications nécessitant le temps réel. Le temps de calcul de cette approche peut possiblement être modéré dans le cas de prédiction linéaire via la définition d'un système inverse mieux conditionné ou l'utilisation d'un environnement numérique hautes performances, mais l'estimation des non-linéarités reste hors de portée en temps réel. Étant données les propriétés numériques de ICWM, une reformulation du problème inverse utilisant efficacement ces propriétés pourrait significativement améliorer les performances numériques de l'algorithme, probablement au prix d'une procédure de pré-traitement des observations ou d'approximations hydrodynamiques supplémentaires. Plus généralement, les stratégies d'inversion rendant compatibles la non-uniformité et la non-périodicité des mesures de vagues avec des méthodes numériques rapides est un champ de recherche à approfondir, puisque cette incompatibilité représente la contrainte majeure pour l'initialisation en temps réel des modèles de vagues non-linéaires, y compris les modèles les plus efficaces à l'heure actuelle. Néanmoins, l'algorithme de prédiction non-linéaire proposé ici peut d'ores et déjà être utilisé dans des systèmes "off-line" à des fins d'analyse, donnant accès à une représentation de la surface plus précise que celle issue de l'approche linéaire. De plus, la limitation de la disponibilité de mesures

optiques de la surface, telles que produites par une caméra lidar, n'affecte pas l'applicabilité de la méthode développée, car elle peut être facilement adaptée à tout autre type de mesures de vague, qu'elles soient issues de capteurs in situ (par exemple une bouée) ou mesurant à distance (par exemple un radar).

Même si la méthode de prédiction a été développée dans un cadre entièrement compatible avec les champs de vagues directionnels, le travail présenté dans cette thèse traite principalement de champs unidirectionnels. Compte tenu du fait que les propriétés non-linéaires de ICWM ne sont pas affectées par le caractère directionnel du champ de vagues (c'est-à-dire qu'il n'y pas d'émergence de nouveaux phénomènes non-linéaires dans le cas d'une distribution directionnelle des vagues), les résultats étendus aux champs de vagues directionnels devraient être comparables avec ceux issus des champs unidirectionnels. Cependant, dans le but de saisir en profondeur le potentiel de ICWM dans des conditions réalistes, une étude dédiée à ses performances en prédiction de champs de vagues directionnels est nécessaire. À cette fin, la procédure d'analyse et les connaissances apportées par cette thèse sont d'un grand intérêt et doivent servir de base. La plupart des aspects de la méthode d'analyse utilisée ici peuvent être étendus sans difficulté aux champs de vagues directionnels. Les questions relatives à la quantification des indicateurs de précision des prédictions doivent être adressées prudemment afin d'éviter des temps de calcul excessifs dus au grand nombre requis de surfaces à analyser. Le jeu de données décrit dans le dernier chapitre, généré pour des états de mer directionnels et déjà en partie caractérisé, fournit un cadre approprié pour poursuivre la validation expérimentale.

Enfin, malgré l'application de la méthode de prédiction dans les conditions expérimentales de bassin de houle, qui fournit déjà une preuve de concept solide, l'utilisation de données acquises en milieu océanique est requise pour aboutir à une vue complète de son potentiel.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations and Acronyms

1D, 2D, 3D	one-, two-, three-Directional
ADCP	Acoustic Doppler Current Profiler
ANR	Agence Nationale de la Recherche
CWM	Choppy Wave Model
CWM1	first-order Choppy Wave Model
CWM2	second-order Choppy Wave Model
DFT	Discrete Fourier Transform
DHWM	Directional Hybrid Wave Model
ECN	École Centrale de Nantes
EMLM	Extended Maximum Likelihood Method
FFT	Fast Fourier Transform
GPS	Global Positioning System
HOS	High-Order Spectral
HWM	Hybrid Wave Model
ICWM	Improved Choppy Wave Model
IFFT	Inverse Fast Fourier Transform
JONSWAP	JOint North Sea WAve Project
lidar	LIght Detection And Ranging
LS	Linear Schrödinger
LWT	Linear Wave Theory
LWT-CDR	Linear Wave Theory with a Corrected Dispersion Relation
MNLS	Modified NonLinear Schrödinger
MTF	Modulation Transfer Function
NLS	NonLinear Schrödinger
NSR	Noise-to-Signal Ratio
PREDEMO-Nav	PREdiction DEterministe des MOuvements de Navire sur houle
RMS	Root Mean Square
SVD	Singular Value Decomposition

Mathematical Symbols

α_a, β_a	vertical and horizontal aperture angles of the optical sensor
$(a, b) = (A\cos\varphi, A\sin\varphi)$	wave model parameters
A	wave amplitude
c	phase velocity
c_g	group velocity
$\epsilon = kA$	wave steepness
η	free surface elevation
$E(\omega)$	angular frequency wave spectrum
$E(f)$	frequency wave spectrum
$E(k)$	wavenumber wave spectrum
f (f_p)	(peak) wave frequency
G	directional spreading function
g	acceleration of gravity
h	water depth
H (H_s)	(significant) wave height
J	number of spatial observations at each measurement instant
$\boldsymbol{k} = (k\cos\theta, k\sin\theta)$	wavenumber vector
$\hat{\boldsymbol{k}} = \boldsymbol{k}/k$	unit wavenumber vector
K	number of assimilated instantaneous spatial datasets
k (k_p)	(peak) wavenumber
λ (λ_p)	(peak) wavelength
L	total number of assimilated observations
L_o	spatial extent of the accessible prediction zone at reconstruction time
M	HOS order of nonlinearity
m_0	zeroth-order moment of the wave spectrum
N	number of wave components in the considered wave model
N_s	number of surface realizations
ω (ω_p)	(peak) wave angular frequency
$\boldsymbol{p}$	wave model parameters that are solution of the model inversion
$\mathcal{P}$	accessible prediction zone
φ	wave phase
p	pressure
$\boldsymbol{r} = (x, y)$	horizontal location
r	regularization parameter
ρ	water density
S	directional wave spectrum
θ (θ^{dir})	(main) wave direction of propagation
t	time
T (T_p)	(peak) wave period
T_a	assimilation time span
t_r	reconstruction time
τ	sampling period of the assimilated data
$\boldsymbol{\mathcal{U}}_s$	Stokes drift vector
$\boldsymbol{\xi}$	horizontal reference location for the Lagrangian expansion
x, y, z	space coordinates
x_c, y_c, z_c	space coordinates of the optical sensor

General Introduction

In the ocean environment, loads induced by surface gravity waves constitute one of the main natural phenomena that drive the motion of marine structures and their structural fatigue. Hence, for a vast range of marine applications, the availability of the real-time deterministic prediction of the incoming waves is key for their optimization. The prediction and control of wave-induced motion is of crucial importance to extend the operational envelope of operations such as optimal maneuvering of surface vessels, aircraft takeoff/landing, ship-to-ship transfer or installation and maintenance of marine structures. For ocean renewable energy harvesting systems, the knowledge of the incoming waves conditions the performances of control strategies. Such deterministic (or phase-resolved) ocean waves predictions rely on the retrieval of the wave field characteristics from a set of observations, i.e., measurements that contain information about the space and/or time evolution of the ocean surface. A parameterized wave model is then used for the propagation of the wave field to the region of interest, which can be mobile (e.g., a moving ship).

This thesis takes part in a project that aims at predicting, deterministically and in real time, the motion of a ship exposed to ocean waves. The currently widely used method to achieve such predictions is based on the analysis of the past ship motion, and allows the ship crew to anticipate the future motion during a horizon of prediction on the order of several seconds. In contrast, motion predictions based on the remote observation of the ocean surface are able to significantly increase the horizon of prediction to reach several tens of seconds, namely several characteristic wave periods of the sea state. This implies, however, to address three major issues. The first one is the measurement of the appropriate wave quantities (e.g., surface elevations, pressure variations, velocity profiles) to perform a prediction later in time and around a ship that can be moving. The second issue is the extraction of the relevant wave information (e.g., amplitudes, phases) from the measurements, in order to initialize a specific physical model for the space and time propagation of the underlying wave field. Both the choices of the wave information extraction method and of the wave model have to be made with respect to the constraint of real-time computation, i.e., we want to know the incoming waves before they actually arrive, with a certain time span in advance that allows a control strategy to be executed or the crew to use a decision support system during a sensitive operation. The third one is the calculation of the ship response from the wave-induced loads, also subjected to the real-time constraint.

Operational systems that provide real-time deterministic predictions of ocean wave fields (e.g., Hilmer & Thornhill, 2015; Kusters *et al.*, 2016; Naaijen *et al.*, 2018) rely on X-band radar images and linearized physical models. These choices come from the capacity of nautical radars to measure some local properties of the ocean surface over a large domain (up to $\sim$ 3–4 km^2) around the structure upon which they are mounted, and on the real-time constraint that strongly restrains from using a complex physical modeling of the wave propagation and wave/structure interaction processes. Hence, limitations arise in conditions for which nonlinear physical processes cannot be neglected (e.g., large prediction horizons, strong seas), affecting the reliability of the predictions. Even if these limitations could be eliminated (to a certain extent) by the use of more complex wave models (e.g., Blondel *et al.*, 2010; Klein *et al.*, 2020; Qi *et al.*, 2018*a*; Simanesew *et al.*, 2017), the initialization of such nonlinear models is still a major issue (e.g., Fujimoto & Waseda, 2020; Köllisch *et al.*, 2018; Yoon *et al.*, 2016), preventing their use for the

real-time prediction of wave fields.

This thesis focuses on the problem of the prediction of nonlinear wave fields around a moving structure at a moderate computational cost, in order to achieve a typical forecast horizon of several characteristic wave periods. As we will see in the first chapter of this thesis, only measurements that are done with remote sensing methods can provide the required wave information to deterministically predict wave fields around a moving structure. Past investigations on the technological solutions for the acquisition of remote wave measurements addressed the capabilities of optical sensors, such as lidar cameras, to provide surface elevation measurements (Belmont *et al.*, 2007; Kabel *et al.*, 2019). Grilli *et al.* (2011) and Nouguier *et al.* (2014) demonstrated the usability of lidar-like measurements for deterministic prediction of nonlinear ocean waves. Following the latter studies, we consider in this thesis similar optical measurements for the parameterization of the physical model that describes the time and space evolution of the observed wave fields.

Studies by Grilli *et al.* (2011) and Nouguier *et al.* (2014) also propose the use of a computationally efficient weakly nonlinear wave model, based on the Lagrangian description of the ocean surface evolution, to represent some nonlinear features with a numerical efficiency that is close to that belonging to the linear theory. This approach is further developed and qualified in this thesis, and the prediction performances of an improved version is investigated throughout the presented work.

Thesis summary

Chapter I first describes the classical methodology to characterize ocean waves statistical properties, and the wave measurement methods that are used for deterministic prediction. It is seen that lidar cameras have inherent qualities – the nature of their working principle that allows a direct measurement of wave quantities (free surface elevations), their high sampling rate (in both space and time) and the recent developments of operational systems – that make them a relevant choice for the investigation of new prediction algorithms. We then review the existing deterministic ocean waves prediction approaches, emphasizing on both the hydrodynamic and computational performances of the used wave models and methodologies to invert them from the wave measurements. We see that the Linear Wave Theory (LWT) is the most convenient model for inversion from any set of wave measurements, even if precautions have to be taken regarding non-uniformity and non-periodicity of the measured data. It appears, however, that the hypotheses pertaining to this modeling approach restrict its application to the prediction of wave fields of small steepness and of relatively short prediction horizons. Weakly nonlinear wave models are able to significantly improve the prediction accuracy if nonlinear corrections of wave velocities are taken into account. Fully nonlinear approaches, such as the High-Order Spectral (HOS) method, require complex model inversion procedures that are not easily applicable to real wave data and computable in real-time. In this context, the weakly nonlinear Lagrangian approach appears to be a convenient method to model relevant nonlinear wave properties (wave velocity corrections) at a low computational cost thanks to a simple analytical formulation of the free surface elevation.

Chapter II presents the Lagrangian approach for ocean gravity waves modeling and some of its intrinsic properties. We detail the derivation of the first-order solution, and show how the Lagrangian approach (up to the second-order) compares to its Eulerian counterpart in terms of free surface description. The second-order Lagrangian solution includes nonlinear effects that both affect the shape of the free surface and the dispersion relation. The former creates wave shape asymmetry (i.e., sharper crest and flatter trough), while the latter results in a modification of the wave velocity. Then, we introduce the improved Lagrangian second-order solution, referred to

as Improved Choppy Wave Model (ICWM), that corrects the nonlinear dispersion relation modeled by the classical second-order Lagrangian solution. It is shown that ICWM allows retrieving the third-order Eulerian solution for regular waves. In the case of irregular waves, it leads to an appropriate correction for energetic wave components but misses the modulation effects of shorter waves. We then provide an Eulerian form of the corresponding free surface elevation that is required for the reconstruction of the wave field from the wave measurements. We show that a simple Eulerian approximate that does not necessitate complex mathematical developments nor iterative procedures can be formulated and leads to consistent results. In order to validate the relevance of the "improved" formulation of the Lagrangian solution, we compare various results of free surface elevation to a fully nonlinear (i.e., converged in order of nonlinearity) HOS solution. It is found that the nonlinear correction of the phase velocity is crucial for accurate wave representation after several periods of propagation. Note that this specific work has been published in Guérin *et al.* (2019). A characterization of some properties of ICWM through the derivation and computation of the wave induced pressure field is finally presented.

Chapter III focuses on the data assimilation process, which consists in getting access to relevant information from the wave measurements (called observations) to properly define the initial conditions of our wave model. The developed assimilation method is based on typical optical measurements, such as a lidar camera observing the ocean surface would make. First, the distribution of lidar-like measurements is characterized through a geometrical approach, highlighting the strong spatial non-uniformity due to the grazing incidence of the sensor's rays on the surface. Then, the wave field reconstruction (i.e., wave model inversion) problem is formulated for the nonlinear wave model ICWM studied in the previous chapter, and we detail the iterative solving process. The analytical formulation of ICWM allows us to write the inverse problem in an explicit matrix form, and the solution is found using a Tikhonov regularization procedure and a Singular Value Decomposition. Once the wave field is reconstructed, the spatio-temporal zone within which the assimilated information remains valid during the model propagation is bounded. Based on the extent and sampling characteristics of the observations, we formally describe the evolution of the theoretically accessible prediction zone. The spatio-temporal distribution of a theoretical prediction error is derived from the amount of truncated energy by the accessible measured information, and shows that the extent of the prediction zone strongly depends on the expected prediction accuracy. Criteria are presented for the selection of the frequency and direction bandwidths for the modeled wave field as well as for the prediction zone evolution, allowing us to close the wave-model inverse problem and to define a clear region within which the prediction performance will be evaluated. Finally, issues related to the determination of the optimal measurement zone (that leads to the desired prediction horizon) are exposed and partly addressed. From the characterization of remote optical free surface observations to the derivation of the ICWM inversion and the description of the accessible prediction region, this chapter completes the theoretical basis of the prediction method employed in this thesis.

Chapter IV discusses the capabilities of the developed nonlinear wave model and prediction method from non-uniform datasets of realistic – but synthetic and error free – optical measurements. High-fidelity numerical simulations of wave fields are used to provide reference ocean surfaces, from which the observations are generated. After a brief description of the method for the generation of synthetic lidar observations, we study the dependence of the proposed prediction algorithm, in terms of reconstruction/prediction accuracy, on some assimilation parameters, namely the number of wave components in the wave models for the wave field representation, the high cutoff boundary of the frequency bandwidth, the non-uniformity of the observations' distribution, the assimilation time, and the directional extent of spatial observations in case of short-crested waves. We found that the reconstruction accuracy converges quickly with the number of wave components (a few dozens are enough for long-crested waves), as well as with the

high cutoff frequency. The main characteristic of optical measurements, which is to yield highly sparsely distributed datasets, makes for a particularly challenging spatial reconstruction, but including several spatial datasets acquired from different observation times in the wave model inversion allows overcoming this issue. For a properly parameterized assimilation procedure, the impact of the inclusion of the nonlinear properties of ICWM on the prediction accuracy is then investigated for different characteristic wave steepnesses. It is shown that, despite its simple analytical formulation, ICWM is able to model relevant nonlinear effects in ocean wave fields that impact the accuracy of the wave prediction. The improvements using ICWM, as compared to LWT, reaches a relative error reduction of 25% for the largest steepness investigated, i.e., $H_s/\lambda_p \approx 3.8\%$. A significant part of the results from this chapter is also presented in Desmars *et al.* (2018*a*,*b*).

Chapter V investigates the application of our wave reconstruction and prediction algorithm to experimental data generated in the ocean wave tank of École Centrale de Nantes. Sea states of different characteristic wave steepness and/or directional spreading have been generated physically in a wave tank, as well as using a numerical wave tank based on the HOS formalism, providing a digital twin of the experimental setup. A number of test cases were defined with different directional spreading and characteristic wave steepness, from mildly to more strongly nonlinear. The set of wave probes (physical or numerical, measuring time series of surface elevation at one point in space) was non-uniformly distributed in space, to mimic the typical uneven sampling made by a lidar camera at grazing incidence. In order to validate the methodology for the data generation, the adequacy between the generated sea states and their theoretical characteristic quantities (namely, significant wave heights) is first verified. Then, the dataset pertaining to cases of long-crested waves is further analyzed for the quantification and physical qualification of the perturbations belonging to the experimental conditions. Through the examination of the residual oscillations measured by the physical probes, these perturbations are found to be the consequence of the excitation of transverse modes of the wave tank by the wavemaker. They are identified as the principal cause for the observed differences between the experimental and numerical data. Next, the sensitivity of the developed prediction algorithm to assimilation parameters is evaluated. Our study show that the prediction accuracy converges with the duration of the assimilated surface elevation time series (i.e., amount of data used). In addition, smaller data acquisition time steps yields higher convergence rates, and the larger the characteristic wave steepness, the larger the prediction error, in the form of nonlinear phase shifts (related to wave phase velocity) and wave shape discrepancies. These two type of disagreements are highlighted by the comparison of predicted and reference surface elevation profiles of a specific wave group. After that, we assess the performances and relative improvements of ICWM compared to LWT and LWT-CDR (LWT with wave velocity correction similar to that of ICWM) under different sea state configurations. We show that, for the prediction misfit indicator defined for this study, ICWM yields improved predictions within the accessible prediction zone, as compared to LWT, with a maximum prediction error reduction of $\sim 35\%$ for an intermediate wave steepness $H_s/\lambda_p \approx 3.18\%$ (based on numerical data). In the spatio-temporal region corresponding to the observations, the wave shape asymmetry represented in ICWM improved the surface elevation representation. For waves that had propagated (in space and/or time), the main factor for reducing the prediction misfit is the accurate modeling of nonlinear wave phase velocity. For surface slope predictions, however, the improved representation of wave shape asymmetry in ICWM allows to achieve an average prediction misfit reduction of 20% over the prediction zone, for a low wave steepness of $\sim$ 1–2%, as compared to a linear model with a phase speed correction (LWT-CDR). Through a procedure that has been fully validated experimentally, these relative improvements demonstrate the benefits of using ICWM for the prediction of deterministic fields of the ocean surface. Results presented in this chapter have been published in Desmars *et al.* (2020).

Résumé du premier chapitre

Le premier chapitre de cette thèse commence par décrire la méthodologie classiquement utilisée pour la caractérisation des propriétés statistiques des vagues, puis les différentes méthodes de mesure envisageables pour la prédiction déterministe de champs de vagues sont présentées. Parmi elles, les caméras lidar ont plusieurs avantages – le principe de fonctionnement qui permet une mesure directe et à distance d'une grandeur d'intérêt (élévation de surface), l'échantillonnage spatial et temporel élevé, et les récents développements en vue de systèmes opérationnels – qui font d'elles une solution intéressante pour l'élaboration d'un nouvel algorithme de prédiction. Les méthodes de prédiction déterministes existantes sont ensuite présentées, en insistant sur les performances hydrodynamiques et computationnelles des modèles de vague utilisés et des méthodologies d'inversion de ces modèles à partir des mesures. Nous remarquons que la théorie linéaire – Linear Wave Theory (LWT) en anglais – permet une inversion aisée à partir de n'importe quel jeu de mesure de vagues, même si des précautions doivent être prises concernant le caractère non-uniforme et non-périodique des mesures. Cependant, les hypothèses relatives à l'approche linéaire restreignent son application à la prédiction de champs de vagues en cas de mer calme et pour un horizon de prédiction relativement court. Les modèles de vague faiblement non-linéaires permettent, jusqu'à un certain degré, d'améliorer la précision des prédictions si des corrections non-linéaires des vitesses de vague sont modélisées, ce qui nécessite un développement à l'ordre trois au moins pour les modèles Eulériens. Les approches complètement non-linéaires, comme la méthode High-Order Spectral (HOS), requièrent des procédures d'inversion complexes qui ne sont pas facilement applicable à des données réelles et dans des temps compatibles avec le temps réel. De part son formalisme, l'approche Lagrangienne (Choppy Wave Model) permet de modéliser des propriétés non-linéaires avec un faible coût en calcul. N'ayant pas été étudiée dans le contexte de prédiction déterministe, la solution Lagrangienne au second ordre de perturbation, qui comprend une correction (partielle) de la vitesse des vagues, apparaît donc comme un bon compromis complexité/efficacité et sera étudiée par la suite.

Chapter I

Existing Approaches for Deterministic Ocean Waves Prediction

Introduction

In this first chapter, we propose to review the existing deterministic ocean waves prediction algorithms. After defining the classical methodology to characterize ocean waves statistics and the wave measurement methods that are used for deterministic prediction, we then focus on different approaches for phase-resolved wave modeling. Each measurement method and wave model implies constraints on the prediction algorithm in terms of performance and computational effort. In the following, we highlight those constraints and describe the existing approaches to address them.

I.1 Waves Statistical Characterization

We briefly remind in this first section the general method to study ocean waves and identify their statistical properties. In many cases, ocean waves are analyzed by means of the power spectral density of the surface elevation. This so-called wave spectrum is largely dominated by surface gravity waves and, even if waves are highly irregular, a typical wave spectrum has a characteristic time scale of several dozens of minutes to a few hours to exhibit significant variations. By this way, a wave field is reduced to a *sea state* characterized by the properties of the directional wave spectrum, which describes the way the variance (or equivalently the energy) is spread over wave frequencies f and directions of propagation θ. In the following, we will use the term "long-crested waves" to refer to a purely unidirectional wave field (i.e., all waves propagate along the same direction), while the term "short-crested waves" refers to a wave field exhibiting a directional distribution of the waves' direction of propagation. The most widely used variable is the elevation variance spectrum $S(f, \theta)$ (unit: $\mathrm{m}^2.\mathrm{Hz}^{-1}$), which is frequently (and abusively) called the "wave energy spectrum" in the literature. In the following, we will adopt this practice, even if the true energy spectrum (unit: $\mathrm{N.m}^{-1}.\mathrm{Hz}^{-1}$) is in fact $\rho g S$, with ρ the water density and g the acceleration of gravity.

Usually, a sea state is characterized through the identification and separation of spectral components in the measured wave spectrum. Each component is supposed to represent a specific type of sea (e.g., swell, wind waves) that can then be described by a theoretical model with appropriate parameters related to statistical properties of the wave spectrum. This methodology further simplifies the description of wave fields and allows for predictive numerical modeling. In ocean engineering, the two main variables used to parameterize a wave spectrum are the significant wave height H_s and the dominant wave frequency f_p. With only a few percents of differences, the significant wave height can be compared to $H_{1/3}$, the mean height of the

highest third of the waves (calculated from the wave height distribution). It is also close to the instinctive estimate of the wave height by a human being staring at the sea surface. It is defined as $H_s = H_{m0} = 4\sqrt{m_0}$ where m_0 is the zeroth-order moment of the wave energy spectrum following

$$m_0 = \int_0^{+\infty} E(f)\,\mathrm{d}f, \quad \text{where} \quad E(f) = \int_0^{2\pi} S(f,\theta)\,\mathrm{d}\theta,$$

is the frequency spectrum. The dominant (or peak) frequency $f_p = \arg\max_f E(f)$ is the frequency of maximal energy. For the sake of simplicity, the directional wave spectrum is often decomposed as the product of $E(f)$ and a directional spreading function $G(f,\theta)$ having the properties of a probability density function, as

$$S(f,\theta) = E(f)\,G(f,\theta), \tag{I.1}$$

in which

$$G(f,\theta) \geq 0, \quad \theta \in [0,2\pi], \quad \text{and} \quad \int_0^{2\pi} G(f,\theta)\,\mathrm{d}\theta = 1.$$

The standard Pierson-Moskowitz frequency spectrum (Pierson Jr. & Moskowitz, 1964), describing fully-developed sea (i.e., produced by wind blowing steadily over a large area and over a long time), only uses f_p and an estimate of the wind velocity in the atmospheric boundary layer (which can be directly related to H_s) as input parameters. The more widely used JONSWAP (JOint North Sea WAve Project) spectrum (Hasselmann *et al.*, 1973) can be parameterized with f_p, H_s and a peak enhancement factor γ to model non-fully developed seas. Since this JONSWAP spectrum is extensively used in the applications presented in this work, we explicit here its formulation. It follows

$$E(f) = \alpha_s \frac{g^2}{f^5} \exp\left[-\frac{5}{4}\left(\frac{f_p}{f}\right)^4\right] \gamma^{\exp\left[-\frac{(f-f_p)^2}{2\sigma_s^2 f_p^2}\right]}, \qquad \sigma_s = \begin{cases} 0.07, & \text{if } f \leq f_p, \\ 0.09, & \text{if } f > f_p, \end{cases} \tag{I.2}$$

where g is the gravity acceleration, α_s is a tuned coefficient to ensure $m_0 = H_s^2/16$. To obtain a realistic model of a directional spectrum, the frequency spectrum is multiplied by a directional spreading function, such as the standard cosine squared function (e.g., Socquet-Juglard *et al.*, 2005)

$$G(\theta) = \begin{cases} \frac{1}{\nu}\cos^2\left[\frac{\pi\left(\theta-\theta^{\mathrm{dir}}\right)}{2\nu}\right], & \text{if } \left|\theta-\theta^{\mathrm{dir}}\right| \leq \nu, \\ 0, & \text{if } \left|\theta-\theta^{\mathrm{dir}}\right| > \nu, \end{cases} \tag{I.3}$$

where θ^{dir} is the main wave direction of propagation and ν is a spreading parameter. This way, the wave energy spectrum is simply the multiplication of two independent functions describing the frequency and the directional energy distributions.

Figure I.1 presents an example of a measured directional wave spectrum at the open-ocean test site of École Centrale de Nantes (ECN), SEM-REV, a 1 km^2 grid-connected area, located 20 km off from the coast from Le Croisic. We identify two wave systems, both swell, characterized by peaked energy distributions in direction and frequency. One of the main directions corresponds to $\theta^{\mathrm{dir}} \approx 270°$, the other slightly differs and is about 260°. One of the peak frequencies is $f_p \approx 0.073$ Hz ($T_p = 1/f_p \approx 13.7$ s), the other is about 0.05 Hz, or 20 s in period.

Depending on the measurement method, numerical model or application, it might be more suitable to perform the spectral analysis using the wavenumber k instead of the wave frequency

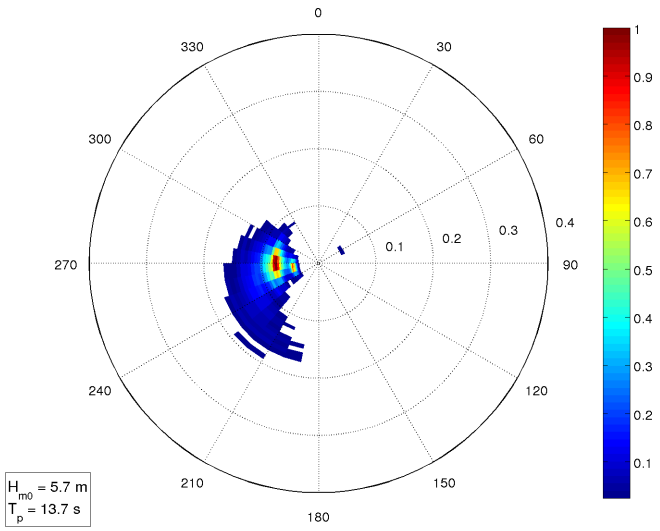

Figure I.1: Example of a directional wave spectrum at SEM-REV test site (ECN) on January, 6$^{\text{th}}$ 2014 at 12 a.m.

f. Since the total wave energy must not depend on which variable is used, it is possible to write

$$\int_k \int_\theta S(k,\theta)\,\mathrm{d}k\,\mathrm{d}\theta = \int_f \int_\theta S(f,\theta)\,\mathrm{d}f\,\mathrm{d}\theta,$$

from which we deduce the relation

$$S(k,\theta) = S(f,\theta)\,\frac{\partial f}{\partial k} = S(f,\theta)\,\frac{c_g}{2\pi},$$

where $c_g = \partial\omega/\partial k$ is the wave group velocity (i.e., the velocity of the wave energy transport), and $\omega = 2\pi f$ is the angular frequency.

According to the linear wave theory (i.e., physical model derived with the hypothesis of waves of infinitely small steepness, see section I.3.2 for details), the free surface elevation η of any (linear) wave field can be represented by a linear superposition of $n = 1, ..., N$ elementary sine waves of wavenumbers k_n, directions of propagation θ_n , amplitudes A_n and phases φ_n following

$$\eta(x,y,t) = \sum_{n=1}^{N} A_n \cos\left[k_n\left(x\cos\theta_n + y\sin\theta_n\right) - \omega_n t - \varphi_n\right],$$

in which the angular frequencies and wavenumbers respect the linear dispersion relation $\omega^2 = gk\tanh(kh)$ (where h is the water depth). Now, assuming that φ_n are random phases uniformly distributed over $[0, 2\pi]$, the amplitudes of the wave components are related to the directional wave spectrum through

$$S(f_n,\theta_n) = \lim_{\Delta f\to 0}\lim_{\Delta\theta\to 0}\frac{1}{\Delta f\Delta\theta}\left(\frac{1}{2}A_n^2\right), \tag{I.4}$$

where Δf and $\Delta\theta$ are the spectral resolutions in frequency and direction. Note that it is not possible to reconstruct the surface of the wave field directly from its spectrum, since it misses phase information. However, assuming that the elevation is a Gaussian process, random phases uniformly distributed can be used to generate a sea surface realization that is statistically similar to the wave field described by the wave spectrum.

I.2 Wave Measurement Methods

Wave spectra are calculated from space/time series of surface elevation measurements. We classify here the wave measurement systems in two categories: in situ measurement systems, which are directly in contact with the ocean, and remote measurement systems, which use electromagnetic waves that interact with the water and remotely give access to physical characteristics of the ocean. We restrict the scope of this section to solutions that can be installed in the open ocean and give access to instant surface elevation measurements, which are the conditions in case of applications to deterministic wave field reconstruction around offshore structures. More detailed information related to wave measurement systems are proposed by Benoit *et al.* (1997), Barstow *et al.* (2005), Neill & Hashemi (2018) or Ardhuin (2019). Measurement methods adapted to ship motion prediction are reviewed by Giron-Sierra & Jimenez (2010).

I.2.1 In Situ Technologies

Here, we focus on systems that are installed into the ocean, directly interacting with the water to measure wave properties that can be used to retrieve the surface elevation.

I.2.1.1 Wave Buoys

Wave buoys are single-point systems that are the most widely used instruments to measure waves at the ocean surface. Wave buoys (moored to the seabed or freely drifting) equipped with an on-board accelerometer can deliver, based on an estimate of the horizon to define the vertical and after a double time integration of the vertical acceleration, time series of the heave (i.e., vertical) motion of the buoys at their specific location. Modern systems are able to send their data via satellite communication. Recent systems rely on GPS (Global Positioning System) instead of accelerometers to measure the buoy's motion, which reduces the cost and ensure a long-term calibration of the instrument, despite a higher energy consumption and the need of continuous connection with GPS satellite.

Directional information can be obtained through the measure of the two components $\partial z/\partial x$ and $\partial z/\partial y$ of the surface slope or the two horizontal motion in x (surge) and y (sway) of the buoy. While horizontal motion can be recorded via accelerometers or a GPS, the surface slope estimation necessitates inclinometers to measure the tilt of the buoy. The former can be measured by small buoys (e.g., Datawell $<$ 1-m diameter buoys), whereas the latter is measured by larger models (e.g., US National Data Buoy Center 3-m diameter buoys). For non GPS-based wave motion sensor, another instrument (e.g., magnetic compass) is used to monitor direction of north in order to geographically locate the wave direction. The determination of the directional spectrum is done using statistical methods such as Maximum Likelihood Method (Isobe *et al.*, 1984; Krogstad, 1988), Maximum Entropy Principle (Nwogu, 1989) or Bayesian Directional Method (Hashimoto & Kobune, 1988).

Figure I.2 shows an operational wave buoy at the SEM-REV test site, which measures the wave height, period and direction of propagation, as well as the water temperature at the surface. The accuracy of this widely used system, the Datawell Waverider^{TM}, is advertised as follows: the mean error of the heave motion measurement is 0.5% of the measured value, and the mean heading error of the wave direction is 0.5°, for wave periods ranging between the limiting values of 1 s and 30 s. Note that in real operating conditions, this precision can be reduced due to the influence of the ocean environment (e.g., biofouling, currents). If additional precision about the wave field is required, such as the distinction between incident and reflected waves (for instance, upstream a wave energy converter to quantified its radiation/diffraction properties), a network a several sensors at different locations becomes necessary.

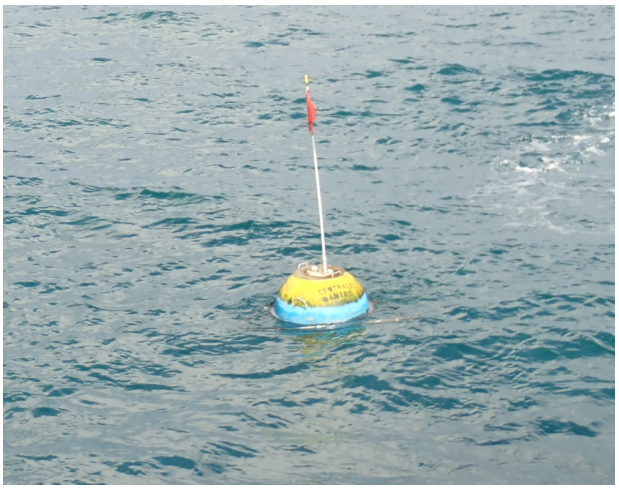

Figure I.2: Datawell Waverider MKIII^{TM} buoy installed at SEM-REV, measuring wave height, period and direction of propagation using accelerometers. It also measures the water temperature at the surface. Credit: ECN.

I.2.1.2 Acoustic and Pressure Sensors

Instead of relying on the motion of a moving structure, other instruments relate pressure variations or Doppler effects of acoustic waves to ocean wave properties in order to estimate the surface elevation. These instruments are either bottom mounted or fixed on a submerged floating platform, thus less subjected to strong current and storm waves than floating buoys.

Pressure transducers retrieve the surface elevation from a pressure signal and a transfer function based, for instance, on the linear wave theory (see details in section I.3.2) (e.g., Bishop & Donelan, 1987; Tsai *et al.*, 2005) or on recently developed methods that include nonlinear physics (e.g., Bonneton & Lannes, 2017; Bonneton *et al.*, 2018). However, the limited extent of the pressure variations induced by surface waves in the water column strongly impacts the sensibility of the instrument to wave-induced pressure variations, especially from high frequency components. If mounted on the seabed, these systems are thus limited to shallow water area (typically with less than 20-m water depth). Pressure transducers provide a measurement of one quantity at only one point in space, thus cannot be used to calculate a directional wave spectrum. However, the simultaneous use of several sensors allows capturing directional phenomena and calculate the directional spreading function $G(f, \theta)$. For instance, Howell (1998) describes and tests an array of three pressure transducers deployed in shallow water depth to derive time series spectral parameters such as significant wave heights, peak frequencies and mean directions of propagation for each frequency.

Acoustic Doppler Current Profilers (ADCPs) are used to estimate the three-dimensional velocity profile in the water column. Based on the measurement of the reflected echos of acoustic signals by water particles, the Doppler drift from multiple divergent beams is estimated. Under the hypothesis of homogeneous horizontal velocity within the cone formed by the beams, the Doppler shift gives access to the full three-dimensional profile at many locations in the water column. Similar to the pressure field, the surface elevation can be retrieved from the estimate of the vertical velocity based on a transfer function, generally based on the linear wave theory. Despite the fact that only three beams are needed for the velocity profile reconstruction process, ADCPs are usually equipped with four beams, each titled with an angle of $\sim 25°$ with respect to the vertical (the "Janus" configuration), to provide an estimate of the measurement error. However, even if the homogeneous velocity hypothesis is valid to evaluate water displacement in the case of currents, it may be no longer reliable for waves orbital velocities for which the space variation rate exceeds the limitation imposed by the distance between the beams, espe-

cially for short waves and for long measuring distances from the ADCP. This prevents from accurately separating the horizontal and vertical components of the measured velocities, but, under the assumption that the wave field is statistically stationary in time, the cross-spectra between velocities measured at various locations contain relevant information about wave direction. Terray *et al.* (1997) employed the Maximum Likelihood Method to determine directional wave spectra from such measurements, and found consistent results with spectra calculated from a co-located array of pressure transducers. Jeans *et al.* (2003) also found good agreement between ADCP-based spectral wave parameters and similar quantities derived from directional wave buoy measurements. Work (2008) further reports comparisons between wave spectra calculated from ADCP and directional wave buoy measurements and the Maximum Entropy Method, and highlighted the dependence of the wave data quality on the type of system used.

Still, the homogeneous velocity hypothesis limits the physical description of the water kinematics related to short waves, and the lack of direct measurement of surface elevation time series prevents from accurately calculating extreme value statistics. Terray *et al.* (1999) showed that the use of surface tracking through echo-ranging (i.e., detection of the time for the acoustic signal to reflect on the water-air interface and get back to the ADCP) for the direct determination of the free surface displacement is possible and that the derived surface elevation spectra compare well to those computed from ADCP velocity measurements or from an array of pressure transducers. Accordingly, to measure waves, some ADCPs include a vertically oriented beam, dedicated for surface tracking, such as the widely used AWAC (Acoustic Wave And Current Profiler) system developed by Nortek^{TM} and showed by Pedersen *et al.* (2002) to improve the surface elevation spectrum accuracy (better resolution of short waves) and completeness (possibility to calculate the distribution of wave heights) compared to previous ADCPs systems relying on velocity measurements. The benefits of a (vertical) dedicated beam for surface tracking to calculate the surface elevation spectrum is confirmed by Birch *et al.* (2004) who compared the performances of a similar AWAC system and an ADCP that uses the standard Janus configuration to measure velocity profiles as well as surface elevations through echo ranging. The advantage of the vertical beam configuration for the determination of the vertical velocity component in the water column is demonstrated by Wanis (2013) using comparisons with velocity profiles derived from inclined beams. Bouferrouk *et al.* (2016) also provide consistent conclusions about the advantages of including a surface tracking with a vertically oriented beam for the wave parameter estimates.

I.2.1.3 Wave Probes

Wave probes (also called wave gauges) is the most frequently used system to directly measure the surface elevation in controlled environment (i.e., laboratory wave tanks). By means of the measurement of the resistance or capacity variation related to the submerged portion of the probe, a very accurate estimate of the free surface elevation is calculated. Another widely used wave probe system relies on ultrasonic waves to measure the distance between a transceiver and the free surface elevation. Even if this technique is non-intrusive, it is not considered as a remote sensing technique since the transducer needs to be located close to the surface. Similar to pressure transducers, a network of several (at least three) wave probes is necessary to calculate of the directional wave spectrum.

I.2.2 Remote Sensing Technologies

When wave measurements have to be made along a specific trajectory, e.g., at the location of a moving ship, in situ measurement systems become not usable since they mainly depend on non-moving structures to operate (e.g., mooring lines). Moreover, in situ measurement systems would need to be far from the source of motion (e.g., a ship) in order to be unaffected by its scattered wave field. Instead, remote systems measure at a distance the scattering characteristics of electromagnetic waves by the water surface. Remote sensing instruments can be divided into

two categories: passive systems, that only rely on natural sources of electromagnetic waves emission (e.g., the sun) and active systems, that emit pulses with specific characteristics and measure the backscattered signal.

One major advantage of remote sensing instruments over in situ systems is that they provide a measurement of directional surface elevation fields $\eta(x, y, t)$, from which the directional wave spectrum can be calculated directly, without requiring a specific wave model, ensuring a correct spectral representation of every directional wave component.

I.2.2.1 X-Band Radar

X-band radars are mounted on many vessels and offshore platforms for navigation, ship traffic control and target identification purposes. They operate in the X-band microwave regime, corresponding to the frequency range 8–12 GHz of the electromagnetic spectrum. The radar backscattered signal includes echoes from structures such as ships and coastlines, but also includes an ensemble of echoes by the ocean surface called the sea clutter (figure I.3). Although

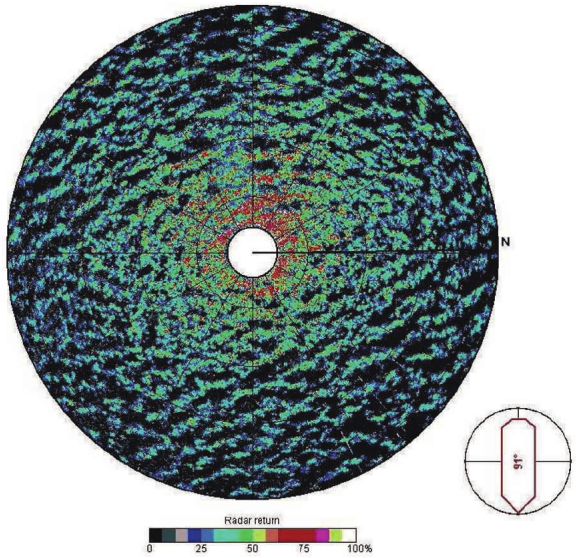

Figure I.3: Example of the backscatter signal from an X-band radar (Giron-Sierra & Jimenez, 2010). Tilt modulation (intensity variations of the radar backscattered signal depending on the incidence angle) and wave shadowing effects (regions without return signal in black) are clearly noticeable, suggesting the ocean surface deformation.

the sea clutter is a disturbing signal for navigation applications, it contains information that can be used to characterize the surrounding surface wave field (Alpers *et al.*, 1981). Sea-clutter echoes are generated by the Bragg resonance between the radar microwaves ($\sim$ 3 cm wavelength) and short-wavelength capillary-gravity waves ($\sim$ 1.5 cm wavelength) covering the ocean surface due to wind generation. Hydrodynamic and tilt modulations of such short ripples by longer gravity waves carrying them, that affect the backscattered signal, allow inverting for surface elevations by means of a Modulation Transfer Function (MTF) (Dankert & Rosenthal, 2004; Nieto Borge *et al.*, 2004; Young *et al.*, 1985). Moreover, since the angle of incidence radar pulses is grazing, a wave shadowing phenomenon occurs, i.e., the wave troughs are hidden behind the crests, improving long waves identification, therefore helping the wave field inversion through the implementation of a shadowing mask in the MTF. This way, radars are able to generate large spatio-temporal instantaneous datasets of wave elevations surrounding the structure upon which they are mounted, with a typical space resolution (limited by their range resolution) of about 5–10 m at sampling frequency 0.5–1 Hz. This technology has been implemented in commercial products such as WaMoS II developed by OceanWaveS GmbH^{TM} (Hilmer & Thornhill, 2015), the

prediction systems of Next Ocean™ (Naaijen *et al.*, 2018), or FutureWaves™ (Kusters *et al.*, 2016). A selection of radar images, together with the dispersion relation, can be used to estimate the directional wave spectrum. However, the estimate of the surface elevation, both deterministically and for integral parameters, directly from the gain of radar images is not reliable due to the unknown MTF over the specific wave field observed. Improvement in wave height estimation can be done using an additional measurement of, for instance, the amplitude of the ship motions, or from a finer analysis of the shadowing mask (e.g., Wijaya & van Groesen, 2016). Ocean current and wind information are also available through a suitable analysis of radar images.

The main drawback of the use of radar images for the surface elevation measurement is the incomplete understanding of the interaction processes between the radar microwaves and the ocean surface (Plant & Farquharson, 2012), making radar images not easy to relate the ocean surface properties. Hence, methods for the inversion of radar images to retrieve the surface elevation thus rely either on empirical formulations for the MTF (e.g., Nieto Borge *et al.*, 2004), or on simplified models of the radar backscattered intensity (e.g., Dankert & Rosenthal, 2004; Lyzenga & Walker, 2015).

I.2.2.2 Lidar Camera

Similar to radars, large spatio-temporal surface elevation datasets can be obtained by the way of lidar (LIght Detection and Ranging) cameras, which operate in the visible light (e.g., green; 532 nm wavelength) or close to it (e.g., near-infrared; 1064 nm). Instead of using modulation properties of Bragg waves to estimate gravity waves characteristics, lidar cameras provide direct geo-referenced measurements of free surface elevations, computed based on laser beam travel times. One advantage of this technique, as compared to X-band radars, is the higher spatial resolution resulting from the smaller divergence of the light beams compared to the microwave beams (Sviridov, 1993), providing a more accurate phase resolved (instantaneous) measurement of the ocean surface. Depending on their wavelength, laser beams are scattered by the air/water interface, by the water column if it has a sufficient turbidity (due to the presence of plankton or other suspension particles), as well as by local surface phenomena such as capillary waves and foam.

Two lidar technologies are available: rotating and flash lidars. Rotating lidars are based on a single rotating beam, pointing in every desired direction at a very high rate thanks to the beams going through a smart association of rotating mirrors. A set of measurements can be considered as instantaneous since the time difference between each measurement points is short compared to the characteristic time of waves propagation processes. For flash lidars, a defined number of beams are distributed on a Cartesian grid, and all pulses are sent at the same time, providing an instantaneous picture of the surface (time differences between instants of pulse reception are negligible). The resulting map from a flash lidar can be compared to a surface photo in which a beam represents a pixel, except data are 3D geo-referenced points. Figure I.4 shows examples of typical instantaneous ocean surface samples using both rotating and flash lidars. Results with and without shadowing effects are presented to highlight the induced irregularity on sampled data.

Even if the lidar technology is widely used for bathymetry measurement from quite a long time, in the form of airborne lidars, systems capable of measurement surface elevation are at their early stage of development. Although no lidar-based commercial product for surface profiling has been proposed so far, several prototypes has been tested, both in laboratory (Allis *et al.*, 2011; Belmont *et al.*, 2007) and in real ocean conditions. Since installation is easier on the coastline, surface elevation in swash/surf zones was investigated on the basis of point measurements or one-directional profiles (i.e., lidar scans along only one direction) (Belmont *et al.*, 2007; Blenkinsopp *et al.*, 2010; Harry *et al.*, 2018; Irish *et al.*, 2006; Martins *et al.*, 2017). As part of the ANR-Astrid PREDEMO-Nav project, a lidar camera was designed and tested at Ifremer, Brest, for surface elevation measurement at one spatial point with a varying angle of incidence (figure I.5). Two-

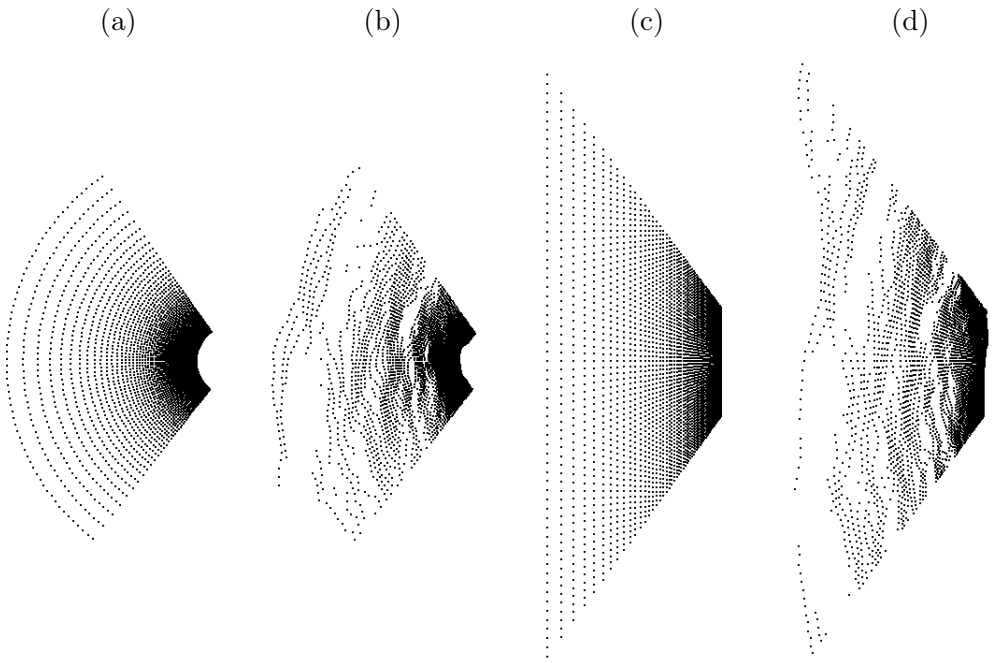

Figure I.4: Typical instantaneous surface sampling from rotating (a,b) or flash (c,d) lidar technologies. For each one, surface sampling without (a,c) and with (b,d) shadowing effects are shown. The exact same sensor's geometrical parameters (viewing and aperture angles, height) and number of points (64×64) are used in all cases.

Figure I.5: Testing of surface elevation measurement using a lidar camera with grazing incidence at Ifremer, Brest, in November 2018. This image corresponds to the calibration phase of the instruments; during the measurements the laser target point was located on the right side of the tower, below a ultrasound wave probe that was used to provide data to compare with. Credit: ECN.

dimensional free-surface maps were obtained by Harry *et al.* (2010), but with a limited rotational scan that introduced a time shift between measurements made at the beginning and at the end of the scan, preventing from having 'instantaneous' spatial maps. Recently, open-ocean tests

were conducted by Kabel *et al.* (2019) using a modern lidar camera from an offshore platform in the North Sea, providing instantaneous two-directional maps of surface elevation.

I.2.2.3 Stereo Video

Stereo-video imagery is a passive technique that, based on a correlation analysis between two images of the same region with different points of view, can lead to detailed estimate time-varying surface elevation fields (see figure I.6) at a relatively high frequency (up to 15–20 Hz). It is particularly efficient to qualify nonlinear spectral properties of the observed wave field (e.g.,

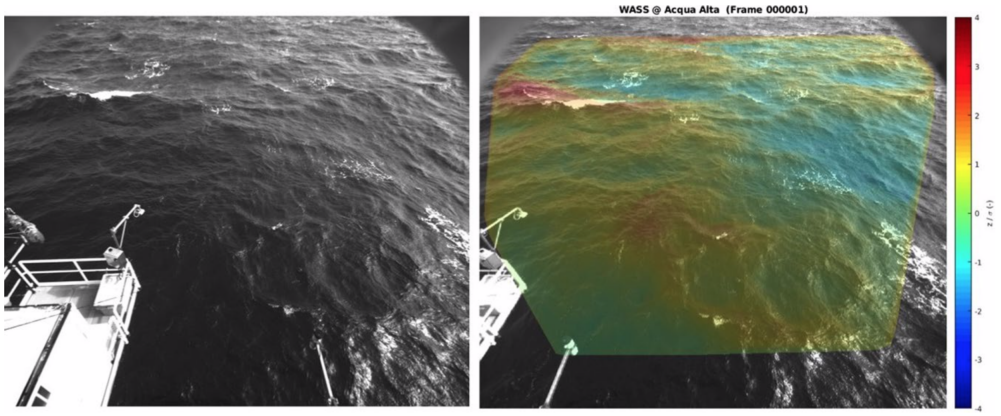

Figure I.6: Example of images from a stereo-video system with the calculated surface map on top of the right image (Guimarães *et al.*, 2020).

the energy distribution as a function of both frequency and wavenumber), wave breaking (as white-capping appears explicitly), air-sea interaction or wave-current interaction. The dependence on light conditions (i.e., not operational during the night) of stereo-video systems is a disadvantage against radars and lidars for remote free surface elevation measurements. However, its use becomes more common (e.g., Filipot *et al.*, 2019; Guimarães *et al.*, 2020), and a significant development of stereo-video systems is expected in the near future.

We have seen in this section the working principles of the main technological solutions currently employed to measure deterministic ocean wave quantities, and their respective assets and limitations have been highlighted. In the following section, after presenting the general governing equations for the physical description of ocean waves, we will explore the methods that have already been developed to post-process wave measurements for deterministic prediction.

I.3 Wave Models and Associated Inversion Methods

Several wave models have been considered for wave prediction applications over the past. The most used, which is also the simplest, is certainly the Linear Wave Theory. However, as we will see below, this approach has several shortcomings regarding the prediction accuracy when it comes to large horizon of prediction and large wave steepness. More complex approaches were thus proposed and investigated in terms of prediction performances and calculation time. Since the prediction algorithms must be able to run in real time, all considered wave models rely on the potential flow theory, for which we briefly recall the governing equations. Only the methods that were used in algorithms to deterministically predict a wave field are reviewed here (we refer for instance to Blondel (2009) and Le Touzé (2003) for a detailed review of existing approaches for wave fields modeling).

I.3.1 General Governing Equations

In ocean waves analysis, fluid and flow hypotheses reasonably apply to the problem description, drastically simplifying the corresponding mathematical modeling, and allowing the use of the potential flow theory. These hypotheses are:

- Since the flow induced by ocean waves is associated with a high Reynolds number, viscous effects are neglected, i.e., the induced dissipation on the spatio-temporal scale of interest is negligible;
- Typical Mach number is very low in ocean waves flow and the temperature variations are negligible within the spatio-temporal scale we are interested in here, the fluid can thus be assumed isovolume (i.e., incompressible and indilatable);
- We do not consider cases of wave breaking and wave-structure interaction, which implies that, under inviscid fluid assumption, there is no source of vorticity (i.e., irrotational flow).

We now consider an infinitely large fluid domain, of finite constant depth h, representing the ocean. We choose a Cartesian coordinate system $(\boldsymbol{r} = (x, y), z)$ with x, y-axes on the mean water surface (i.e., the fluid level at rest) and z-axis positive upward so that the seabed is located at $z = -h$. Governing equations for the velocity potential $\phi(\boldsymbol{r}, z, t)$ in the fluid domain are the Laplace equation (continuity equation) $\Delta\phi = 0$ and the Bernouilli-Lagrange equation (momentum conservation equation) that reads

$$\frac{\partial\phi}{\partial t} + \frac{1}{2}\left|\widetilde{\boldsymbol{\nabla}}\phi\right|^2 + \frac{p}{\rho} + gz = c(t), \qquad \text{in the fluid domain,} \tag{I.5}$$

where $\widetilde{\boldsymbol{\nabla}}$ is the three-dimensional gradient, t is time, $c(t)$ is a time-dependent constant (also known as the Bernoulli constant), ρ is the water density, p is the pressure and g is the gravity acceleration. We further assume that the air above the ocean surface has no viscosity and a constant pressure (i.e., $c(t)$ is set to the normalized atmospheric pressure, there is no wind forcing), and surface tension effects are neglected so that the only force on the water is the gravitational force (i.e., no capillary waves). The problem reduces to the study of a single phase with a single-valued free surface defined by the equation $z = \eta(\boldsymbol{r}, t)$, on which it is possible to write equation (I.5) to obtain the following dynamic free-surface boundary condition

$$\frac{\partial\phi}{\partial t} + \frac{1}{2}\left|\widetilde{\boldsymbol{\nabla}}\phi\right|^2 + g\eta = 0, \qquad \text{on } z = \eta(\boldsymbol{r}, t).$$

In order to close the problem, we join to the latter equation a kinematic boundary condition on the free surface, stating that the fluid velocity $\boldsymbol{V} = \widetilde{\boldsymbol{\nabla}}\phi$ and the surface velocity $\boldsymbol{U}$ have equal normal components at the free surface (according to the impermeability property of the surface), following

$$\boldsymbol{V}\cdot\boldsymbol{n} = \boldsymbol{U}\cdot\boldsymbol{n},$$

where $\boldsymbol{n} = (-\partial\eta/\partial x, -\partial\eta/\partial y, 1)$ is a free-surface normal vector. The free-surface velocity $\boldsymbol{U}$ is calculated by the material derivative of the free surface location as

$$\boldsymbol{U} = \left|\begin{matrix}\boldsymbol{u}\\ U_z\end{matrix} = \frac{\mathrm{D}}{\mathrm{D}t}\right|\begin{matrix}\boldsymbol{r}\\ \eta(\boldsymbol{r},t)\end{matrix} = \left(\frac{\partial}{\partial t} + \boldsymbol{U}\cdot\widetilde{\boldsymbol{\nabla}}\right)\left|\begin{matrix}\boldsymbol{r}\\ \eta(\boldsymbol{r},t)\end{matrix} = \right|\begin{matrix}\boldsymbol{u}\\ \dfrac{\partial\eta}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\eta\end{matrix},$$

in which $\boldsymbol{\nabla}$ is the horizontal gradient, and from which we deduce

$$\boldsymbol{U}\cdot\boldsymbol{n} = \frac{\partial\eta}{\partial t}, \qquad \text{and} \qquad \boldsymbol{V}\cdot\boldsymbol{n} = \frac{\partial\phi}{\partial z} - \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\eta.$$

The kinematic boundary condition thus reads

$$\frac{\partial \eta}{\partial t} = \frac{\partial \phi}{\partial z} - \boldsymbol{\nabla}\phi \cdot \boldsymbol{\nabla}\eta, \qquad \text{on } z = \eta\left(\boldsymbol{r}, t\right).$$

On the seabed, we suppose a slipping condition $\partial\phi/\partial z = 0$. Finally, the system of equations for the velocity potential yields

$$\begin{cases} \Delta\phi = 0, & \text{in the fluid domain,} \\ \dfrac{\partial \phi}{\partial t} = -g\eta - \dfrac{1}{2}\left|\widetilde{\boldsymbol{\nabla}}\phi\right|^2, & \text{on } z = \eta\left(\boldsymbol{r}, t\right), \\ \dfrac{\partial \eta}{\partial t} = \dfrac{\partial \phi}{\partial z} - \boldsymbol{\nabla}\phi \cdot \boldsymbol{\nabla}\eta, & \text{on } z = \eta\left(\boldsymbol{r}, t\right), \\ \dfrac{\partial \phi}{\partial z} = 0, & \text{on } z = -h. \end{cases} \tag{I.6}$$

Additionally, the lateral boundaries of the ocean domain will be assigned conditions related to the considered wave models.

I.3.2 Linear Wave Theory

Linear Wave Theory (LWT) is derived from the linearization of the system of equations (I.6). The linearization is classically done with respect to the normalized variable $\epsilon = kA$ characterizing the wave steepness (Stokes, 1847), in which k is the wavenumber and A the wave amplitude. Using a Taylor expansion of the elevation and the potential about the still water level $z = 0$, we write

$$\begin{cases} \eta = \eta_1 + \eta_2 + \eta_3 + ..., \\ \phi = \phi_1 + \phi_2 + \phi_3 + ..., \end{cases} \tag{I.7}$$

in which $()^{(m)}$ are mth-order terms. Note that the zeroth-order terms are equal to zero since they correspond to the surface at rest. Then, the free-surface boundary conditions are expanded and sorted to keep only the first-order terms in $\mathcal{O}\left(\epsilon\right)$, to obtain the following linearized system

$$\begin{cases} \Delta\phi_1 = 0, & \text{in the fluid domain,} \\ \dfrac{\partial \phi_1}{\partial t} = -g\eta_1, & \text{on } z = 0, \\ \dfrac{\partial \eta_1}{\partial t} = \dfrac{\partial \phi_1}{\partial z}, & \text{on } z = 0, \\ \dfrac{\partial \phi_1}{\partial z} = 0, & \text{on } z = -h. \end{cases}$$

Assuming the waves periodic in time and space, the following analytical Airy solution is found for the free surface elevation and velocity potential

$$\begin{aligned} \eta_1 &= A\cos\left(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t\right), \\ \phi_1 &= \frac{Ag}{\omega}\sin\left(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t\right)\frac{\cosh\left[k\left(z+h\right)\right]}{\cosh\left(kh\right)}, \end{aligned}$$

in which $\boldsymbol{k} = k\hat{\boldsymbol{k}} = \left(k\cos\theta, k\sin\theta\right)$ where θ is the wave direction of propagation with respect to the x-axis, and the wavenumber k and angular frequency ω are related through the linear dispersion relation

$$\omega^2 = kg\tanh\left(kh\right). \tag{I.8}$$

This solution is then generalized to a superposition of individual waves to get the following polychromatic solution

$$\eta(\boldsymbol{r}, t) = \sum_{n=1}^{N} A_n \cos(\boldsymbol{k}_n \cdot \boldsymbol{r} - \omega_n t - \varphi_n), \tag{I.9}$$

$$\phi(\boldsymbol{r}, z, t) = \sum_{n=1}^{N} \frac{A_n g}{\omega_n} \sin(\boldsymbol{k}_n \cdot \boldsymbol{r} - \omega_n t - \varphi_n) \frac{\cosh[k_n(z+h)]}{\cosh(k_n h)}.$$

According to ocean observations, φ_n are wave phases that can be, as a first approximation, assumed uniformly distributed in $[0, 2\pi]$. The free surface elevation η is thus approximated by a linear summation of N sinusoidal waves evolving independently, with their own amplitude A_n, direction θ_n and speed, without interacting with other wave components. These waves are called *free waves*.

Considering the simplicity of this solution to deterministically describe a phase-resolved wave field, LWT was employed in many studies that aim at predicting surface waves, either as the solution chosen in a specific prediction algorithm or as a reference to compare with more complex wave models. Based on LWT, Morris *et al.* (1998, 1992) were the first to propose a deterministic prediction method. They initiate the new discipline of Deterministic Sea Wave Prediction and introduced the fundamental concept of prediction zone, i.e., a spatio-temporal region within which the wave field is theoretically predictable from a specific set of wave measurements (see section III.3 for details). In these papers, they state that the phase velocity is the criterion to define the boundaries of the prediction zone. Later, Wu (2004) showed that it would rather be related to the wave group velocities, from which follows a long-term discussion (reported in here section III.3) about the correct formulation.

I.3.2.1 DFT-based Inversion Methods

If the measurement method allows generating observations that exhibit appropriate specifications for the application of Fourier analysis, i.e., uniform sample rates and periodicity, a linear wave field can be reconstructed with Discrete Fourier Transforms (DFT), such as Fast Fourier Transform (FFT) algorithms. Fourier coefficients then correspond to an estimate of free wave amplitudes A_n and their associated phases. Only one inverse FFT (IFFT) is then necessary to retrieve the surface elevation at any location $\boldsymbol{r}_f$ and at any time instant t_f in the future (i.e., wave components are just shifted by an amount $\boldsymbol{k}_n \cdot (\boldsymbol{r}_f - \boldsymbol{r}) - \omega_n (t_f - t)$). This property makes LWT a very good candidate for real-time applications. However, requirements for the application of FFT on wave measurements is never met in practice, due to the non-periodicity of real wave fields, and usually to the non-uniformity of measurement grids (especially for remote sensing technologies that are affected by shadowing effects). 1D-FFTs can be applied to one-dimensional data, such as time series of surface elevation from a fixed wave probe, and used to reconstruct a long-crested wave field, given that (i) the signal has a constant time sampling rate, and (ii) the signal is sufficiently long and finely sampled to contain enough spectral information (and to limit the impact of windowing effects).

Because of the necessity to have access to measurements with a regular spatial sampling, the process to retrieve a directional (short-crested) wave field through Fourier analysis is way more challenging. Even if a constant time sampling is easy to set on measurement systems, regularity of the spatial sampling is not compatible with the required measurement footprint size and density. Indeed, the observations have to both cover a sufficiently large area to be able to measure long waves and be dense enough for short waves to be detectable (i.e., Nyquist's limitation). From section I.2, it follows that no current systems are able to provide such data with regularly sampled observations. Remote sensing technologies do give access to large and dense spatial samples of surface elevations, but wave shadowing effects prevent from a priori

defining measurement points locations, since it is related to the local surface elevations which determination is the purpose of the measurement itself.

Adaptation to non-uniform input data

Still based on LWT, several studies have been done on how to deal with the non-uniform distribution of wave measurements. Belmont & Morris (1994) described a "non-uniform to uniform mapping" adaptive strategy based on a measure of non-uniformity and an approximate interpolation function to use in the context of remotely sampled surface elevations. Using simulated instantaneous spatial scans, they found that the higher the waves amplitudes, the stronger the non-uniformity and the higher the interpolation error. Belmont (1995) investigated the process to recover the spectrum of non-uniformly sampled functions and proposed an extension of the Nyquist's theorem. Later, Belmont & Horwood (1999) quantified the discrepancy arising from the non-uniform frequency distribution of the wave spectrum that is used to generate the input data. In order to better compare prediction results using datasets generated with different spectrum, they suggested a procedure to produce consistent spectra leading to a similar amount of discrepancy. To avoid errors during the phase shifting process due to discrete spectral description of the wave field, Belmont *et al.* (2003) developed analytical shifting functions that do not require a discrete formulation of the frequency spectrum.

Adaptation to non-periodic input data

LWT and the other spectral models that are usually employed for wave modeling rely on periodic functions to represent the ocean surface elevation. However, raw wave measurements are always non-periodic. Considering wave measurements as periodic signals thus induces leakage errors in the wave field estimate, producing spectral distortion within the bandwidth of the original sea spectrum and creating new components outside of it. Naaijen & Huijsmans (2008) applied a "tanh-shaped window function with steep slopes" to their input signal before performing 1D-FFTs for long-crested waves predictions in order to reduce leakage errors. Experimental data were used to predict wave elevation, ship motion (in heave and pitch) and their respective envelopes. Encouraging results were found while leaving room for significant improvements in prediction accuracy. In fact, classical signal processing windowing methods that work to improve the spectral estimate in non-periodic conditions are not applicable for signal reconstruction since information near the signal boundaries cannot be discarded. To assess this issue, Edgar *et al.* (2000) proposed a preprocessing method consisting of producing a periodic subdomain of the signal by matching both surface elevation and slope between two parts of the signal before analysis. This end-matching method, however, leads to a reduction of the length of the signal, causing information to be lost, thus reducing the time for which predictions are available. Instead, Belmont *et al.* (2006) adopted an approach based on convolution functions to obtain the required phase shift (in time or space) for prediction at the desired location and time instant. Later, Abusedra & Belmont (2011) introduced a "data extension method" based on an iterative scheme that extends the input signal with adequately synthesized data to produce a dataset which has the same bandwidth of the original sea spectrum.

Applications to X-band radar images

Despite the difficulties described above, DFTs are widely used for linear wave field reconstructions and predictions. The use of X-band radar images for deterministic wave field reconstruction through consecutive 2D-FFTs (Atanassov *et al.*, 1985) or 3D-FFTs (i.e., 2D space + time) and a filtering method based on the linear dispersion relation (Nieto Borge *et al.*, 2004; Young *et al.*,

1985) opened the field for many developments of directional wave fields forecasting systems using LWT and DFTs applied to radar images (for the most part in the context of ship motion prediction) (e.g., Belmont *et al.*, 2014; Blondel-Couprie & Naaijen, 2012; Dannenberg *et al.*, 2010; Hilmer & Thornhill, 2014; Hilmer & Thornhill, 2015; Naaijen & Blondel-Couprie, 2012; Serafino *et al.*, 2011). Qi *et al.* (2016) included a nonlinear wave model (third-order HOS, see section I.3.4.1 for details about the HOS method) in the radar images inversion procedure in order to better catch the observed surface dynamics and improve the surface elevation reconstruction.

On the basis of the theoretical relation between the angle of incidence of radar pulses and the intensity of the backscattered signal (tilt modulation), Dankert & Rosenthal (2004) developed a method to retrieve the surface elevation without requiring an empirical calibration of the MTF. They later used their method to derive significant wave heights values from the obtained free surface elevation fields, and compare the results with Waverider buoy and laser measurements (Dankert *et al.*, 2005).

Following a comparable approach, Naaijen & Wijaya (2014) developed an inversion method that uses 1D-FFTs applied on tilt radar signal to retrieve the (unscaled) surface elevation along each of the radar beam direction (i.e., in a polar coordinate system). An interpolation of η on a Cartesian grid is then performed to apply 2D-FFTs on the interpolated map. The main direction of propagation of the waves is calculated from the directional wave spectrum, which can be obtained, for example, from a 3D-FFT of radar images or an auxiliary directional wave buoy. A scaling factor is then applied to the wave coefficients that is calculated from another measurement source (e.g., amplitude of a ship motion). The prediction is updated regularly according to the procedure detailed by Wijaya *et al.* (2015), which consists in averaging the current prediction with propagated radar images from previous time steps, except in the radar blind area (i.e., center of the radar image). The inaccuracies pertaining to the shadowing effects are partly compensated by the averaging process, and after a certain simulation time, the prediction in the blind area of the radar is 'initialized' with propagated solutions. Naaijen *et al.* (2018) adopted this approach to predict ship motion and validated it against field tests.

I.3.2.2 Alternative approaches

Although these DFT-based methods lead to accurate sea state statistics, their limited performance in retrieving deterministic wave characteristics led to the development of alternative approaches, still using the LWT. For example, Belmont *et al.* (2014) proposed the "multiple fixed point method" that was shown to yield better performances than the classical DFT approach in the case of sea-state with narrow directional spreading. The principle of this method is to calculate the incoming waves using wave measurements in the form of time series at a reasonable number (e.g., three) of fixed spatial locations, which has been optimized (in a way to reduce errors related to the wave field inversion method). Regarding the fact that it is based on fixed point measurements, this method is not applicable to predict the incoming waves on moving structures. Still on the basis of fixed-point measurements, Wada *et al.* (2016) and Takagi *et al.* (2017) investigated the "Multi-Point Method", which considers the surface velocity potential as a distribution of sources using the free-surface Green function (Newman, 1985), and uses a convolution method to shift the reconstructed quantities to the point of interest for prediction. Numerical simulations are used as wave measurements, and their results seem to give accurate predictions of long-crested wave fields. In the short-crested case, predictions are of increasing accuracy with the number of measurement points (varying from 1 to 20). However, no comparison is done with respect to other inversion approaches.

An additional method developed by Mérigaud & Ringwood (2019) considers the wave description as a stationary, homogeneous Gaussian random field, and provides free surface predictions as a function of the covariance matrix between the measurements and predictions, which is completely derived from the surface elevation spectrum and the linear dispersion relation. The use

of a nonlinear dispersion relation would allow this method to take into account some nonlinear effects. In a least mean square sense, the optimal prediction that corresponds to a minimal error covariance matrix is found. Redundant statistical information are identified and discarded to form a prediction matrix that only has to be multiplied to the observations to provide the prediction, ensuring a very short computation time. This method is relatively simple, but it relies on an a priori accurate estimate of the wave spectrum, and is not able to take into account non-stationary processes. Also, since the calculation of the covariance matrix (which represents the most computational effort in the prediction chain) is related to specific measurements locations, this method is not directly compatible with predictions around moving structures that necessitate a constant update of the measurements locations. An experimental validation of this approach was proposed by Mérigaud *et al.* (2018) based on ADCP measurements of wave elevation. Due to the very short distances between the locations of their measurements, the prediction performance of their setup was very limited. They however found results that are quite consistent with the theoretical framework. Note that a comparable approach has been employed by Simanesew *et al.* (2017) and compared to other methods for the prediction of wave field with various characteristic using wave tank data. In their study, the covariance approach leads to limited performances compared to the others. However, two major upgrades brought by Mérigaud & Ringwood (2019) are the use of the wave spectrum for the calculation of the covariance matrix and the elimination of redundant statistical information, which may have an impact on the prediction performances.

Even for LWT, there are methods adapted to X-band radar images that do not use DFT for wave inversions. Least square minimization was used by Naaijen *et al.* (2009), Connell *et al.* (2015) and Kusters *et al.* (2016) to fit surface elevation maps. Paalvast *et al.* (2014) applied a 3D panel method using a frequency domain Green function and directly fed by either surface elevation or tilt angle radar observations. They found that, in the case of truncated observations due to shadowing effects, the method that uses tilt information performs better than the one uses elevation information. This method is limited by the radar-images resolutions for the size of the panels, which can lead to convergence issues, especially for short waves.

The use of LWT is limited to sea states exhibiting small characteristic steepness (typically $H_s/\lambda_p \sim 0.01$, where $\lambda_p = 2\pi/k_p$ is the peak wavelength and k_p the peak wavenumber, i.e., of maximal energy), assuming that appropriate modeling of *bound waves* (i.e., harmonic waves that do not obey the dispersion relation) can be neglected, and that the space and time scales of the observations and prediction horizon do not allow time-dependent nonlinear wave-wave interactions (e.g. nonlinear phase shift) to play a significant influence in the description of the wave dynamics. As soon as these restrictions breach, partial or complete modeling of nonlinear wave properties becomes important for an accurate prediction.

I.3.3 Weakly Nonlinear Models

In order to increase the range of physical phenomena described by the wave model, one method consists in increasing the order of expansion of the perturbation approach described in the previous section. To further simplify the resolution, the perturbation approach can be supplemented with developments of the surface elevation and velocity potential in harmonic series around a carrier wave. The evolution of the envelope of the surface elevation is then described by the NonLinear Schrödinger (NLS) equation or one of its modifications. A different approach is to use the Lagrangian framework to describe the fluid dynamics, which is intrinsically appropriate for describing steep waves. A similar perturbation approach to the one described in the previous section can be done to derive solutions at different orders, leading to wave properties that differ from (and are shown to exhibit more physical content than) the classical Eulerian approach.

As we will see, the drawback of these perturbation approaches is that the mathematical complexity drastically increases with the order of expansion, making it not usable to model

highly nonlinear effects, thus limiting its range of application. Only solutions expanded up to the second-, third- or fourth-order (depending on the approach) are used in practice, which makes this approach weakly nonlinear. Nevertheless, for practical ocean engineering applications, accuracy of weakly nonlinear wave models is often sufficient.

I.3.3.1 Second-Order Solution

Similar to the first-order expansion of the governing wave equations (I.6), second-order solution can be obtained by keeping terms up to $\mathcal{O}\left(\epsilon^2\right)$ in the elevation and potential expansions (I.7). The first approach to derive nonlinear interactions between wave components is the *mode coupling* method as it was proposed by Phillips (1960), Longuet-Higgins (1962) and Hasselmann (1962) under deep water assumption. This method is based on a Fourier decomposition of the free surface elevation and velocity potential, and considers nonlinear interactions as coupling terms between modal components of the wave field that is initially inferred from a superposition of linear components. This approach was then extended to finite depth by Dalzell (1999). We recall here Dalzell's expression for the second-order free surface elevation, generalized to a discrete directional spectrum, which follows

$$\eta\left(\boldsymbol{r},t\right)=\eta_1+\eta_2+\mathcal{O}\left(\epsilon^3\right),\qquad\text{with}\qquad\eta_2=\eta^{\text{stokes}}+\eta^{+}+\eta^{-}. \tag{I.10}$$

Each term in $\mathcal{O}\left(\epsilon^2\right)$ can be seen as the result of wave-wave interactions: η^{stokes} corresponds to the interaction of the wave with itself, whereas η^{+} and η^{-} correspond interactions between waves of different frequencies. They read

$$\eta^{\text{stokes}}=\sum_{n=1}^{N}\frac{A_n^2k_n}{4\tanh\left(k_nh\right)}\left[2+\frac{3}{\sinh\left(k_nh\right)}\right]\cos\left(2\psi_n\right),$$
$$\eta^{\pm}=\sum_{m<n}^{N}A_mA_nB_{mn}^{\pm}\cos\left(\psi_m\pm\psi_n\right),$$

where $\psi_n=\boldsymbol{k}_n\cdot\boldsymbol{r}-\omega_nt-\varphi_n$, and the kernel $B_{mn}^{\pm}$ yields

$$\begin{aligned}B_{mn}^{\pm}=&\frac{\left(\omega_m^2+\omega_n^2\right)}{2g}\mp\frac{\omega_m\omega_n}{2g}\left[1\mp\frac{\cos\left(\theta_m-\theta_n\right)}{\tanh\left(k_mh\right)\tanh\left(k_nh\right)}\right]\\&\times\left[\frac{\left(\omega_m\pm\omega_n\right)^2+g\left|\boldsymbol{k}_m\pm\boldsymbol{k}_n\right|\tanh\left(\left|\boldsymbol{k}_m\pm\boldsymbol{k}_n\right|h\right)}{D_{mn}^{\pm}}\right]\\&+\frac{\left(\omega_m\pm\omega_n\right)}{2gD_{mn}^{\pm}}\left[\frac{\omega_m^3}{\sinh^2\left(k_mh\right)}\pm\frac{\omega_n^3}{\sinh^2\left(k_nh\right)}\right],\end{aligned}$$

in which

$$D_{mn}^{\pm}=\left(\omega_m\pm\omega_n\right)^2-g\left|\boldsymbol{k}_m\pm\boldsymbol{k}_n\right|\tanh\left(\left|\boldsymbol{k}_m\pm\boldsymbol{k}_n\right|h\right).$$

The dispersion relation (I.8) remains unchanged at second-order. We notice that the mathematical complexity pertaining to the second-order terms is higher compared to that of the linear solution (I.9), and efficient wave field reconstruction/prediction methods based on (I)FFTs are no longer adequate to estimate second-order quantities.

Blondel *et al.* (2010) and Blondel-Couprie *et al.* (2013) used equation (I.10) to reconstruct and predict long-crested nonlinear wave fields, using both numerical and experimental data. They employed the iterative procedure proposed by Duncan & Drake (1995) to initialize the wave model, in which at each step, linear waves (free waves) and second-order effects (bound waves) are decoupled using equation (I.10), to eventually converge (according to a prescribed tolerance) leading to the knowledge of the free waves characteristics (namely their amplitude and

phase). Once the free waves are known, the surface evolution is predicted at any time instant by simply updating the time quantity in the model formulation. They found that the overall improvements of the prediction made using such second-order wave model are very limited compared to the predictions using LWT. These results lead them to include a third-order correction to the dispersion relation, hence modifying waves velocities, as we will see in the next section.

Zhang *et al.* (1993) showed that the modal coupling approach for the calculation of nonlinear interactions between two wave components having frequencies located far away in the wave spectrum (i.e., large short-to-long wavelength ratio) converges slowly and eventually diverges. Instead, they adopted a *phase-modulation method* to calculate second-order quantities. This method consists in dividing the wave spectrum in several frequency bands and describing the nonlinearities in terms of a phase-modulation of the short-wavelength components by the long-wavelength ones. The phase-modulation method being less efficient than the modal coupling method in the case of components having close frequencies, Zhang *et al.* (1993) proposed a hybrid model based on modal coupling for the calculation of interactions between components from the same or joint frequency bands, or on phase modulation for components from disjoint frequency bands.

Such a Hybrid Wave Model (HWM) is used by Zhang *et al.* (1996) and Spell *et al.* (1996) to reconstruct and predict long-crested wave fields, and was then extended by the same authors to a Directional Hybrid Wave Model (DHWM) in the case of short-crested waves (Zhang *et al.*, 1999*a*,*b*). In the case of long-crested waves, irregular wave trains with a JONSWAP spectrum were generated in narrow wave tanks, and times series of free surface elevation at different fixed locations, and of orbital velocities at different water depths were recorded. A decoupling iteration procedure based on FFTs applied to surface elevation measurements was used to retrieve the amplitudes and phases of the free-wave components. Predictions of surface elevation and orbital velocities were then performed using HWM to compute the effects of the nonlinear interactions. Surface elevation predictions using the LWT were also performed, as well as velocity predictions using the Wheeler stretching method (Wheeler *et al.*, 1970) and the linear extrapolation (or delta-stretching) method (Rodenbusch & Forristall, 1986). They found that HWM is more accurate than LWT for predicting surface elevation than Wheeler stretching and linear extrapolation for predicting orbital velocities, particularly for high wave steepness.

For short-crested waves, the first step is the use of an Extended Maximum Likelihood Method (EMLM) according to Isobe *et al.* (1984) to obtain the wave amplitude distribution as a function of the wave direction. Since EMLM does not retain the phases of the wave components, the second step is to determine the phases by fitting the reconstructed wave field with the input data, which consist of sets of laboratory data (surface elevation and pressure measurements) as well as sets of field data (surface elevation, pressure and velocity measurements). Once a first estimate of the free-wave amplitudes and phases is done, DHWM is used to compute the wave-wave interactions, which are then subtracted from the measurements that are used for the next EMLM, phase fitting, etc. Iterations are performed until convergence is achieved according to a preset error tolerance between two estimates of the nonlinear effects. It is shown that the predictions are more accurate when laboratory data are used compared to that from field data, due to the uncontrolled environment leading to measurement errors. Also, better performances were found when surface elevation times series are used for DHWM inversion than pressure measurements, probably due to the wave induced pressure decay in the water column, making components of high frequency hard to measure and globally increasing the noise to signal ratio. Even if increasing the number directions per frequency (up to 7) allows to increase the accuracy of the wave field reconstruction, due to the increase of degrees of freedom in the phase fitting step, it has no significant effect on the prediction. The number of wave measurements (varying from 3 to 5) only slightly affects the accuracy of the predictions.

Note that Prislin *et al.* (1997) were the first to predict second-order nonlinear short-crested waves. However, since they used reference wave measurements recorded by pressure sensors

deployed in the ocean, they were not able to properly measure the short wave dynamics. Hence, they considered that the interactions between wave components of short-to-long wavelength ratio were not significant, leading them to use the conventional mode coupling approach.

Even if the concept of accessible prediction zone was already introduced by Morris *et al.* (1998, 1992), the studies described above (Prislin *et al.*, 1997; Spell *et al.*, 1996; Zhang *et al.*, 1996, 1999*a*,*b*) do not use it in their analysis.

I.3.3.2 Enhanced Second-Order Solution

As noticed in the previous section, the second-order terms modify the wave shape through the modeling of bound waves resulting from self-interactions and interactions between two waves of different frequencies. However, the dispersion relation, that governs waves velocity, is similar to that of LWT. In order to further improve the modeling of nonlinear features and include some wave-phase modifications, a third-order expansion has to be performed. Developing the expansion up to the third order (e.g., Zakharov, 1968) produces both non-resonant terms (highly complex to derive in practice since their expressions involve many cross interaction terms) and resonant terms that model the energy exchange between wave components and necessitate a temporal resolution. Some of these resonant terms result in a modification of the dispersion relation that gives rise to a nonlinear wave velocity in amplitude. These resonant effects have been implemented in a second-order solution (Blondel, 2009; Blondel *et al.*, 2010; Blondel-Couprie *et al.*, 2013; Perignon, 2011), leading to the so-called enhanced second-order solution.

The Eulerian approach leads to a phase velocity correction at the third-order of expansion, in the form of a wave-wave interactions effects. Third-order terms are calculated by considering interactions between four waves in two distinct cases: if all wave frequencies are equal (i.e., $k_1 = k_2 = k_3 = k_4$) or if they are equal two-by-two (i.e., $k_1 = k_2$ and $k_3 = k_4$). The case of interaction of waves having equal frequency was derived by Stokes (1847, p. 211), and corresponds, for a wave of amplitude A_n, to the phase velocity correction

$$c^{\text{stokes}} = \frac{1}{2} A_n^2 \omega_n k_n. \tag{I.11}$$

The case of different wave frequencies was derived by Longuet-Higgins & Phillips (1962, equation (2.11)). The same quantity is written for a discrete distribution of directional waves by Perignon (2011, equation (3.12) which includes the Stokes term[1]) and yields the phase velocity correction

$$c_n^{\text{LH}} = \frac{K}{2 A_n \omega_n^2}, \tag{I.12}$$

with

$$K = \sum_{\substack{m=1 \\ m \neq n}}^{N} A_m^2 A_n \omega_m \omega_n \left\{ \begin{array}{l} (\omega_m - \omega_n) \left| \boldsymbol{k}_m - \boldsymbol{k}_n \right| \cos^3\left(\frac{1}{2}\theta\right) \left[1 + \dfrac{4\omega_m \omega_n \sin^2\left(\frac{1}{2}\alpha\right)}{(\omega_m - \omega_n)^2 - g\left| \boldsymbol{k}_m - \boldsymbol{k}_n \right|} \right] \\ + (\omega_m + \omega_n) \left| \boldsymbol{k}_m + \boldsymbol{k}_n \right| \sin^2\left(\frac{1}{2}\theta\right) \left[1 - \dfrac{4\omega_m \omega_n \sin^2\left(\frac{1}{2}\beta\right)}{(\omega_m + \omega_n)^2 - g\left| \boldsymbol{k}_m + \boldsymbol{k}_n \right|} \right] \\ + \omega_m \left[k_m - k_n + 2k_n \cos^2\left(\frac{1}{2}\theta\right) \sin^2\left(\frac{1}{2}\theta\right) \right] + \omega_n \left(k_m + k_n\right) \cos\left(\theta\right) \end{array} \right\},$$

in which θ, α and β are angular relations between wavelength vector $\boldsymbol{k}_m$ and vectors $\boldsymbol{k}_n$, $\boldsymbol{k}_m - \boldsymbol{k}_n$ and $\boldsymbol{k}_m + \boldsymbol{k}_n$, respectively (Longuet-Higgins & Phillips, 1962; Perignon, 2011). The nonlinear,

[1] Note that, in the expression of Stokes term by Perignon (2011, equation (3.14)) in K, the term should read $k_{ij}\omega_{ij}^3 A_{ij}^3$ instead of $k_{ij}^2\omega_{ij}^3 A_{ij}^2$ to yield the appropriate dimension and be consistent with the Stokes result.

third-order, phase velocity is thus

$$\begin{aligned} c_n^{\mathrm{nl}} &= c_n + c^{\mathrm{stokes}} + c_n^{\mathrm{LH}} \\ &= \frac{\omega_n}{k_n} + \frac{1}{2}A_n^2\omega_n k_n + \frac{K}{2A_n\omega_n^2} \\ &= \frac{\omega_n}{k_n}\left[1 + \frac{1}{2}k_n^2\left(A_n^2 + \frac{K}{A_n k_n \omega_n^3}\right)\right], \end{aligned}$$

or equivalently, the nonlinear angular velocity is

$$\omega_n^{\mathrm{nl}} = \omega_n\left[1 + \frac{1}{2}k_n^2\left(A_n^2 + \frac{K}{A_n k_n \omega_n^3}\right)\right]. \tag{I.13}$$

This model was employed by Perignon (2011) to study its performance compared to a higher-order model. It was found that this model brings a significant improvement compared to the LWT and can be considered as a free surface representation of high accuracy during the first characteristic periods of the wave field. Blondel (2009), Blondel *et al.* (2010) and Blondel-Couprie *et al.* (2013) used the enhanced second-order model in the purpose of long-crested waves prediction. Instead of formulating third-order corrections in terms of a nonlinear angular velocity as it is done here in equation (I.13), they fixed the angular velocity and formulated a third-order correction of the wavenumber. This way, the assimilation scheme, that relies on one single time trace, is equivalent to the one employed in the classical second-order case (see previous section) since wavenumbers do not influence the wave field when it is calculated at one location. Results were first compared to synthetic data (Blondel *et al.*, 2010), then to experimental data (Blondel-Couprie *et al.*, 2013), and they used a third-order model (HOS at order three, see section I.3.4.1 for details on the HOS method) as reference model to compare the prediction accuracy with. In the synthetic case (input data generated by a fourth-order HOS), they found that, in contrast to the classical second-order model, the enhanced formulation improves considerably the prediction accuracy for moderate wave steepness by decreasing the prediction error by almost a factor two compared to the LWT and the second-order model. The full third-order solution still allows decreasing further the prediction error by 40% compared to the enhanced solution (60% compared to LWT and the second-order solution) for high wave steepness.

These observations are validated in the case of experimental data. The enhanced second-order solution provides the most accurate predictions for moderate wave steepness and for a limited prediction distance (i.e., distance from the recording location of input data). For small steepness, the LWT, the second-order and the enhanced second-order prediction models give equivalent results, while the third-order model seems to be less accurate. When the prediction distance becomes large, nonlinear effects become significant and the full third-order model provides the lowest prediction errors. Also, as a consequence of nonlinear effects, increasing the wave steepness reduces the prediction distance for the third-order model to exhibit the most accurate prediction. The fact that the third-order model does not return the most accurate results for short prediction distances and short steepness seems to be explained by an inadequate estimate of the wave phases.

In conclusion, this model has shown reasonable performances in terms of accuracy of the ocean surface representation with a limited computational cost. It significantly improves the prediction accuracy compared to the second-order model that has shown to be almost equivalent to the LWT when used in a prediction algorithm. This means that nonlinear phase modeling is a crucial feature that needs to be properly taken into account in deterministic prediction.

I.3.3.3 NonLinear Schrödinger Equation

By assuming that the wave spectrum is narrow-banded, centered around the frequency of a characteristic wave of small steepness, nonlinear wave properties can be modeled using the NLS

equation or one of its modifications, which describes the evolution of the surface elevation envelope of the characteristic wave by means of modulations. This approach is widely used for analysis of nonlinear wave-wave interaction phenomena such as the generation of rogue waves (e.g., Chabchoub *et al.*, 2011). Even if the hypothesis of narrow-banded wave spectrum is not trivial for ocean waves, the performance of this approach has been investigated in the context of deterministic prediction of typical ocean wave fields (Klein *et al.*, 2020; Simanesew *et al.*, 2017; Trulsen, 2007; Trulsen & Stansberg, 2001).

Envelope equations

Assuming the slow variation of a wave train, the velocity potential and surface elevation can be developed in harmonic series

$$\begin{aligned}\eta &= \bar{\eta} + \mathrm{Re}\left(Be^{\mathrm{i}(k_c x-\omega_c t)} + B_2 e^{\mathrm{i}2(k_c x-\omega_c t)} + B_3 e^{\mathrm{i}3(k_c x-\omega_c t)} + ...\right), \\ \phi &= \bar{\phi} + \mathrm{Re}\left(Ae^{\mathrm{i}(k_c x-\omega_c t)} + A_2 e^{\mathrm{i}2(k_c x-\omega_c t)} + A_3 e^{\mathrm{i}3(k_c x-\omega_c t)} + ...\right),\end{aligned} \tag{I.14}$$

where k_c and ω_c are the characteristic wavenumber and angular velocity of the wave field, respectively, and are related by the dispersion relation. $\bar{\eta}$ is the mean surface elevation and $\bar{\phi}$ is the mean flow. We further assume a narrow-banded frequency bandwidth centered around the frequency of a characteristic wave of steepness $\epsilon = k_c a_c$ (with a_c its amplitude), i.e., the relative bandwidth follows $\mu = |\Delta \boldsymbol{k}| / k_c \ll 1$ in which $\Delta \boldsymbol{k}$ is a characteristic modulation wave vector.

A perturbation approach with respect to ϵ and μ allows to describe the evolution of the surface elevation envelop B by Schrödinger-type equations. The NLS equation is third-order in wave steepness, but we describe hereafter the higher-order Modified NonLinear Schrödinger (MNLS) equation (also known as the Dysthe equation) that neglects terms of relative order higher than ϵ^4. Note that there exists a Broad MNLS equation formulation (Trulsen & Dysthe, 1996) that relaxes the narrow banded constraint. Considering that the water depth is large, and following Dysthe (1979) and Trulsen & Dysthe (1997), the space-domain formulation of the MNLS equation describes the spatial evolution of B according to

$$\frac{\partial B}{\partial x} + \mathcal{L}B + \mathrm{i}\,|B|^2 B - 8\,|B|^2 \frac{\partial B}{\partial t} - 2B^2 \frac{\partial B^*}{\partial t} - 4\mathrm{i}\frac{\partial \bar{\phi}}{\partial t} B = 0, \tag{I.15}$$

where the superscript * denotes the complex conjugate and each term is normalized in a similar way to that of Simanesew *et al.* (2017). The operator $\mathcal{L}$, which corresponds to the exact linear dispersive part, is given by Trulsen *et al.* (2000) and yields

$$\mathcal{L} = -\mathrm{i}\left\{\left[\left(1 + \mathrm{i}\frac{\partial}{\partial t}\right)^4 + \frac{\partial^2}{\partial y^2}\right]^{1/2} - 1\right\}.$$

For the MNLS equation, only terms up to the fourth order are retained, hence every harmonic in equation (I.14) is used to calculate the surface elevation, with $\bar{\eta}$, B_2 and B_3 varying slowly with time and space as

$$\bar{\eta} = -\frac{\partial \bar{\phi}}{\partial t}, \quad B_2 = \frac{1}{2}B^2 + \mathrm{i}B\frac{\partial B}{\partial t}, \quad \text{and} \quad B_3 = \frac{3}{8}B^3,$$

and the governing equations for the potential $\bar{\phi}$ of the mean flow read

$$\begin{cases} \dfrac{\partial \bar{\phi}}{\partial z} = -\dfrac{\partial\,|B|^2}{\partial t}, & \text{on } z = 0, \\ \Delta\bar{\phi} = 0, & \text{in the fluid domain,} \\ \dfrac{\partial \bar{\phi}}{\partial z} = 0, & \text{on } z = -h. \end{cases}$$

Other Schrödinger-type equations can be formulated from equation (I.15): the cubic Nonlinear Schrödinger (NLS) equation is obtained by discarding the last three terms, and the Linear Schrödinger (LS) equation only retains the first two terms. The numerical procedure to solve the above MNLS equation is not detailed here but examples can be found in Lo & Mei (1985, 1987).

Ocean wave fields prediction performance

Trulsen & Stansberg (2001) and Trulsen (2007) used the LWT, the NLS and MNLS equations to deterministically predict the evolution of bichromatic waves by means of probes measurements. One wave probe was used to initialize the models and predictions were compared to downstream time series at fixed locations. It was found that the NLS equation gives almost equivalent results than the LWT in terms of root mean square error, whereas the MNLS results are significantly better.

Later, Simanesew *et al.* (2017) compared, still based on surface elevation experimental data, the prediction performances of six prediction methods in the condition of bichromatic waves, long-crested irregular waves, and short-crested waves with different directional spreading. Two of the prediction methods rely on a simple least squares formulation of the problem in order to calibrate a set of weights leading to an optimal prediction. One is just an empirical linear calibration, while the second allows the predictions and the observations to be interdependent stochastic variables, and the least squares minimization is performed using covariance matrices. In practice, both least squares methods have been shown to give equivalent results. The other four methods are based on the LWT and on the LS, NLS and MNLS equations. The initialization of LWT is simply done through a 2D-FFT of the spatio-temporal input signal. The initialization of LS, NLS and MNLS requires first an estimate of the characteristic angular frequency ω_c (or equivalently the characteristic wavenumber k_c), which is typically calculated as a mean angular frequency with respect to the wave spectrum S following

$$\omega_c = \iint \omega S^p(\omega,\theta)\,\mathrm{d}\theta\,\mathrm{d}\omega \Big/ \iint S^p(\omega,\theta)\,\mathrm{d}\theta\,\mathrm{d}\omega,$$

where p is a peak-enhancement parameter (e.g., $p=1$ in Simanesew *et al.* (2017) and 5 in Klein *et al.* (2020)). The second step consists in extracting the surface elevation envelope B from the measurements η_m, which they do by low-pass filtering the Fourier transform of the quantity $2\,\eta_\mathrm{m} e^{\mathrm{i}\omega_c \mathrm{t}}$ $\left(=(B+B^*)\,e^{-2\mathrm{i}\omega_c \mathrm{t}}\right)$. For bichromatic waves, similar results to those of Trulsen & Stansberg (2001) are found concerning LWT and MNLS (actually the same dataset was used in both studies and similar initialization procedures were employed). Despite their different initial estimates, LWT and LS were found to give similar prediction results for all wave conditions. Except in the case of short-crested waves with small directional spreading at short distances, NLS does not improve the predictions from the linear wave theory and even leads to worse predictions in case of long-crested irregular waves. The least square approach is relatively efficient only in the case of bichromatic waves. Overall, the MNLS equation gives better predictions, but the directional spreading significantly affects the prediction accuracy and only short prediction distances can be reached for short-crested waves. Authors indicate that this is mainly due to the lack of directional observations.

The Schrödinger-equation approach was also investigated by Klein *et al.* (2020) by means of experimental measurements of long-crested waves with different wave steepness and spectrum peakedness factor. Since the objective is to evaluate the prediction performances of the methods using radar images, input data were chosen to be spatial series of surface elevation. Hence, they adopted a "semi-experimental" procedure: the input data are numerically generated on the basis of digital twin of the wave tank experiments by means of the time-domain Finite Element

Method potential flow solver waveTUB (Steinhagen, 2001). This technique allows using finely sampled spatial series of surface elevation that could be very difficult to produce experimentally. Accuracy of the predictions is then estimated by comparison with experimental time series at three downstream locations and during time spans corresponding to the theoretically accessible prediction zone. The prediction error indicator is a normalized "surface similarity parameter" (Perlin & Bustamante, 2016).

Three NLS-type models are investigated: NLS and MNLS with a full linear dispersion relation (that includes terms of higher-order than the limited truncation of (M)NLS) and NLS with a consistently truncated linear dispersion relation. They showed that, due to the inaccurate narrow-banded spectrum hypothesis, the modeling of the exact linear dispersion is crucial and improves significantly the predictions compared to the truncated version in which the dispersion effects are only approached by a Taylor series expansion about the carrier wave. Still, due to the narrow-banded spectrum hypothesis, results show that increasing the peakedness factor (which concentrates the wave energy towards a peak of the spectrum) makes the prediction error decrease. The benefits of MNLS arise for the highest steepness. The authors then compare the MNLS results to the LWT and a fourth-order HOS solution (see section I.3.4.1 for details about the HOS method). The overall better performance of HOS and the fact that HOS and the input waveTUB records exhibit similar prediction error in every case indicates (i) that the use of a complex wave model extends significantly the range of efficiency of a wave prediction tool and (ii) that relative accuracy of the input data with respect to the actual wave quantities is a crucial step in the prediction chain. They also note that results from LWT and NLS with full dispersion are almost equivalent, with a slightly better prediction accuracy for the latter model. This small difference could be explained by the underlying hypothesis of narrow-banded wave spectrum pertaining to NLS, even if the full dispersion is modeled.

The MNLS equation appears to be able to compete with the classical perturbation approach leading to the enhanced second-order solution described in section I.3.3.2. However, since the wave spectrum distribution plays an important role in the model validity, its performance would need to be further assessed in short-crested waves (i.e., using a setup that provides more directional observations than the one used by Simanesew *et al.* (2017)) to conclude on the actual benefits of its use in the context of deterministic prediction.

I.3.3.4 Lagrangian Approach

Alternatively to the Eulerian description of the fluid dynamics which seeks the value of the fluid velocity and pressure fields at fixed spatial locations, the Lagrangian approach uses the motion of fluid particles and pressure at their location to describe the fluid behavior. Due to its intrinsic properties, the Lagrangian representation of the ocean surface of a given order can model nonlinear wave properties that are not included in Eulerian developments of the same order (as explicitly explained by Nouguier *et al.*, 2015). For instance, the linear approximation of the Lagrangian and Eulerian descriptions of an irregular wave field have similar mathematic complexity, but the Lagrangian solution exhibits slope properties that correspond to a third-order Eulerian solution (e.g., Lindgren, 2009; Nouguier *et al.*, 2009). This characteristic comes from the fact that the Lagrangian formulation satisfies the free surface boundary conditions on the actual free surface, leading to the inclusion of interactions that are nonlinear in the Eulerian formulation.

The first solution to the Lagrangian equations was found more than two centuries ago by Gerstner (1809), who derived a solution that exactly satisfies the nonlinear equations for a periodic wave. Because his solution relies on the deep-water assumption and includes a second-order vorticity term, it has not been investigated in depth: the Eulerian solution proposed by Stokes (1847) was largely preferred by researchers. It is only one and a half centuries later that Pierson (1961) proposed irrotational solutions based on a Lagrangian perturbation expansion

with respect to the wave steepness up to the second order for long-crested waves.

Still, due to the theoretical complexity of the mathematical developments in a Lagrangian framework, the majority of ocean wave investigations rely on the Eulerian formalism. A detailed and complete description of the Lagrangian equations of motion is provided by Gjosund (2000). Nevertheless, due to its suitableness to represent very steep waves and to evaluate wave field statistical quantities at a moderate cost in terms of analytical complexity, a renewed interest in the Lagrangian formalism has emerged (Nouguier *et al.*, 2015) and its physical implications has been investigated in the context of water waves. For instance, Gjosund (2003) used a first-order solution to describe the surface elevation and kinematics of irregular long-crested wave fields of intermediate water depth. Interesting features were highlighted such as the fact that it favorably compares to the Wheeler method (Wheeler *et al.*, 1970), providing a more accurate description of the kinematics, within a theoretically consistent framework. Another notable study is provided by Clamond (2007) who proposed a detailed overview of the physical and mathematical benefits pertaining to the Lagrangian formulation in the case of a standing wave.

Nouguier *et al.* (2009) developed the "Choppy Wave Model" (CWM), an extension to the Gerstner's theory for the description of irregular short-crested waves. They showed that CWM correctly approaches second-order (in an Eulerian sense, Longuet-Higgins, 1963) elevation and slope properties for narrow-banded spectra. Later, Nouguier *et al.* (2015) derived an analytical framework to describe short-crested Lagrangian waves at second order denoted CWM2. They demonstrated that CWM2 is fully consistent with the classical second-order Eulerian expansion by Longuet-Higgins & Phillips (1962) and even captures nonlinear effects that can be related to third-order properties. Moreover, their analytical framework allows to improve the Lagrangian second-order formulation by Pierson (1961) for long-crested waves by incorporating an essential second-order term to ensure a zero-mean level.

CWM was used in for deterministic prediction by Grilli *et al.* (2011) and Nouguier *et al.* (2014) on the basis of synthetic lidar measurements. They first compared the performance of LWT and CWM to accurately reconstruct and predict a long-crested, uniformly sampled CWM surface elevation, using a reconstruction method relying on the minimization of a quadratic cost function. It was found that the higher the wave steepness, the better the performance of CWM over LWT, suggesting that CWM was suitable for nonlinear wave field prediction. For long- and short-crested waves, non-uniform lidar-like elevation measurements are then used as input for the CWM inversion. It was shown that the use of multiple surface snapshots are crucial to reduce the prediction error, countering the error induced by wave shadowing effects. The inclusion of time information also leads to get access to the wave directionality, which directly affects and improves the estimation of the wave amplitudes.

At second order, Lagrangian expansion includes a horizontal particle shift that affects the dispersion relation, whereas, as explained in section I.3.3.2, nonlinear corrections of the dispersion relation appear at third order with the Eulerian approach. At equivalent computational cost, the Lagrangian description thus allows a more accurate modeling of wave velocity, and as seen in section I.3.3.2, taking into account nonlinear corrections of the wave velocity is crucial to issue accurate predictions. This suggests that this Lagrangian second-order model could be well suited for fast and efficient wave predictions. Since this approach is the one chosen in this work to perform nonlinear predictions, a detailed presentation of the Lagrangian equations and their solutions is given in section II.1.

I.3.4 High-Order Spectral Method for Fully Nonlinear Predictions

Further nonlinear properties of the surface wave dynamics can be modeled by means of numerical methods. The High-Order Spectral (HOS) method solves the fully nonlinear wave equations (I.6)

using a pseudo-spectral approach. Because of its computational efficiency (numerical schemes using FFTs) and high accuracy, this method is widely used for deterministic investigations of highly nonlinear wave fields. The inversion of HOS, however, needs a significantly higher computational efforts than the inversion of LWT or weakly nonlinear models, leading to calculation times that are not compatible with the real-time constraint of prediction systems. Even if no operating prediction systems based on HOS has been proposed so far, the increasing computational efficiency of modern machines allows considering making such real-time computations in the near future.

I.3.4.1 Core of the HOS Method

The HOS method was developed independently by Dommermuth & Yue (1987) and West *et al.* (1987). The core of the method consists in expressing the free surface boundary conditions according to the Hamiltonian formalism formulated by Zakharov (1968), i.e., using the surface velocity potential ϕ^s given by

$$\phi^s(\boldsymbol{r},t) = \phi(\boldsymbol{r}, z=\eta(\boldsymbol{r},t),t).$$

Using the relation $\partial\phi^s/\partial x^* = \partial\phi/\partial x^* + \partial\eta/\partial x^* \times \partial\phi/\partial z$ for any variable $x^* \in \{\boldsymbol{r}, z, t\}$, the modified surface boundary conditions follow

$$\begin{cases} \dfrac{\partial\phi^s}{\partial t} = -g\eta - \dfrac{1}{2}|\boldsymbol{\nabla}\phi^s|^2 + \dfrac{1}{2}\left(1+|\boldsymbol{\nabla}\eta|^2\right)\left(\dfrac{\partial\phi}{\partial z}\right)^2, \\ \dfrac{\partial\eta}{\partial t} = \left(1+|\boldsymbol{\nabla}\eta|^2\right)\dfrac{\partial\phi}{\partial z} - \boldsymbol{\nabla}\phi^s\cdot\boldsymbol{\nabla}\eta, \end{cases} \quad \text{on } z=\eta. \tag{I.16}$$

The problem is now expressed with only free surface quantities η and ϕ^s, which reduces to two dimensions a problem that was initially expressed in the fluid volume. Assuming that values of these quantities are known at a given instant of time, the advance in time can be easily performed using conditions (I.16) together with a numerical scheme, e.g., Runge-Kutta method. The remaining unknown (and last volume quantity) is the vertical speed $W = \partial\phi/\partial z|_{z=\eta}$, which is evaluated through a high-order iterative process.

The problem is solved through the spectral approach, which is based on the a priori assumption that our quantities of interest η, ϕ^s and ϕ can be decomposed, at each time, in sets of eigenfunctions $\psi_n(\boldsymbol{r})$ that suit the domain geometry and boundary conditions. Quantities are thus expressed in a new basis which yields for the velocity potential

$$\phi(\boldsymbol{r},z,t) = \sum_{n=1}^{N} A_n(t)\,\psi_n(\boldsymbol{r})\,\frac{\cosh\left[k_n(z+h)\right]}{\cosh\left(k_n h\right)}.$$

The surface elevation and surface velocity potential follow similar expressions without the z-dependent term. This approach is numerically very efficient as it converges rapidly and allows, in our case, the use of FFTs to determine the modal amplitudes.

To calculate W, the potential ϕ is written as a perturbation series in η to a given arbitrary order M (called the HOS order) as

$$\phi(\boldsymbol{r},z,t) = \sum_{m=1}^{M} \phi^{(m)}(\boldsymbol{r},z,t).$$

At $z=\eta$, the potential is further expanded in a Taylor series around $z=0$, so that we have

$$\phi^s(\boldsymbol{r},t) = \phi(\boldsymbol{r},\eta,t) = \sum_{m=1}^{M}\sum_{k=0}^{m-1} \frac{\eta^k(\boldsymbol{r},t)}{k!}\frac{\partial^k\phi^{(m-k)}}{\partial z^k}(\boldsymbol{r},0,t). \tag{I.17}$$

Since η and ϕ^s are known, equation (I.17) corresponds to a Dirichlet boundary condition for ϕ on the free surface. It results a sequence of boundary conditions that yields, for each order $\phi^{(m)}$

$$\phi^{(1)}(\boldsymbol{r},0,t)=\phi^s(\boldsymbol{r},t)\,,\quad\text{and}\quad\phi^{(m)}(\boldsymbol{r},0,t)=-\sum_{k=1}^{m-1}\frac{\eta^k(\boldsymbol{r},t)}{k!}\frac{\partial^k\phi^{(m-k)}}{\partial z^k}(\boldsymbol{r},0,t)\quad\text{for } 1<m\leq M.$$

The resulting system is solved iteratively for the modal amplitudes $A_n^{(m)}(t)$ at increasing orders m. Finally, and similar to the potential, the vertical velocity is expressed as a perturbation series

$$W_M(\boldsymbol{r},z,t)=\sum_{m=1}^{M}W^{(m)}(\boldsymbol{r},z,t)\,,$$

which is derived on the free surface from a Taylor expansion according to

$$\begin{aligned}W^{(m)}(\boldsymbol{r},\eta,t)&=\frac{\partial}{\partial z}\left(\phi^{(m)}(\boldsymbol{r},0,t)+\eta(\boldsymbol{r},t)\frac{\partial\phi^{(m-1)}}{\partial z}(\boldsymbol{r},0,t)+\frac{\eta^2(\boldsymbol{r},t)}{2}\frac{\partial^2\phi^{(m-2)}}{\partial z^2}(\boldsymbol{r},0,t)+...\right)\\&=\sum_{k=0}^{m-1}\frac{\eta^k(\boldsymbol{r},t)}{k!}\frac{\partial^{k+1}\phi^{(m-k)}}{\partial z^{k+1}}(\boldsymbol{r},0,t)\,.\end{aligned}$$

At a converged order of nonlinearity M (a typical value of $M=5$ is often sufficient), this method gives access to the fully nonlinear wave dynamics. It has been used for applications related to highly nonlinear wave phenomena such as rogue waves modeling (Ducrozet *et al.*, 2007) or wave propagation over variable bathymetry (Gouin *et al.*, 2016), and validated many times against experiments (e.g., Bonnefoy *et al.*, 2010; Ducrozet *et al.*, 2012).

I.3.4.2 Model Inversion and Numerical Performances

All comparative studies have demonstrated that HOS gives the best prediction accuracy in all waves conditions (Blondel, 2009; Blondel *et al.*, 2010; Blondel-Couprie *et al.*, 2013; Klein *et al.*, 2020; Qi *et al.*, 2018*a*; Wu, 2004). However, the use of HOS for the deterministic prediction involves considerably more calculations than the lower-order approaches, which makes it difficult to use in practice. Using the same semi-experimental procedure as described by Klein *et al.* (2020), Clauss *et al.* (2014) showed the capabilities of HOS to deterministically predict highly nonlinear phenomenon, namely the propagation of solitons and Peregrine breathers (Peregrine, 1983). Later, Clauss *et al.* (2015) used the same approach to demonstrate the technical feasibility of ship motion predictions under unidirectional irregular sea states using HOS, at least when the initial HOS solution that is propagated is known. The complexity of the HOS model for prediction, however, lies in the model inversion step, which, due to the need for the simultaneous estimation of the HOS surface elevation and velocity potential, as well as due to the HOS numerical complexity, requires more complicated methods than the ones used for lower-order models. It is within the framework of *data assimilation* that the inversion of HOS models is addressed. Given observations (measurements) of a specific process and a mathematical model that describes its dynamics, the point of data assimilation is the determination of the optimal model parameters that better describe the process through the minimization a global error that accurately represents the misfit between the theory and the observed data. Detailed presentation of classical data assimilation procedures can be found in Blondel (2009).

Before going into inversion methods based on assimilation, let us mention that a novel approach is being developed by Huchet *et al.* (2018) at ECN that aims at predicting wave fields with HOS on the basis of horizontal velocity measurements, directly used as boundary conditions of the wave model. This approach has the advantage of avoiding the time-consuming step of data

assimilation, that is also an important source of errors in prediction algorithms. Developments are made around the open source code HOS-NWT[2] that simulates a numerical wave tank with the HOS method. A wavemaker, modeled on one side of the wave tank and driven by horizontal velocity conditions in the water column, deterministically generates waves that propagate using the HOS method in the wave tank before being absorbed by a numerical beach on the other side of the tank. A proper initialization of the wavemaker would theoretically lead to very accurate predictions with the minimal computational effort pertaining to the propagation of the solution. The main difficulty with this approach relies on the determination of the initial conditions in a real environment, for which Huchet *et al.* (2018) propose a method based on ADCP measurements. Specifically, the horizontal velocity profile in the water column is approached from profiles of vertical velocity having close horizontal coordinates provided by at least two ADCPs or two rows of multiple ADCPs along a sufficiently large distance in the case of long- or short-crested waves, respectively. The deployment of ADCPs being not compatible with a constant update of the measurement locations, this method is not yet applicable in the context of predictions around a moving structure.

Two main approaches are used in data assimilation: the sequential and the variational approaches. In the former, the model is adjusted observation by observation, i.e., an estimate of the model state (or parameterized model) is made from a first observation, then updated at each new observation available. When all observations have been used, the last model state estimate is used as initial condition for the model forecast. In this method, each correction step does not take into account future observations. In contrast, the variational approach tries to assimilate all observations at once through the minimization of a cost function. The obtained solution, which serves as initial condition for the forecast, has the closest trajectory to the observations. This method is thus more capable of capturing the dynamics of the observed process and is also more adapted to nonlinear problems. As we will see below, most of the HOS inversion procedures are based on variational assimilation, and it is also the approach employed in this work. A detailed presentation of the variational cost function that is minimized is done in section III.2.

Variational approach with gradient-free optimization

The first procedure to invert HOS is proposed by Wu (2004). His method consists in minimizing the classical quadratic variational cost function for the initial η and ϕ^s through an optimization process that determines their complex modal amplitudes. His cost function formulation includes a weighting function that emphasizes the large surface displacements so that the features near the wave peaks are better captured. Since HOS relies on a numerical integration scheme, there is no analytical formulation of the derivatives of the cost function to the optimization variables. Gradient-based optimization methods thus necessitate numerical calculation of the derivatives and lead to a high number of computations. Instead, he used the gradient-free optimization method proposed by Rosenbrock (1960). At each step of the optimization, the HOS solution is propagated over the spatio-temporal region covered by the observations to obtain the reconstructed wave signal. The initial surface elevation and velocity potential for the HOS propagation are found by applying a FFT on the reconstructed wave elevation signal (from the previous iteration) in order to find the free-waves components used in the calculation of the second-order quantities using the mode coupling formulation (equation (I.10) for the surface elevation). The surface potential is typically calculated with a Taylor series expansion of the potential around $z = 0$ up to the second order. An intermediate step consists in making the second-order solution periodic over the observation zone to ensure a consistent initialization of the HOS solution (which is periodic). This process is repeated until the cost function converges to a prescribed precision. Wu (2004) checked the uniqueness of the solution as well as the convergence of this approach

[2] https://github.com/LHEEA/HOS-NWT

with increasing HOS order M.

Blondel (2009); Blondel *et al.* (2010) used a similar approach. The method proposed by Duncan & Drake (1995) is used for the free-waves decoupling procedure, which is extended to the enhanced second-order formulation and to finite depth. The code HOS-ocean[3] (Ducrozet *et al.*, 2016) is used as numerical solver for the HOS simulations. They also used a complementary weighting function which gives more weight to the observations near the end of the observation times. Almost similar results as the ones provided by Wu (2004) are obtained. They later validated this approach against experimental data (Blondel-Couprie *et al.*, 2013).

The main downside of the gradient-free optimization technique is the high-number of iterations needed to converge (e.g., $\sim$ 2500 cost function evaluations to reconstruct long-crested waves of steepness $H_s/\lambda_p = 2\%$ in Blondel *et al.* (2010)).

Variational approach with gradient-based optimization

Qi *et al.* (2018*a*) developed a method that is almost similar to the one from Wu (2004), except that they do not use weighting functions for the cost function formulation, and that they use the quasi-Newton optimization procedure based on a gradient evaluation. The numerical details concerning the calculation of the cost function derivatives are reported in Qi (2017). Numerical and experimental validations are performed using surface elevation data and horizontal velocity measurements in the water column.

Aragh & Nwogu (2008) and Aragh *et al.* (2008) chose an optimization method based on the evaluation of the gradient of the cost function using adjoint equations. Before deriving the adjoint equations, the "tangent linear model" equations that describe the evolution of an initial perturbation of the solution, have to be calculated. This step is done by linearizing the HOS evolution equations with respect to a small perturbation of the initial model parameters. Tangent equations return the discrepancy on the HOS solution after forward propagation. Adjoint equations are derived from the tangent equations and describe the backward evolution of a discrepancy δy and allow to find the initial HOS solution. This way, the comparison between an initial HOS solution y_i and the solution of the adjoint equations parameterized by the discrepancy $y_o - y_f$ between an observation y_o and the propagated HOS simulation y_f, allows the determination the gradient of the cost function for a specific set of parameters. With this approach, estimating the gradient only necessitates one forward (HOS) and one backward (adjoint model) model propagations at each iteration, which significantly improves the computational efficiency compared methods that only rely on a forward model propagations.

In their work, Aragh & Nwogu (2008) used a third-order HOS model and derived and used the adjoint equations up to the third order, together with a conjugate gradient optimization procedure, to predict unidirectional wave fields. However, since their observations are surface elevation series, their optimization parameters are the complex amplitudes of the surface elevation. Hence, after an optimization iteration, they only have access to the updated surface elevation amplitudes and do not know the corresponding velocity potential. They thus chose to estimate the velocity potential using the linear formulation based on the free waves components to initialize the HOS model for forward propagation. Aragh *et al.* (2008) applied the same method to directional wave data, but using an adjoint Newton optimization procedure and first-order adjoint equations for the gradient calculation, which is lower than the order of the considered wave model (third-order HOS) and limits the interest of this approach which was to calculate the gradient with a high accuracy. Even if the convergence is achieved very rapidly compared to the gradient-free optimization procedures, another drawback of this method is the increasing complexity of the adjoint equations with the HOS order, which makes their implementation very challenging.

[3] https://github.com/LHEEA/HOS-ocean

In order to get rid of the complex adjoint derivation and implementation, Fujimoto & Waseda (2020) used an ensemble-based method in the optimization process. This method consists in computing the gradient of the cost function using perturbed ensemble simulations. More specifically, they used the a4dVar method which finds the optimal perturbations on the initial wave model parameters that minimize the cost function. A particular set of perturbations is selected at each optimization iteration based on an analysis of the estimate model at previous iteration, leading to a more efficient search for optimum. It can also be implemented in a parallelized way to decrease the computational time. They named this method Surface Waves reconstruction by Ensemble Adjoint-free Data assimilation. Besides the constraints imposed by the observations, they applied a regularization of inverse problem through the modeling of the background error covariance using the wave spectrum. In fact, they considered a diagonal background covariance matrix where each term corresponds to the variance of one wave component amplitude, which should be proportional to the power spectrum. Several values of the regularization parameter were tested. With a third-order HOS model, they successfully reconstructed and predicted waves fields that contain extreme wave events. Here, the surface velocity potential is estimated from the linear wave theory, without a free-wave decoupling method, assuming the nonlinear surface elevation amplitudes to be free wave components.

Jacobian-Free Newton-Krylov method

Remote observations consist of surface elevation fields (e.g., from radar or lidar measurements), thus do not allow for a direct estimate of the velocity potential, which is necessary to initialize HOS simulations. Inversion methods thus often rely on a linear or second order estimate of the potential, reducing both the numerical performance of the inversion technique and the accuracy of the result. Assuming that the temporal sampling of the observations is high enough to calculate the surface elevation derivatives, Köllisch *et al.* (2018) developed a method to directly estimate the surface velocity potential from the observations. Their estimate is done through the minimization of the residual F arising from the kinematic surface boundary condition (I.16) and follows

$$F\left(\phi^s_{\text{rec}}\right) = \frac{\partial \eta_{\text{obs}}}{\partial t} - \left(1 + |\boldsymbol{\nabla}\eta_{\text{obs}}|^2\right) W\left(\eta_{\text{obs}}, \phi^s_{\text{rec}}\right) + \boldsymbol{\nabla}\phi^s_{\text{rec}} \cdot \boldsymbol{\nabla}\eta_{\text{obs}},$$

where η_{obs} corresponds to the surface elevation observations and ϕ^s_{rec} to the reconstructed surface velocity potential. They used the Newton method to iteratively minimize F, but since F is nonlinear, the calculation of its Jacobian is non-trivial. Instead, they employed a Krylov subspace method that allows solving the nonlinear minimization problem only through evaluations of F.

Concerning the computational efficiency, they obtained a calculation time of 0.6 s to invert a JONSWAP sea state of steepness $H_s/\lambda_p \approx 3.2\%$ on a 512×512 points grid with a HOS order $M = 4$, which is satisfactory. The forward propagation of the inverted sea state is then performed approximately five times faster than real time.

Hence, this method seems promising for real-time applications, but it necessitates two major prerequisites. First, the time derivative of every surface elevation observation needs to be available from the measurements. Considering radar images which have a time sampling rate of typically 1–2 s and exhibit shadowing effects that lead to constantly updated measurement points locations, this is unlikely to be the case. The measurement configurations around a structure with a forward speed would also lead to practical limitations for the estimate of $\partial\eta/\partial t$. The second requirement is the Cartesian layout of the observations, which, as seen in section I.2, would necessitate interpolation.

Kalman filter

The sequential approach was investigated by Yoon *et al.* (2016), who developed an efficient Kalman filter (Kalman, 1960) that is analytically calculated. Their method associates the non-linear HOS model at order three for forward propagation of the wave field, and LWT for the calculation of the error covariance matrix that is used to determine the Kalman gain applied on the model parameters at each update step. The key point of their approach is that the error covariance matrix can be analytically formulated and efficiently calculated since it does not need to be integrated in time but only estimated at each update instants.

Their method is shown to be stable, even when the observations are contaminated with a normal random noise. However, for large steepnesses, the reconstruction procedure seems to have difficulties to accurately represent the nonlinear dynamics of the wave field, which is mainly due to the fact that the Kalman filter is based on a linearized wave model. Also, the sequential nature of Kalman filtering may not ensure the physical consistency in case of highly nonlinear events.

I.4 Use of Machine Learning

Due to the recent interest and impressive developments of machine learning approaches to address engineering issues, several studies envisaged the prediction of ocean waves by a data-trained parametric model instead of a physics-based model.

Desouky & Abdelkhalik (2019) investigated the prediction of waves ahead a controlled wave energy converter. Synthetic linear long-crested waves are used first, and successful predictions are obtained within a limited horizon of prediction. Real field data from surface buoys are then used without achieving accurate prediction. This can be explained by the fact that they used only one buoy to predict a real (thus directional) wave field.

Law *et al.* (2020) used an artificial neural network to predict synthetic nonlinear (HOS) long-crested waves and compared the prediction results with the LWT. With input data generated in different configurations (surface elevation time series at a fixed position, two times series of same time interval at fixed positions and one surface series at a time instant), prediction error maps are computed and the error over the theoretically accessible prediction zone are analyzed. It is found that the developed data-driven model gives lower error than LWT in all configurations, especially in case of high wave steepness, suggesting that the neural network approach is capable of capturing some nonlinear features. After the necessary training step, the prediction can be easily issued in real-time since no model inversion is required. However, this new approach needs to be further investigated regarding the short-crested waves prediction capabilities. An interesting improvement of this approach would be the integration of physics into a data-driven model to increase its performances (e.g., a weakly nonlinear wave model supplemented by a data-generated correction of highly nonlinear effects).

Conclusion

The benefits and drawbacks of the existing prediction methodologies have been reviewed. Each methodology depends both on a specific wave model which has its own hydrodynamic properties and on a specific wave measurement methods. We saw that the linear wave theory is the most convenient model for inversion from any set of wave measurements, even if precautions have to be taken regarding non-uniformity and non-periodicity of the measured data.

It appears, however, that the linear wave theory is only sufficient to predict wave fields of small steepness and during limited prediction horizons. Investigations of weakly nonlinear wave models showed that improvements are significant if nonlinear corrections of wave velocities are taken into account, which are third-order quantities. Further increased hydrodynamic performances rely on

numerical methods and require complex (and time-consuming) model inversion procedures, based on data assimilation schemes, that are not easily applicable to real wave data measurements and computable in real-time.

The objectives of this work being the deterministic prediction of wave fields around a moving structure such as a ship with a forward velocity, we must rely on a remote sensing technology. Datasets consisting of remotely measured free surface elevation fields will thus be assimilated for prediction. The direct measurements of surface heights, the high sampling rate (in both space and time), and the recent interests and developments for surface elevation profiling are inherent qualities of lidar cameras that make them a relevant choice for the purpose of this work. Through the generation of synthetic datasets, this technique will thus be considered are the source of wave measurements for the following investigation of prediction algorithms. Regarding wave models, we have seen that the Lagrangian approach has not been deeply investigated concerning its phase-resolved prediction performances and seems to be able to model nonlinear wave velocity corrections at a low computational cost. The development of an efficient Lagrangian model is the object of the next chapter.

Résumé du deuxième chapitre

Le deuxième chapitre présente l'approche Lagrangienne pour la modélisation des ondes de gravité à la surface de l'océan, ainsi que les propriétés intrinsèques à cette approches. La dérivation de la solution au premier ordre de perturbation est exposée, puis nous comparons, jusqu'au second ordre, les descriptions Lagrangienne et Eulérienne de la surface libre pour différentes configurations d'état de mer. La solution Lagrangienne inclut au second ordre des effets non-linéaires qui affectent à la fois la forme de la surface libre et la relation de dispersion (reliant fréquence et longueur d'onde des vagues). Le premier effet crée une asymétrie (c'est-à-dire des crêtes plus prononcées et creux plus plats), alors que le second engendre une modification de la vitesse de propagation des vagues. Nous introduisons ensuite une version améliorée de cette solution au second ordre, appelée Improved Choppy Wave Model (ICWM), qui corrige la relation de dispersion non-linéaire afin de mieux représenter les phases lors de la propagation du champ de vagues. Nous montrons que ICWM permet de retrouver la solution Eulérienne au troisième ordre pour une vague régulière. Dans le cas de champs de vagues irréguliers, ICWM corrige convenablement les composantes énergétiques du champ de vagues, mais manque les effets de modulation des vagues plus courtes. Afin d'utiliser ICWM lors de la reconstruction de champ de vagues à partir des mesures, nous en proposons une formulation Eulérienne pour l'élévation de surface. Nous montrons qu'une simple approximation dans un cadre Eulérien, ne nécessitant pas de développements mathématiques complexe ni de procédure itérative, permet d'obtenir des résultats consistants. Dans le but de valider l'intérêt de la formulation améliorée de la solution Lagrangienne, nous comparons divers résultats d'élévation de surface à une solution HOS complètement non-linéaire (c'est-à-dire convergée en ordre de non-linéarité). Il est montré que la correction non-linéaire de la vitesse de phase des vagues est cruciale pour une représentation correcte du champ de vagues après seulement quelques périodes de propagation. Ce travail a été publié dans Guérin *et al.* (2019). Pour finir, une caractérisation de quelques propriétés du champ de pression issu de ICWM est présentée.

Chapter II

Choppy Wave Modeling

Introduction

The need for real-time prediction is not compatible with complex (i.e., highly nonlinear) wave models, such as High-Order Spectral methods. Hence, we seek a model that is able to properly take into account the nonlinear effects that play important roles in the propagation of the considered wave fields, while being suitable for the real-time constraint of the prediction problem. It has been shown in last chapter that the Lagrangian approach could lead to interesting nonlinear properties at a reasonable computational cost. In this chapter, we present and assess numerically a novel derivation of the Lagrangian second-order solution adapted to wave prediction applications. It has been published in Guérin *et al.* (2019).

We first present the Lagrangian approach and some of its intrinsic properties. We detail the derivation of the first-order solution, show how the Lagrangian approach (up to the second-order) compares to the Eulerian approach in terms of free surface description, and explain which parameter bounds its applicability. Then, we introduce the improved Lagrangian second-order solution, referred to as Improved Choppy Wave Model (ICWM). We explain the method employed to derive the model, and provide a simple Eulerian form of the corresponding free surface elevation that is required for the reconstruction of the wave field. In order to validate the relevance of the "improved" formulation, we then compare various results of free surface elevation to a fully nonlinear (i.e., converged in order of nonlinearity) HOS solution. Last, we study the dynamical properties of ICWM through the derivation of the wave induced pressure field. Note that all the model developments that are done in this chapter are for infinite water depth, but the extension to finite depth is conceptually equivalent.

II.1 Lagrangian Approach

Alternatively to the Eulerian approach, where we seek the value of the fluid velocity and pressure fields at fixed spacial locations, the Lagrangian approach uses the motion and pressure of fluid particles to describe the fluid behavior. Even if a similar perturbation approach to that described in section I.3.2 is used to obtain solutions at different orders, the intrinsic formulation of the Lagrangian equations allows to express the free surface elevation boundary conditions on the actual free surface. This makes Lagrangian wave models more adapted to steep waves, and we will see in the following that the Lagrangian representation of a given order can model nonlinear wave properties that are not included in Eulerian developments of the same order.

II.1.1 General Equations

We consider the spatial coordinates of a particle $\boldsymbol{R} = (x, y, z)$ which depend on their independent reference labels (reference location coordinates) $\boldsymbol{\zeta} = (\alpha, \beta, \delta)$ and time t. Explicitly, that is

$x = x(\alpha, \beta, \delta, t)$, $y = y(\alpha, \beta, \delta, t)$ and $z = z(\alpha, \beta, \delta, t)$. $\boldsymbol{\zeta}$ is chosen to be the locus of particles at rest. A particle motion is driven by Newton's law on dynamics for an inviscid fluid, which is

$$\frac{\partial^2 \boldsymbol{R}}{\partial t^2} + \boldsymbol{g} = -\frac{1}{\rho} \boldsymbol{\nabla}_{\boldsymbol{R}} p, \tag{II.1}$$

where $\boldsymbol{g} = (0, 0, g)$. This equation, through the operator $\nabla_{\boldsymbol{R}} = (\partial_x, \partial_y, \partial_z)$, contains differential coefficients with respect to the particle coordinates $\boldsymbol{R}$, instead of the particle position at rest $\boldsymbol{\zeta}$. To eliminate these terms, we multiply equation (II.1) by the Jacobian matrix of $\boldsymbol{R}$ defined as

$$J = \overline{\overline{\nabla}}^T \boldsymbol{R} = \begin{bmatrix} \partial_\alpha x & \partial_\alpha y & \partial_\alpha z \\ \partial_\beta x & \partial_\beta y & \partial_\beta z \\ \partial_\delta x & \partial_\delta y & \partial_\delta z \end{bmatrix},$$

and we obtain the Lagrangian form of the dynamic equation

$$J \frac{\partial^2 \boldsymbol{R}}{\partial t^2} + \boldsymbol{\nabla} (gz) + \frac{1}{\rho} \boldsymbol{\nabla} p = \boldsymbol{0}, \tag{II.2}$$

where the spatial derivative operator $\boldsymbol{\nabla}$ relates to the independent Lagrangian variables $\boldsymbol{\zeta}$.

The determinant of matrix J expresses the volume ratio between the displaced fluid particle and the particle at rest. Since we are assuming that the fluid is incompressible, we get the continuity equation

$$|J| = \frac{\partial (x, y, z)}{\partial (\alpha, \beta, \delta)} = 1. \tag{II.3}$$

One may express the solution as a perturbation expansion in wave steepness $\epsilon = kA$ for the position and pressure

$$\begin{cases} \boldsymbol{R} = \boldsymbol{R}_0 + \boldsymbol{R}_1 + \boldsymbol{R}_2 + ..., \\ p = p_0 + p_1 + p_2 + ..., \end{cases} \tag{II.4}$$

the zeroth-order component being the particle at rest

$$\begin{aligned} \boldsymbol{R}_0 &= \boldsymbol{\zeta}, \\ p_0 &= p_a - \rho g \delta, \end{aligned}$$

where p_a is the atmospheric pressure. This perturbation approach will lead to the expression of the solution at the first then the second order.

Note that the form of the Lagrangian governing equations are very different from the Eulerian ones and may appear less explicit since the free surface does not appear explicitly and the velocity is described through the derivatives of the particle coordinates. As we will see, the mathematical framework needed to find consistent solutions to these equations is indeed not trivial, which could explain the modest number of studies regarding these equations in the context of surface waves.

II.1.2 First-Order Solution

We describe here the process to find a solution of the linearized Lagrangian equations system. At first order, equation (II.2) writes

$$\frac{\partial^2 \boldsymbol{R}_1}{\partial t^2} + \boldsymbol{\nabla} (gz_1) + \frac{1}{\rho} \boldsymbol{\nabla} p_1 = \boldsymbol{0}.$$

The continuity equation (II.3) leads to

$$\frac{\partial x_1}{\partial \alpha} + \frac{\partial y_1}{\partial \beta} + \frac{\partial z_1}{\partial \delta} = \boldsymbol{\nabla} \cdot \boldsymbol{R}_1 = 0.$$

In order to find a solution that fulfills the irrotational flow condition, we investigate solutions of the form $\boldsymbol{R}_1 = \boldsymbol{\nabla} w$ in a way that the function $w(\boldsymbol{\zeta}, t)$ exists, and satisfies the following dynamic and continuity equations

$$\boldsymbol{\nabla}\left(\frac{\partial^2 w}{\partial t^2} + g\frac{\partial w}{\partial \delta} + \frac{p_1}{\rho}\right) = \boldsymbol{0}, \tag{II.5}$$

$$\Delta w = 0. \tag{II.6}$$

Setting p_1 to zero at the free surface ($p_1|_{\delta=0} = 0$) leads to a solution of the dynamic equation (II.5) of the form

$$w = \cos\left(k_\alpha \alpha + k_\beta \beta - \omega t\right) e^{k\delta}, \tag{II.7}$$

with k_α, k_β and k constant numbers, and $\omega^2 = gk$. Injecting (II.7) in continuity equation (II.6) gives

$$\Delta w = \left(-k_\alpha^2 - k_\beta^2 + k^2\right) w = 0,$$

which leads to $k^2 = k_\alpha^2 + k_\beta^2$. Since the components of $\boldsymbol{R}_1$ have a spatial dimension, a suitable solution is $\boldsymbol{\nabla}\left(Ak^{-1}w\right)$, leading to the following spatial components at first-order expansion

$$\begin{cases} x = \alpha - A\dfrac{k_\alpha}{k}\sin\left(k_\alpha \alpha + k_\beta \beta - \omega t\right) e^{k\delta}, \\ y = \beta - A\dfrac{k_\beta}{k}\sin\left(k_\alpha \alpha + k_\beta \beta - \omega t\right) e^{k\delta}, \\ z = \delta + A\cos\left(k_\alpha \alpha + k_\beta \beta - \omega t\right) e^{k\delta}, \\ p = p_a - \rho g \delta. \end{cases} \tag{II.8}$$

Even if the irrotationality hypothesis is not explicit here, this first-order solution can be proven to be irrotational. Indeed, showing that there exists a scalar function F that satisfies

$$\frac{\partial \boldsymbol{R}}{\partial t} = \boldsymbol{\nabla}_{\boldsymbol{R}} F, \tag{II.9}$$

(which denotes $\boldsymbol{\nabla} \times \frac{\partial \boldsymbol{R}}{\partial t} = \boldsymbol{0}$) implies that F is a velocity potential and the flow is irrotational. Plus, by definition, a perfect differential of such a function follows

$$\boldsymbol{\nabla} F \cdot \mathrm{d}\boldsymbol{\zeta} = \boldsymbol{\nabla}_{\boldsymbol{R}} F \cdot \mathrm{d}\boldsymbol{R},$$

and we know that

$$\left(J\frac{\partial \boldsymbol{R}}{\partial t}\right) \cdot \mathrm{d}\boldsymbol{\zeta} = \frac{\partial \boldsymbol{R}}{\partial t} \cdot \mathrm{d}\boldsymbol{R},$$

which means that finding a function F such that $\boldsymbol{\nabla} F = J\frac{\partial \boldsymbol{R}}{\partial t}$ and $\boldsymbol{\nabla} F \cdot \mathrm{d}\boldsymbol{\zeta}$ is a perfect differential is equivalent to show that equation (II.9) is verified. At the first-order of expansion, we have

$$J\frac{\partial \boldsymbol{R}_1}{\partial t} = \boldsymbol{\nabla} F_1 \iff F_1 = Ak^{-1}\frac{\partial w}{\partial t} = \frac{A\omega}{g}\sin\left(k_\alpha \alpha + k_\beta \beta - \omega t\right) e^{k\delta}.$$

Following Schwarz's condition, it is necessary and sufficient to show that

$$\begin{cases} \dfrac{\partial F_{1\alpha}}{\partial \beta} - \dfrac{\partial F_{1\beta}}{\partial \alpha} = 0, \\ \dfrac{\partial F_{1\beta}}{\partial \delta} - \dfrac{\partial F_{1\delta}}{\partial \beta} = 0, \\ \dfrac{\partial F_{1\delta}}{\partial \alpha} - \dfrac{\partial F_{1\alpha}}{\partial \delta} = 0, \end{cases}$$

where $\boldsymbol{\nabla} F_1 = (F_{1\alpha}, F_{1\beta}, F_{1\delta})$, to prove that $\boldsymbol{\nabla} F_1 \cdot \mathrm{d}\boldsymbol{\zeta}$ is a perfect differential. This is easily verified, hence the first order solution is intrinsically related to the description of an irrotational flow.

According to this first-order development, every fluid particle has a circular trajectory (see figure II.1), with a radius $Ake^{k\delta}$ exponentially decreasing as the distance between the particle and free surface increases. One particular feature of this first order solution is that it leads to a

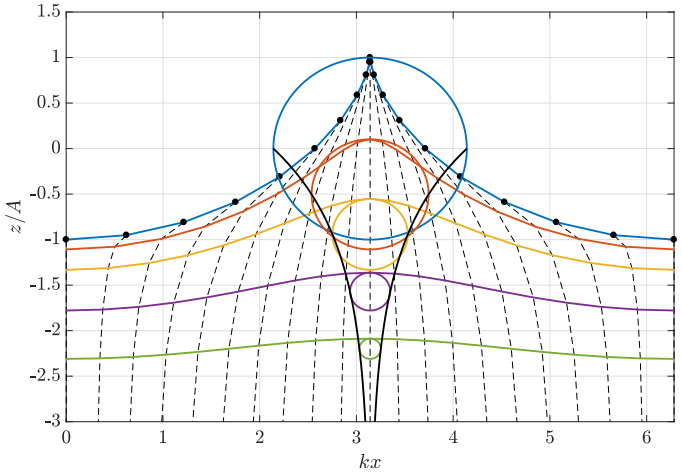

Figure II.1: Snapshot of the fluid particles position according to the first-order solution for a wave of steepness $kA = 1$ (theoretical limit allowed by the model) at a given time $t = \pi/\omega$ and for different depths. The wave is going along the x-axis (i.e., $y = \beta$), and we are looking through a plane perpendicular to the y-axis. Dashed lines represent the form taken by a line of particles which is vertical at rest (and remains vertical when it passes through a crest or a trough).

nonzero mean surface level. Indeed, it introduces a wave shape modification, i.e., the crests are sharper and the troughs flatter, that decreases the mean surface level by an amount $\frac{1}{2}A^2k$. We will see in the next section that a second-order correction corresponding to a wave self-interaction effect removes this discrepancy.

It can be seen from equation (II.4) that the Lagrangian free-surface (i.e., $\delta = 0$) particle position can be expressed in the form

$$\begin{cases} \boldsymbol{r}\left(\boldsymbol{\xi}, t\right) = \boldsymbol{\xi} + \displaystyle\sum_i \boldsymbol{D}_i\left(\boldsymbol{\xi}, t\right), \\ z\left(\boldsymbol{\xi}, t\right) = \displaystyle\sum_i Z_i\left(\boldsymbol{\xi}, t\right), \end{cases} \tag{II.10}$$

where $\boldsymbol{r} = (x, y)$, $\boldsymbol{\xi} = (\alpha, \beta)$ and $\boldsymbol{D}_i$ and Z_i are respectively the horizontal and vertical particle relative displacements with respect to the particle position at rest $\boldsymbol{\xi}$, associated with the i^{th}-order of expansion. According to equation (II.8), the relative displacements at first-order are

$$\begin{cases} \boldsymbol{D}_1\left(\boldsymbol{\xi}, t\right) = -A\hat{\boldsymbol{k}} \sin\left(\boldsymbol{k} \cdot \boldsymbol{\xi} - \omega t\right), \\ Z_1\left(\boldsymbol{\xi}, t\right) = A \cos\left(\boldsymbol{k} \cdot \boldsymbol{\xi} - \omega t\right), \end{cases}$$

where $\boldsymbol{k} = (k_\alpha, k_\beta) = (k\cos\theta, k\sin\theta)$ with θ the direction of propagation of the wave with respect to the x-axis, and $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$. The generalization of this solution to the case of irregular waves is straightforward since it is simply a linear superposition of individual linear regular waves. For a wave field composed of a set of N frequencies, it gives

$$\begin{cases} \boldsymbol{D}_1(\boldsymbol{\xi}, t) = -\sum_{n=1}^{N} A_n \hat{\boldsymbol{k}}_n \sin(\boldsymbol{k}_n \cdot \boldsymbol{\xi} - \omega_n t - \varphi_n) = \mathscr{R}(\eta(\boldsymbol{\xi}, t)), \\ Z_1(\boldsymbol{\xi}, t) = \sum_{n=1}^{N} A_n \cos(\boldsymbol{k}_n \cdot \boldsymbol{\xi} - \omega_n t - \varphi_n) = \eta(\boldsymbol{\xi}, t), \end{cases}$$

where φ_n are wave phases.

Note that the vertical displacement corresponds to the elevation according to the LWT (equation (I.9)), and the horizontal displacement is its Riesz transform $\mathscr{R}$ (generalization of a Hilbert transform for a multi-dimensional signal). The latter feature could be helpful for an efficient calculation of the first-order contributions of the particle position from a known linear free surface elevation. Since its extension for the description of the statistical properties of irregular short-crested waves by Nouguier *et al.* (2009), the Lagrangian solution is commonly referred to as "Choppy Wave Model" (CWM). Accordingly, we will call in the following CWM1 the first-order solution (II.8).

II.1.3 Second-Order Solution

A complete derivation of a solution to equations (II.2) and (II.3) to the second-order can be done using the same method than the one detailed in section II.1.2 for the first-order contributions (Nouguier *et al.*, 2015; Pierson, 1961). In addition, we have shown in Guérin *et al.* (2019) that it is also possible to use Eulerian velocity fields to retrieve those results without having to deal with the Lagrangian equations of motion. In case of long-crested waves, both methods lead to the following second-order contributions

$$\begin{cases} x_2(\alpha, \delta, t) = & \sum_{m<n}^{N} \frac{A_m A_n}{g} \left[\omega_n(\omega_n + \omega_m) e^{(k_n - k_m)\delta} - \left(\frac{\omega_n^3 + \omega_m^3}{\omega_n - \omega_m} \right) e^{(k_n + k_m)\delta} \right] (\sin\Psi_{mn} - \sin\Psi_{mn}|_{t=0}) \\ & + \sum_{n=1}^{N} A_n^2 \omega_n k_n e^{2k_n\delta} t, \\ z_2(\alpha, \delta, t) = & \sum_{m<n}^{N} \frac{A_m A_n}{g} \left[-\omega_n(\omega_n + \omega_m) e^{(k_n - k_m)\delta} + \left(\frac{\omega_n^3 - \omega_m^3}{\omega_n - \omega_m} \right) e^{(k_n + k_m)\delta} \right] \cos\Psi_{mn} \\ & + \sum_{n=1}^{N} \frac{1}{2} A_n^2 k_n e^{2k_n\delta} \\ p_2(\alpha, \delta, t) = & 2\rho \sum_{m<n} A_m A_n \omega_m \omega_n \left[e^{(k_n - k_m)\delta} - e^{(k_n + k_m)\delta} \right] \cos\Psi_{mn} + \rho g \sum_{n=1}^{N} \frac{1}{2} A_n^2 k_n \left(e^{2k_n\delta} - 1 \right) \end{cases} \tag{II.11}$$

where $\Psi_{mn} = (k_n - k_m)\alpha - (\omega_n - \omega_m)t - \varphi_n + \varphi_m$. Note that in x_2, the subtraction of the sinus of the initial phase $\Psi_{mn}|_{t=0}$ is not directly derived from the second-order solution, but is introduced in Guérin *et al.* (2019) to take care of the divergence of x_2 when the frequency bandwidth becomes infinitely small. Similarly to the first-order solution, it is possible to show that the second-order solution leads to an irrotational flow (Nouguier *et al.*, 2015; Pierson, 1961). The formulation of the complete second-order solution for short-crested waves is derived by Nouguier *et al.* (2015).

The expression of these contributions becomes much simpler when we consider only a particle on the free surface (i.e., $\delta = 0$). Following the formalism of equation (II.10), it gives

$$\begin{cases} D_2(\alpha,t) = -\sum_{m<n}^{N} A_m A_n k_m \left(\frac{\omega_n+\omega_m}{\omega_n-\omega_m}\right)(\sin\Psi_{mn} - \sin\Psi_{mn}|_{t=0}) + \sum_{n=1}^{N} A_n^2 \omega_n k_n t, \\ Z_2(\alpha,t) = \sum_{m<n}^{N} A_m A_n k_m \cos\Psi_{mn} + \sum_{n=1}^{N} \frac{1}{2} A_n^2 k_n, \end{cases}$$

There are two terms in each contribution, the sums involving two indices m, n correspond of the wave-wave interaction effects of different components, the other sums are self-interaction terms. The horizontal self-interaction term is a non-periodic drift, linearly increasing with time. The factor in front of the time quantity corresponds to the sum of the Stokes drift (i.e., nonlinear drift experienced by the second-order regular waves) of all components. In the vertical contribution Z_2, the self-interaction term is a correction to the surface level, leading to retrieve zero mean. The Lagrangian solution to the second-order is denoted CWM2.

II.1.4 Highest Wave Steepness

When the wave steepness is large, Lagrangian solutions can lead to nonphysical events (Pierson, 1961), i.e., the particles trajectories on the free surface are making a loop (free surface elevation no longer single valued). Whenever the Lagrangian model is used, it is thus essential to verify that the wave field characteristics fit into their limitations in steepness. For a regular wave, the limiting steepness is $kA = 1$ (far above the theoretical highest steepness of a regular wave field), which corresponds to the case depicted in figure II.1. The absolute limiting steepness for a long crested irregular wave field is $\sum_{n=1}^{N} k_n A_n = 1$. However, this quantity corresponds to the steepness of the steepest wave that could appear in a wave field, i.e., if all components have their phase synchronized. Since in a typical ocean wave field the initial phases are random, uniformly distributed in $[0, 2\pi]$, this event is very unlikely to happen, and this limitation is generally overly conservative. In practice, it is sufficient to verify that the free surface particle position $\boldsymbol{r}$ is monotonically increasing with its reference position $\boldsymbol{\xi}$ during the simulation time over the calculation domain. For a long crested wave field, this condition follows

$$\frac{\partial x(\alpha,t)}{\partial \alpha} > 0 \iff -\sum_{i} \frac{\partial D_i(\alpha,t)}{\partial \alpha} < 1. \tag{II.12}$$

Note that the sum term in equation (II.12) represents a local steepness, or "effective steepness", which has to stay lower than one. Explicitly, the spatial derivatives of the first and second-order contributions are

$$\begin{aligned} \frac{\partial D_1(\alpha,t)}{\partial \alpha} &= \sum_{n=1}^{N} A_n k_n \cos(k_n\alpha - \omega_n t - \varphi_n), \\ \frac{\partial D_2(\alpha,t)}{\partial \alpha} &= -\sum_{m<n}^{N} A_m A_n k_m (k_n - k_m)\left(\frac{\omega_n+\omega_m}{\omega_n-\omega_m}\right)(\cos\Psi_{mn} - \cos\Psi_{mn}|_{t=0}). \end{aligned}$$

In the following numerical applications, we checked that our simulated wave fields respect criterion (II.12). For examples, figure II.2 presents the time evolution of the maximal effective steepness of a second-order unidirectional Lagrangian wave field for the sea states used in section II.3.3 for the numerical assessment of our improved Lagrangian model, namely a fully developed (i.e., broad banded) JONSWAP spectrum of characteristic steepness $H_s/\lambda_p \approx 4\%$, and a narrow-banded Gaussian spectrum of steepness 6%. It seems that, more than the characteristic steepness of the wave field, the frequency spreading of the underlying spectrum plays a significant

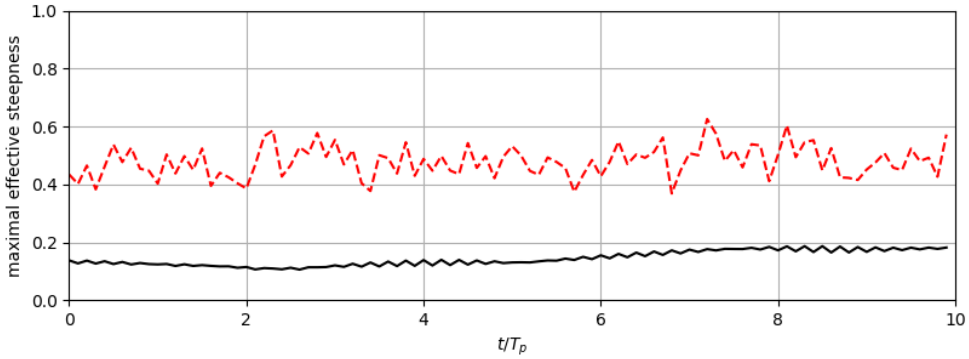

Figure II.2: Time evolution of the maximal effective steepness of unidirectional irregular waves for broad-banded JONSWAP (- - - -) and narrow-banded Gaussian (——) spectra, of characteristic steepness $H_s/\lambda_p \approx 4\%$ and 6%, respectively. A detailed description of both sea states and calculation domains can be found in section II.3.3.

role in the value of the maximal effective steepness, which is about 0.5 for our JONSWAP sea state. This might come from the fact that, in a wave field with a large spectral bandwidth, the probability for short waves to ride long waves, creating a local steep event, is enhanced.

II.1.5 Comparison to the Eulerian Approach

In order to characterize the Lagrangian results, it is interesting to determine the equivalent (however always approached) solution in the Eulerian framework, which is the basis of many of classical wave models. We will derive in this section the expansion of the Lagrangian solution up to the second-order for a regular wave in the Eulerian framework, and compare the results to a Stokes expansion.

The second-order Lagrangian free surface elevation of a regular wave (for the sake of simplicity, the wave is going along the x-axis, i.e., $y = \beta$)

$$\begin{cases} x(\alpha,t) = \alpha + D_1(\alpha,t) + D_2(\alpha,t) = \alpha - A\sin(k\alpha - \omega t) + A^2\omega k t, \\ z(\alpha,t) = Z_1(\alpha,t) + Z_2(\alpha,t) = A\cos(k\alpha - \omega t) + \dfrac{1}{2}A^2 k. \end{cases} \tag{II.13}$$

In order to simplify the following mathematical developments, we recast equation (II.13) to implicitly integrate the particle shift D_2 into the particle position coordinates. It follows

$$\begin{cases} \mathcal{X}(\alpha,t) = x(\alpha - D_2(\alpha,t),t) = \alpha + D_1(\alpha - D_2(\alpha,t),t) = \alpha + \mathcal{D}(\alpha,t) = \alpha - A\sin\phi, \\ \mathcal{Z}(\alpha,t) = z(\alpha - D_2(\alpha,t),t) = A\cos\phi + \dfrac{1}{2}A^2 k, \end{cases} \tag{II.14}$$

where $\phi = k\alpha - \omega_{\mathrm{L}} t$ and the modified angular frequency is $\omega_{\mathrm{L}} = \omega\left(1 + k^2A^2\right)$. Then, to approach this solution (II.14) in the Eulerian framework, we seek a function η_{L} that gives explicitly the free surface elevation at any abscissa α and satisfies $\eta_{\mathrm{L}}(\mathcal{X}) = \mathcal{Z}(\alpha)$. This will give us an explicit expression between $\mathcal{X}$ and $\mathcal{Z}$ (or equivalently x and z). Such a function can be approached with a Taylor series expansion about $\mathcal{X} = \alpha + \mathcal{D}(\alpha)$ as

$$\begin{aligned} \eta_{\mathrm{L}}(\alpha) &\approx \eta_{\mathrm{L}}(\mathcal{X}) + (\alpha - \mathcal{X})\left[\eta_{\mathrm{L}}(\mathcal{X})\right]' + \frac{(\alpha-\mathcal{X})^2}{2}\left[\eta_{\mathrm{L}}(\mathcal{X})\right]'' \\ &\approx \eta_{\mathrm{L}}(\mathcal{X}) - \mathcal{D}(\alpha)\left(\alpha + \mathcal{D}(\alpha)\right)'\eta_{\mathrm{L}}'(\mathcal{X}) + \frac{\mathcal{D}^2(\alpha)}{2}\left[\left(\alpha + \mathcal{D}(\alpha)\right)'\eta_{\mathrm{L}}'(\mathcal{X})\right]' \\ &\approx \mathcal{Z}(\alpha) - \mathcal{D}(\alpha)\left(1 + \mathcal{D}'(\alpha)\right)\mathcal{Z}'(\alpha) + \frac{\mathcal{D}^2(\alpha)}{2}\left[\mathcal{D}''(\alpha)\mathcal{Z}'(\alpha) + \left(1 + \mathcal{D}'(\alpha)\right)\mathcal{Z}''(\alpha)\right]. \end{aligned}$$

We look for an approximation to the second-order ϵ^2 (where $\epsilon = kA$ is the steepness). The derivative terms yield

$$\begin{cases} \mathcal{D}'(\alpha) = -\epsilon\cos\phi, \\ \mathcal{Z}'(\alpha) = -\epsilon\sin\phi, \end{cases} \qquad \begin{cases} \mathcal{D}''(\alpha) = \epsilon k\sin\phi, \\ \mathcal{Z}''(\alpha) = -\epsilon k\cos\phi, \end{cases}$$

and the approximated elevation turns into

$$\begin{aligned} \eta_{\mathrm{L}}(\alpha) &\approx A\left(\cos\phi + \frac{1}{2}\epsilon\right) - A\sin\phi\,(1-\epsilon\cos\phi)\,\epsilon\sin\phi + \frac{1}{2}A^2\sin^2\phi\,[-\epsilon k\sin\phi\epsilon\sin\phi - (1-\epsilon\cos\phi)\,\epsilon k\cos\phi] \\ &\approx A\left\{\cos\phi + \frac{1}{2}\epsilon - \sin\phi\,(1-\epsilon\cos\phi)\,\epsilon\sin\phi + \frac{1}{2}\epsilon\sin^2\phi\,[-\epsilon\sin\phi\epsilon\sin\phi - (1-\epsilon\cos\phi)\,\epsilon\cos\phi]\right\} \\ &\approx A\left[\cos\phi + \frac{1}{2}\epsilon - \epsilon\sin^2\phi\,(1-\epsilon\cos\phi) - \frac{1}{2}\epsilon\sin^2\phi\left(\epsilon^2\sin^2\phi + \epsilon\cos\phi - \epsilon^2\cos^2\phi\right)\right]. \end{aligned}$$

Discarding the third-order terms, we get

$$\begin{aligned} \eta_{\mathrm{L}}(\alpha) &\approx A\left[\cos\phi + \epsilon\left(\frac{1}{2} - \sin^2\phi\right) + \epsilon^2\left(-\sin^2\phi\cos\phi - \frac{1}{2}\sin^2\phi\cos\phi\right)\right] \\ &\approx A\left[\cos\phi + \frac{1}{2}\epsilon\cos(2\phi) - \frac{3}{2}\epsilon^2\sin^2\phi\cos\phi\right] \\ &\approx A\left[\cos\phi + \frac{1}{2}\epsilon\cos(2\phi) - \frac{3}{4}\epsilon^2\left(\cos\phi - \cos\phi\cos(2\phi)\right)\right] \\ &\approx A\left\{\cos\phi + \frac{1}{2}\epsilon\cos(2\phi) - \frac{3}{4}\epsilon^2\left[\cos\phi - \frac{1}{2}\left(\cos\phi + \cos(3\phi)\right)\right]\right\} \\ &\approx A\left[\left(1 - \frac{3}{8}\epsilon^2\right)\cos\phi + \frac{1}{2}\epsilon\cos(2\phi) + \frac{3}{8}\epsilon^2\cos(3\phi)\right]. \end{aligned} \tag{II.15}$$

From equation (II.15), we notice that the second-order Lagrangian expansion allows to retrieve the complete expression of the Stokes third-order expansion. Note that, except the mean water level, the wave shape modification only comes from the first-order Lagrangian terms. However, they manifest different nonlinear dispersion relations leading to different nonlinear wave velocities. The Lagrangian phase velocity can be determined following

$$c_{\mathrm{L}} = \frac{\omega_{\mathrm{L}}}{k} = \frac{\omega}{k}\left(1+\epsilon^2\right),$$

whereas the third-order Eulerian phase velocity is

$$c_E = \frac{\omega}{k}\left(1 + \frac{1}{2}\epsilon^2\right).$$

The correction in the second-order Lagrangian solution is twice the one from the third-order Eulerian expansion. This was already noticed by Pierson (1961) but could not be explained. A novel approach to address this issue is proposed in Guérin *et al.* (2019) and explained in the next section.

It can be shown that the second-order solution derived from the Eulerian approach for a short-crested irregular wave field is fully recovered by the second-order Lagrangian solution (Nouguier *et al.*, 2015). Specifically, only the first-order terms $(\boldsymbol{D}_1, Z_1)$ and second-order horizontal displacement Z_2 (including all wave-wave interaction terms) are needed to consistently represent the second-order Eulerian solution. All effects derived from $\boldsymbol{D}_2$ correspond to Eulerian third- and higher-order effects.

II.2 Improved Choppy Wave Model

As shown in the previous section, the second-order Lagrangian solution does not recover all third-order Eulerian features, specifically the nonlinear correction of the wave velocity. This property influences the phase shift of wave components, it is thus of significant importance in wave propagation process and has to be correctly modeled for an accurate free surface elevation prediction. In this section, we explain the principle behind the derivation of a Lagrangian solution that leads to a modified (or "improved") nonlinear wave velocity correction that is latter shown to be more accurate.

II.2.1 Derivation of an "Improved" Solution

The complete derivation of the Improved Choppy Wave Model (ICWM) is provided in Guérin *et al.* (2019). In this section, we briefly explain the principle of the model correction and provide its final expression.

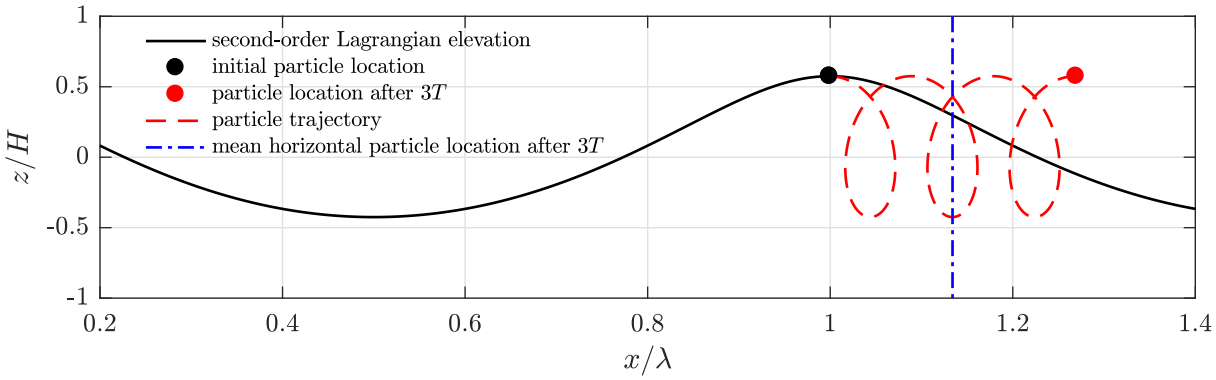

Figure II.3: Particle horizontal shift for a periodic wave with steepness $\epsilon = kA = 0.3$ according to the Lagrangian second-order formulation. As time increases, the distance between the initial particle location and its position grows two times faster than the distance between the initial particle location and its mean horizontal location.

As noticed in the previous section, the nonlinear correction in the dispersion relation from the Lagrangian approach leads to a discrepancy of the nonlinear correction of the wave phase velocity compared to the results from a third-order Eulerian expansion. Even if this result was already noticed by Pierson (1961), he did not provide an explanation of its origin. We have recently shown in Guérin *et al.* (2019) that this discrepancy comes from the fact that the particle reference location for the Lagrangian expansion, set to its initial location $\boldsymbol{\zeta}$, leads the particle to undertake large displacements with respect to $\boldsymbol{\zeta}$. Figure II.3 shows the time evolution of a free surface particle for a periodic wave of steepness $\epsilon = kA = 0.3$ according to the second-order Lagrangian solution. After three periods of propagation, the particle has already been shifted about 0.25λ away its initial location. The idea proposed in Guérin *et al.* (2019) to, at least partly, solve this drifting issue is to perform the Lagrangian expansion about a reference location that minimizes the particle horizontal fluctuation. Adopting Guérin's notation, the particle horizontal motion is expressed as $X(t) = \overline{x} + x(t)$ where the reference location $\overline{x}$ is a linear function of the particle horizontal shift. The magnitude of the horizontal particle fluctuation over a given number n of motion cycles (i.e., over time interval $[0, nT]$ with T the wave period), quantified by the following mean square distance to the reference location

$$\langle x^2 \rangle = \frac{1}{nT} \int_0^{nT} x^2(t)\, \mathrm{d}t,$$

is minimized for the optimal rate of horizontal drift of the reference location. With the optimal

value, we find

$$\overline{x} = \frac{1}{nT}\int_0^{nT} X\left(t\right)\mathrm{d}t,$$

which means that the reference location corresponds to the average particle motion over the completed cycles. For instance, in figure II.3, the particle reference location coincides with the vertical blue line after three completed cycles. For a periodic wave, this result leads to retrieve a Lagrangian solution that exhibits the same nonlinear wave velocity correction than the one appearing at the third-order Eulerian expansion.

Moreover, and as it will be shown in section II.3.3, despite their onerous computational complexity, the second-order wave-wave interaction terms of different components in equation (II.11) have a very limited influence on the solution compared to the self-interaction terms, especially after a long time of propagation. Hence, in order to retain only the relevant effects for the long-time propagation of a wave field, we choose in the following to only retain the horizontal drift term and correction to the mean free surface elevation level in the formulation of ICWM. It follows

$$\begin{cases} \boldsymbol{r}\left(\boldsymbol{\xi},t\right) = \boldsymbol{\xi} - \displaystyle\sum_{n=1}^{N} A_n \hat{\boldsymbol{k}}_n \sin\left(\boldsymbol{k}_n\cdot\boldsymbol{\xi} - \tilde{\omega}_n t - \varphi_n\right) + \boldsymbol{\mathcal{U}}_s t, \\ z\left(\boldsymbol{\xi},t\right) = \displaystyle\sum_{n=1}^{N} A_n \cos\left(\boldsymbol{k}_n\cdot\boldsymbol{\xi} - \tilde{\omega}_n t - \varphi_n\right) + \sum_{n=1}^{N}\frac{1}{2}A_n^2 k_n, \end{cases}$$

where $\tilde{\omega}_n = \omega_n - 1/2\boldsymbol{k}_n\cdot\boldsymbol{\mathcal{U}}_s$ and $\boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N} A_n^2\omega_n\boldsymbol{k}_n$ is the Stokes drift vector. In order to simplify the following mathematical developments pertaining to the wave model inversion, the equations are rewritten in the form

$$\begin{cases} \boldsymbol{r}\left(\boldsymbol{\xi},t\right) = \boldsymbol{\xi} + \displaystyle\sum_{n=1}^{N} \hat{\boldsymbol{k}}_n\left(-a_n\sin\tilde{\psi}_n + b_n\cos\tilde{\psi}_n\right) + \boldsymbol{\mathcal{U}}_s t, \\ z\left(\boldsymbol{\xi},t\right) = \displaystyle\sum_{n=1}^{N}\left(a_n\cos\tilde{\psi}_n + b_n\sin\tilde{\psi}_n\right) + \sum_{n=1}^{N}\frac{1}{2}\left(a_n^2+b_n^2\right)k_n, \end{cases} \tag{II.16}$$

where $\tilde{\psi}_n = \boldsymbol{k}_n\cdot\boldsymbol{\xi} - \tilde{\omega}_n t$, $\boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N}\left(a_n^2+b_n^2\right)\omega_n\boldsymbol{k}_n$, and $(a_n, b_n) = (A_n\cos\varphi_n, A_n\sin\varphi_n)$ are wave parameters describing the ocean surface.

We illustrate in figure II.4 the difference between the surface elevation associated with the LWT and all the wave models described in this section (i.e., CWM1, CWM2 and ICWM) for a regular wave of steepness $H/\lambda \approx 0.04$ (i.e., $\epsilon = kA \approx 0.125$) which stands in the validity domain of the Lagrangian approach (Pierson, 1961). A proper qualification of the models' characteristics is done in section II.3, and in Guérin *et al.* (2019). Figure II.4a shows that at initial time (i.e., $t = 0$), CWM1 corresponds to a simple geometrical transformation of the linear surface (LWT): the crest is sharper and the trough flatter. Since time-dependent terms vanish, CWM2 and ICWM give the exact same elevation, corresponding to a CWM1 surface with a mean surface level correction. After approximately four periods of propagation (see figure II.4b), a phase shift between CWM1, CWM2 and ICWM is already quite noticeable. While CWM1 keeps the same phase value as the linear surface, CWM2 exhibits a nonlinear phase shift which is twice the one experienced by ICWM. After four periods, the spatial horizontal shift is approximately 0.03λ.

Similarly, it is shown in figure II.5 the free surface elevation of an irregular long-crested wave field (JONSWAP spectrum with peakedness $\gamma = 3.3$, equation (I.2)) of characteristic steepness $H_s/\lambda_p \approx 0.04$ (i.e., $k_p H_s/2 \approx 0.125$) at initial time, and then after four peak periods of propagation. Like the periodic wave, we observe a surface shape modification leading to sharper crests

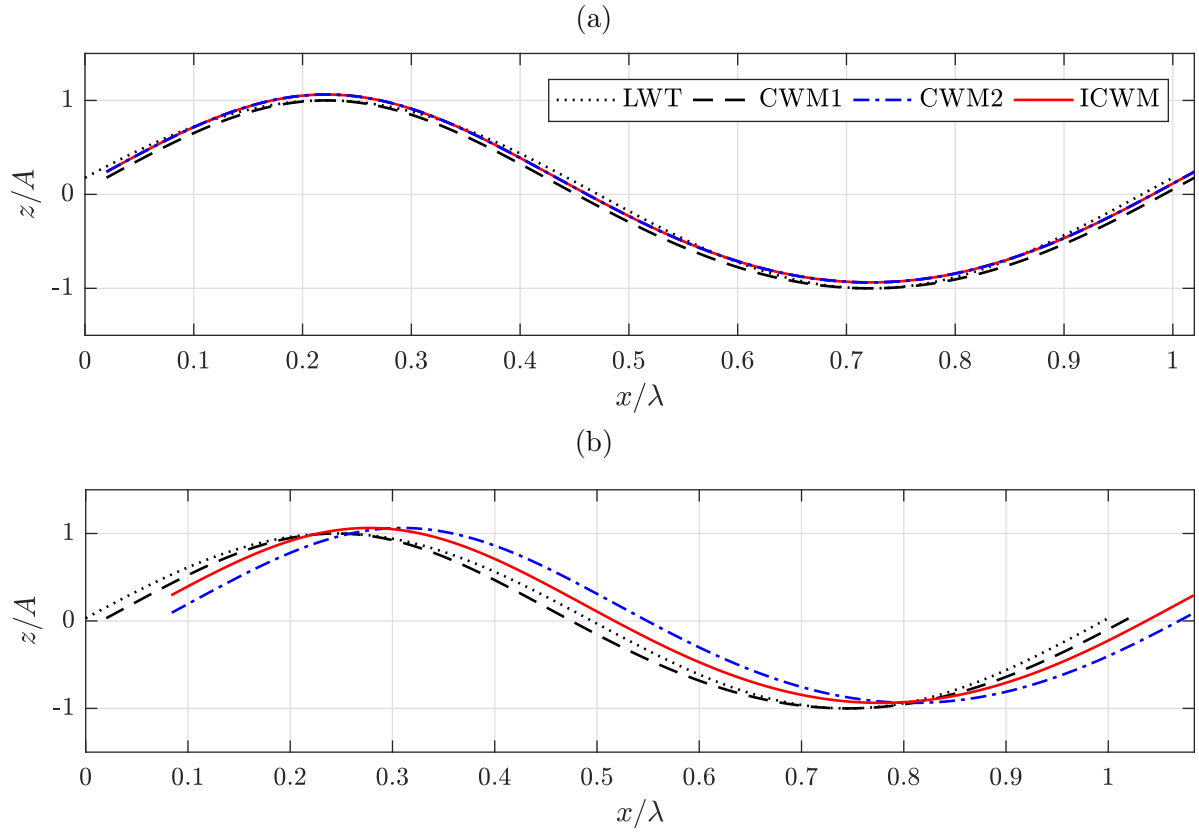

Figure II.4: Elevation associated with LWT, CWM1, CWM2 and ICWM for a periodic wave of steepness $\epsilon = kA \approx 0.125$ (a) at initial time and (b) after approximately four periods of propagation.

and flatter troughs for CWM1 compared to LWT, and a correction of the mean surface level for CWM2 and ICWM. However, CWM2 and ICWM are no longer strictly equivalent at initial time, since CWM2 includes wave-wave interaction effects (barely noticeable on the figure II.5b) of different components that are neglected in ICWM. Even after the relatively short propagation time of four peak periods, nonlinear phase effects are clearly visible, suggesting that they play a crucial role in deterministic prediction. As expected, CWM2 is shifted by twice the amount of ICWM.

II.2.2 Simple Approximate of ICWM in Eulerian Form

Optical measurements of the ocean surface are expected to provide, at each time, the surface elevation measured for a set of spatial points at irregular but defined Eulerian locations in a reference coordinate system. Hence, the wave model used in the reconstruction method must also be able to provide and use comparable information. The Lagrangian form of ICWM (II.16) is not directly usable, and we must find an approximate equivalent model in the Eulerian framework. As explained in section II.1.5, we have to find an explicit relationship between $\boldsymbol{r}$ and z, but this time, we want the formulation to be as simple as possible to avoid losing the benefit of the Lagrangian formulation's simplicity. In order to formulate the approximation, it is more convenient to implicitly incorporate the particle horizontal shift $\boldsymbol{\mathcal{U}}_s t$ to a modified angular frequency, in the

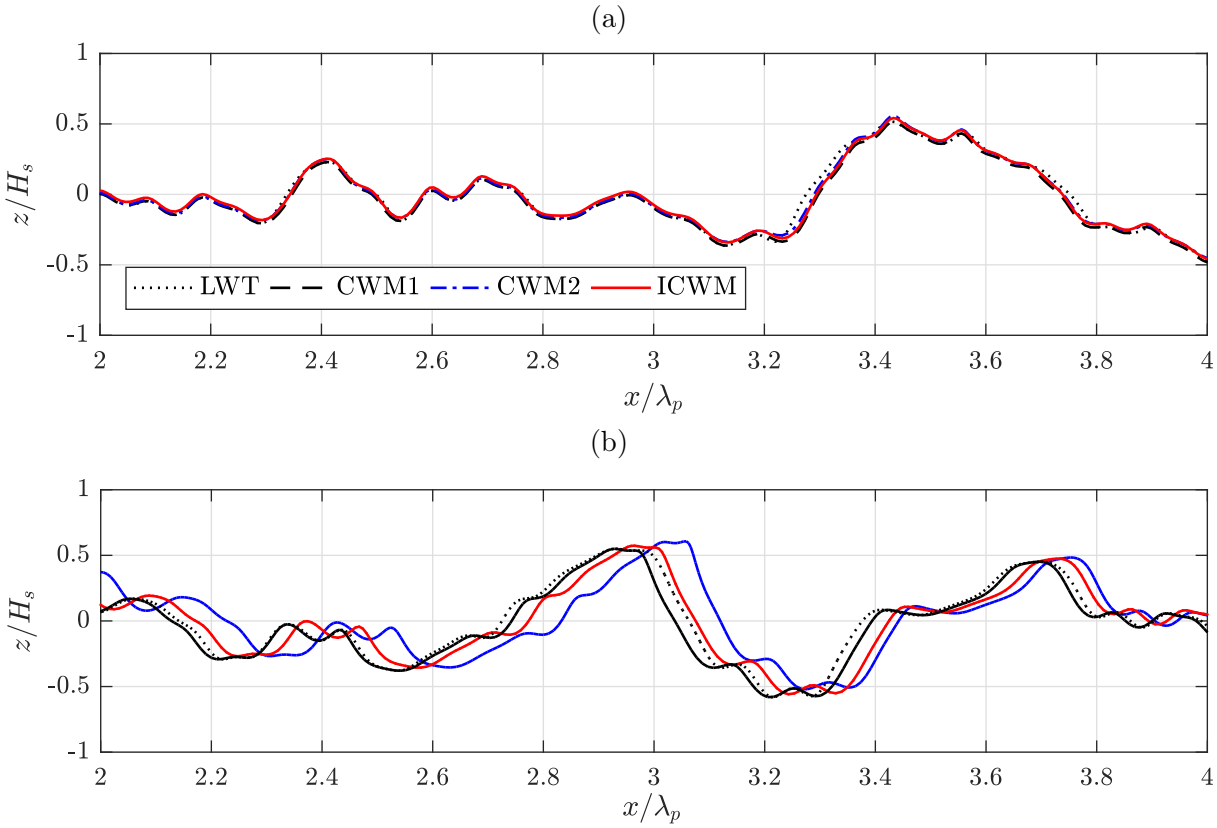

Figure II.5: Elevation associated with LWT, CWM1, CWM2 and ICWM for an irregular wave field (JONSWAP spectrum with peakedness $\gamma = 3.3$) of characteristic steepness $H_s/\lambda_p \approx 0.04$ (i.e., $k_p H_s/2 \approx 0.125$) (a) at initial time and (b) after four peak periods of propagation.

same way as in section II.1.5. From equation (II.16), we get

$$\begin{cases} \boldsymbol{\mathcal{R}}\left(\boldsymbol{\xi},t\right) = \boldsymbol{r}\left(\boldsymbol{\xi} - \boldsymbol{\mathcal{U}}_s t, t\right) = \boldsymbol{\xi} + \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}\right) = \boldsymbol{\xi} + \displaystyle\sum_{n=1}^{N} \hat{\boldsymbol{k}}_n \left(-a_n \sin\tilde{\phi}_n + b_n \cos\tilde{\phi}_n\right), \\ \mathcal{Z}\left(\boldsymbol{\xi},t\right) = z\left(\boldsymbol{\xi} - \boldsymbol{\mathcal{U}}_s t, t\right) = \displaystyle\sum_{n=1}^{N} \left(a_n \cos\tilde{\phi}_n + b_n \sin\tilde{\phi}_n\right) + \sum_{n=1}^{N} \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n, \end{cases}$$

where $\tilde{\phi}_n = \boldsymbol{k}_n \cdot \boldsymbol{\xi} - \tilde{\omega}_{\mathrm{L}n} t$ and $\tilde{\omega}_{\mathrm{L}n} = \omega_n + 1/2 \boldsymbol{k}_n \cdot \boldsymbol{\mathcal{U}}_s$ A simple and direct way to proceed to the approximation is to consider that the particle vertical displacement can be calculated at the particle position instead of its reference location, following

$$\mathcal{Z}\left(\boldsymbol{\xi}\right) = \mathcal{Z}\left(\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}\right)\right) \approx \mathcal{Z}\left(\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{\mathcal{R}}\right)\right) = \tilde{\eta}\left(\boldsymbol{\mathcal{R}}\right). \tag{II.17}$$

Accordingly, we now formulate explicitly the approximate ICWM free surface elevation at any spatial point $\boldsymbol{r}$ following

$$\eta\left(\boldsymbol{r},t\right) = \sum_{n=1}^{N} \left(a_n \cos\Psi_n + b_n \sin\Psi_n\right) + \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n, \tag{II.18}$$

$$\text{in which} \quad \Psi_n = \boldsymbol{k}_n \cdot \left[\boldsymbol{r} - \sum_{i=1}^{N} \hat{\boldsymbol{k}}_i \left(-a_i \sin\tilde{\phi}_i + b_i \cos\tilde{\phi}_i\right)\right] - \tilde{\omega}_{\mathrm{L}n} t.$$

As detailed by Gjosund (2003), the discrepancy introduced in equation (II.17) can be reduced using a convergence scheme, leading to the determination of the actual reference particle location

$\boldsymbol{\xi}_c$ that should be taken to find the appropriate particle displacements at specified location $\boldsymbol{r}$, following

$$\begin{cases} \boldsymbol{\xi}_c^{(0)} = \boldsymbol{r}, \\ \boldsymbol{\xi}_c^{(q+1)} = \boldsymbol{r} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}_c^{(q)}\right), \end{cases} \tag{II.19}$$

where q is the iteration number. With converged results, we would have $\mathcal{Z}\left(\boldsymbol{\xi}_c\right) = \tilde{\eta}\left(\boldsymbol{r}\right)$. However, going through this convergence scheme is not appropriate for the real-time constraint of the prediction process, since it requires a wave model inversion at each step. Moreover, we noticed that the initial approximation (II.17) (i.e., $\boldsymbol{\xi}_c = \boldsymbol{r}$) leads to a very small error compared to the other sources of error in the wave model inversion process. Figure II.6 shows elevations from both the Lagrangian and Eulerian results for a regular wave. Even for the depicted large steepness

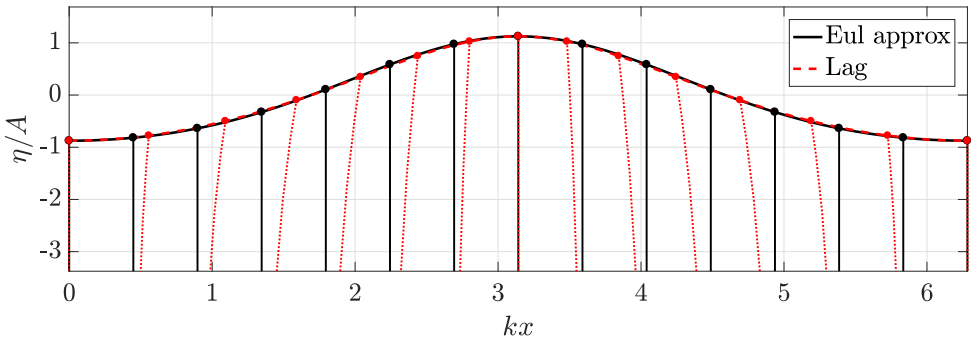

Figure II.6: Comparison of a Lagrangian second-order surface elevation (- - - -) and its Eulerian approximation (——) for a regular wave of steepness $H/\lambda \approx 8\%$ ($kA = 0.25$). Lagrangian reference particle locations in the water column (vertical black lines) as well as displaced locations (red dotted lines) are plotted to highlight the imposed displacements.

$H/\lambda \approx 8\%$ ($kA = 0.25$), the agreement between the two formulations is very good. Accordingly, it was shown by Grilli *et al.* (2011) that this error is on the order of the mean square slope of the surface. Hence, in the following, we will stick to the ICWM formulation (II.18), which leads to the surface elevation calculation with the computational cost of only two IFFTs, provided that wave amplitude coefficients are spectrally distributed in a suitable way.

ICWM formulation (II.18) includes a nonlinear correction of the deep water dispersion relation following

$$\tilde{\omega}_{\mathrm{L}n} = \sqrt{gk_n} + \frac{1}{2}\boldsymbol{k}_n \cdot \boldsymbol{\mathcal{U}}_s. \tag{II.20}$$

This correction induces modifications on both the wave phase and group velocities. The corresponding phase velocity gives

$$c_n = \frac{\tilde{\omega}_{\mathrm{L}n}}{k_n} = \sqrt{\frac{g}{k_n}} + \frac{1}{2}\hat{\boldsymbol{k}}_n \cdot \boldsymbol{\mathcal{U}}_s,$$

and similarly the corresponding group velocity is

$$c_{gn} = \frac{\partial \tilde{\omega}_{\mathrm{L}n}}{\partial k_n} = \frac{1}{2}\sqrt{\frac{g}{k_n}} + \frac{1}{2}\hat{\boldsymbol{k}}_n \cdot \sum_{m \neq n}^{N} \omega_m \boldsymbol{k}_m \left(a_m^2 + b_m^2\right) + \omega_n k_n \left(a_n^2 + b_n^2\right).$$

Note that the nonlinear correction of the group velocity is always larger than the correction of the phase velocity.

II.2.3 Comparison Between ICWM and Eulerian Phase Velocity Corrections

The Eulerian third-order phase velocity correction terms for short-crested waves are given by equations (I.11) and (I.12), corresponding to the Stokes correction term and a modulation correction, respectively. In the case of long-crested waves, it simplifies to

$$\Delta C_{pn}^{E} = \frac{1}{2}\omega_n k_n A_n^2 + \sum_{m\neq n}^{N} \omega_m \min\left(k_m, k_n\right) A_m^2.$$

The modulation term is the sum term in the above equation and corresponds to the modulation of the short waves by the longer waves: as the wave frequency increases, the value of this term monotonically increases as well. On the other hand and as shown in the previous section, the correction introduced in the Improved Choppy Wave Model is independent of the properties of individual wave components and is given by

$$\Delta C_{p}^{L} = \frac{1}{2}\sum_{m=1}^{N} \omega_m k_m A_m^2,$$

which is the sum of every wave component's Stokes correction terms. For a particular wave, the difference between those two corrections is

$$\begin{aligned}\zeta_n = \Delta C_p^L - \Delta C_{pn}^E &= \frac{1}{2}\sum_{m=1}^{N} \omega_m k_m A_m^2 - \frac{1}{2}\omega_n k_n A_n^2 - \sum_{m\neq n}^{N} \omega_m \min\left(k_m, k_n\right) A_m^2 \\ &= \sum_{m\neq n}^{N} \omega_m A_m^2 \left(\frac{1}{2}k_m - \min\left(k_m, k_n\right)\right).\end{aligned}$$

We already see that this difference is bounded between two extreme values, $\zeta_0 = \frac{1}{2}\mathcal{U}_s$ for a frequency approaching zero and $\zeta_\infty = -\frac{1}{2}\mathcal{U}_s$ for a frequency approaching infinity, with $\mathcal{U}_s = \sum \omega_m k_m A_m^2$. Compared to the Eulerian developments, our improved Lagrangian solution will tend to overestimate and underestimate the phase velocity correction of low and high frequencies, respectively.

Figure II.7 shows as a function of the wavenumber the evolution of correction difference $\zeta\left(k\right)$ between the Lagrangian and Eulerian phase velocity corrections, considering the second order (CWM2) and improved (ICWM) Lagrangian formulations, and normalized by the improved Lagrangian phase velocity correction $\Delta C_p^L \left(= \frac{1}{2}\mathcal{U}_s\right)$. The wave spectrum is given on the same figure to facilitate interpretations. We see that ζ crosses zero around the peak of the spectrum for ICWM, which shows that the estimation of the elevation is globally correct (assuming that the third-order Eulerian correction is accurate) for energetic wave components. In contrast, CWM2 always overestimates the phase velocity correction, with ζ approximately equal to ΔC_p^L around the peak and going to zero for high frequencies.

Figure II.8 presents the same quantity for the improved Lagrangian formulation and for two types of wave spectrum, namely a JONSWAP with a peakedness parameter $\gamma = 1$ and a Gaussian spectrum with a normalized standard deviation $\sigma/\omega_p \approx 0.08$. For the JONSWAP spectrum, even if the peak enhancement factor γ is lower, similar results to that of previous paragraph are found. In the Gaussian case, every phase velocity correction is underestimated compared to the Eulerian correction, and is approximately equal to $-0.5\Delta C_p^L$ at the peak wavenumber.

II.3 Numerical Assessment of ICWM

The accuracy and efficiency of the proposed improved second-order Lagrangian model (ICWM), to perform the time updating of nonlinear surface waves, are numerically assessed by comparing results to those of a fully nonlinear reference model based the HOS method (see section I.3.4.1 for

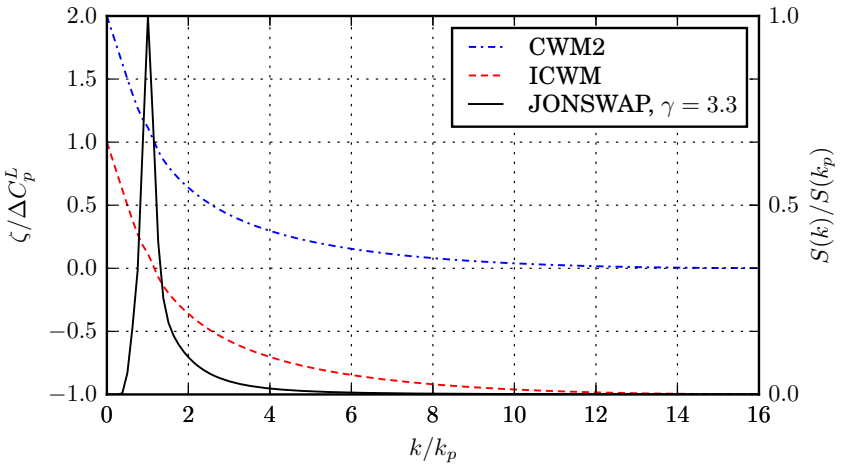

Figure II.7: Evolution of the phase velocity correction difference between the (classical and improved) Lagrangian second-order and Eulerian third-order expansion results (dashed lines, left axis), and wave energy spectrum (continuous line, right axis) as a function of the wavenumber. The spectrum is a JONSWAP with a peakedness parameter $\gamma = 3.3$.

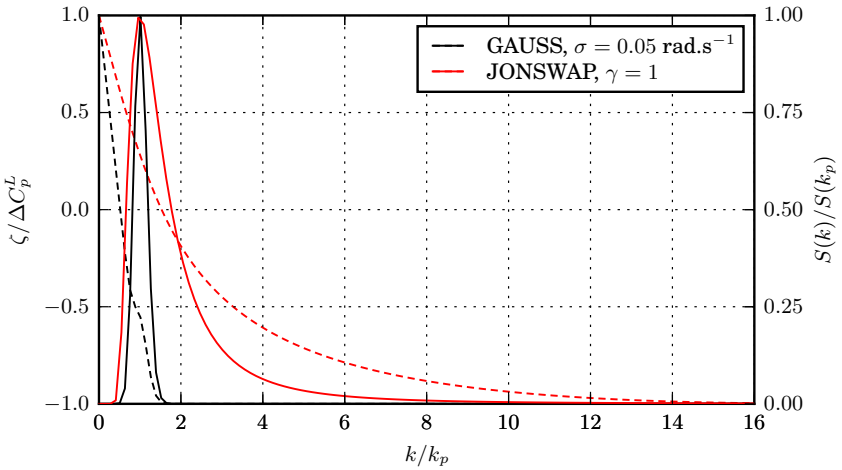

Figure II.8: Evolution of the phase velocity correction difference between the improved Lagrangian second-order and Eulerian third-order expansions (dashed lines, left axis), and wave energy spectra (continuous lines, right axis) as a function of the wavenumber. The spectra follow JONSWAP ($\gamma = 3.3$) and Gaussian ($\sigma/\omega_p \approx 0.08$) formulations.

model details). In this comparison, we also include the complete original second-order Lagrangian model (CWM2, equation (II.11)), where the terms corresponding to wave-wave interactions of different frequencies in x_2 have been discarded. Finally, for completeness, we also compare results to those of a model based on linear wave theory (LWT).

There are at least three conceptual difficulties when comparing the time evolution of phased-resolved Lagrangian and HOS surfaces, namely

(a) the selection of sampling points: Lagrangian and Eulerian surfaces are not evaluated on the same grids;

(b) the selection of initial phases for an irregular surface: these should be identical in all methods, but cannot in practice be strictly so;

(c) non-stationary statistics: because of the nonlinear interactions, the wave spectrum is evolving in time and thus cannot be prescribed except at initial time.

These issues will be further elaborated on below and partial solutions will be implemented in order to achieve a meaningful comparison between the different methods and model results.

II.3.1 Reference Dataset and Error Definition

A synthetic reference nonlinear ocean surface is first generated using the open source code 'HOS-ocean'[1] (Ducrozet *et al.*, 2016), which can accurately simulate the propagation of nonlinear waves over large spatio-temporal domains, up to a specified order in wave steepness. Since its initial development by Dommermuth & Yue (1987) and West *et al.* (1987), the HOS method has been extensively validated and used in various applications, such as the simulation of freak waves (e.g., Ducrozet *et al.*, 2007), the implementation of a numerical wave tank (e.g., Ducrozet *et al.*, 2012), or ocean wave field reconstructions based on measurements (see section I.3.4.2). HOS-ocean solves the fully nonlinear potential flow problem as a function of time, in terms of N complex amplitudes $A_n(t)$, for many wave components n, which are then used to reconstruct the ocean surface at an arbitrary point (α, t) (for long-crested waves) as

$$\eta_{\mathrm{HOS}}(\alpha, t) = \mathrm{Re}\left(\sum_{n=1}^{N} A_n(t)\, e^{\mathrm{i}k_n\alpha}\right).$$

In the following applications, we use a fifth-order HOS model with 32 grid points per peak wavelength λ_p, which was verified to provide converged results in terms of order of nonlinearity as well as discretization. The three conceptual difficulties are addressed in the following ways.

(a) We compare the time evolution of the relative root-mean-square (RMS) difference between the Lagrangian surface η_{LAG} (which can be generated either with CWM2 or ICWM) and the HOS surface elevation η_{HOS} at the same points. Given that the Lagrangian surface is implicitly defined by the surface particle locations, it is much easier to perform this comparison on the Lagrangian grid (that is, the ensemble of points $(x(t), z(t))$) rather than the Eulerian grid $(\alpha, \eta(\alpha))$ which requires an explicit formulation of the Lagrangian surface (which, referring to section II.2.2, can only be attained in an approximate way). The RMS difference between a free surface representation and the HOS solution is thus defined as

$$\epsilon_{\mathrm{LAG}}(t) = \frac{1}{\mathrm{rmsh}}\left[\frac{1}{L}\int_0^L \left(z(t) - \eta_{\mathrm{HOS}}(x(t), t)\right)^2 \mathrm{d}x(t)\right]^{1/2} \tag{II.21}$$

where L is the computational domain length and the normalization constant 'rmsh' is defined as

$$\mathrm{rmsh} = \left(\frac{1}{L}\int_0^L \eta^2_{\mathrm{HOS}}(\alpha, 0)\,\mathrm{d}\alpha\right)^{1/2} = \left(\frac{1}{2}\sum_{n=1}^{N} |A_n(0)|^2\right)^{1/2}$$

and represents the RMS elevation of the initial surface, which we also express as a function of spectral amplitudes $A_n(0)$.

(b) Given an energy density spectrum, the phases must be provided for each wave component to generate an initial free surface, which is then propagated in the various models. A rigorous comparison of time-evolving surfaces should thus be based on identical sets of (random or deterministic) initial phases. However, such phases can only be prescribed for linear surfaces but not for Lagrangian or HOS surfaces, which undergo a nonlinear transformation. To mitigate possible discrepancies arising from the choice of initial phases, the HOS model is initialized with an explicit nonlinear solution, which is close to the Lagrangian solution, namely the second-order Eulerian Stokes solution for irregular wave

[1] https://github.com/LHEEA/HOS-ocean

fields as detailed in section I.3.3.1. The relevance of this method has been investigated by Perignon (2011). The same underlying linear surface can be used to generate Lagrangian surfaces in the second-order expansions (CWM2, ICWM).

This method consists in approaching a nonlinear HOS solution from the spectral distribution of linear waves that also serves for the calculation of the second-order Lagrangian surface elevations. Another approach would have been to retrieve the free wave components from a fully nonlinear HOS simulation through a free-waves decoupling procedure, such as the one based on the Eulerian second-order expansion proposed by Duncan & Drake (1995) and used by Blondel *et al.* (2010) for second-order and enhanced second-order waves reconstruction. In our case, however, it would have required an adaptation to the Lagrangian formulation, which is not trivial and would have been the object of a proper study.

(c) Given the initial wave spectrum of an irregular surface, the corresponding free surface elevation in the HOS model follows a modified energy density spectrum, due to nonlinear wave-wave interactions and the redistribution of energy that occurs during wave propagation. Hence, free surface statistics are not stationary and thus not fully controlled during the nonlinear time evolution. By contrast, our Lagrangian models, which formulate nonlinear properties from a prescribed underlying linear spectrum, have stationary statistics (i.e., their wave spectrum should not evolve in time, even though it is modified by wave interaction terms). Hence, a meaningful comparison between time-evolving surfaces in different models can only be performed within smaller time intervals, within which statistics (essentially the energy density spectrum) can be assumed to be stationary. A measure of differences between the initial spectrum and the evolving spectrum of HOS surfaces is given by the RMS difference

$$\epsilon_{\text{spec}}(t) = \frac{\left(\sum_{n=1}^{N}\left(|A_n(t)| - |A_n(0)|\right)^2\right)^{1/2}}{\left(\sum_{n=1}^{N}|A_n(0)|^2\right)^{1/2}} \tag{II.22}$$

This measure thus provides a lower threshold to the error of Lagrangian methods $\epsilon_{\text{LAG}}(t)$ (defined by equation (II.21)), which therefore provides a relevant indicator of the method performance only if, $\epsilon_{\text{spec}}(t) << \epsilon_{\text{LAG}}(t)$. In practice, for the typical wave spectra under consideration and using the initialization procedure explained in point (b), it was found that this is the case for an evolution time smaller than about 20 to $40\,T_p$, depending on the wave steepness and on the spectral bandwidth.

II.3.2 Periodic Waves

We first consider a strongly nonlinear periodic wave of steepness $H/\lambda \approx 0.08$ and period T, where $H = 2A$ is the wave height and λ the wavelength. For linear deep water waves, $T = \sqrt{2\pi\lambda/g}$. This steepness is over half the deep water limiting steepness and corresponds to $kA \approx 0.25$, which is in the domain of applicability of the Lagrangian approach (Pierson, 1961). For periodic waves, free surfaces represented in the various models can be easily synchronized at initial time, using the same phase as that of the fundamental component of the (stationary) HOS solution. A domain of length equal to one wavelength is used to evaluate the error.

Figure II.9a shows that, at $t = 0$, the CWM2 and ICWM surface elevations are identical since time dependent terms vanish. Both solutions also match quite well the HOS solution (very small errors can be seen in the figure), and much better than the linear solution, whose error is $\epsilon_{\text{LAG}}(0) \approx 0.15$, due to an inaccurate wave shape. In figure II.9a, the errors for the CWM2 and linear model follow a parallel growth: while the linear wave does not experience nonlinear phase velocity corrections, CWM2 is penalized by an overestimated nonlinear correction (close to twice

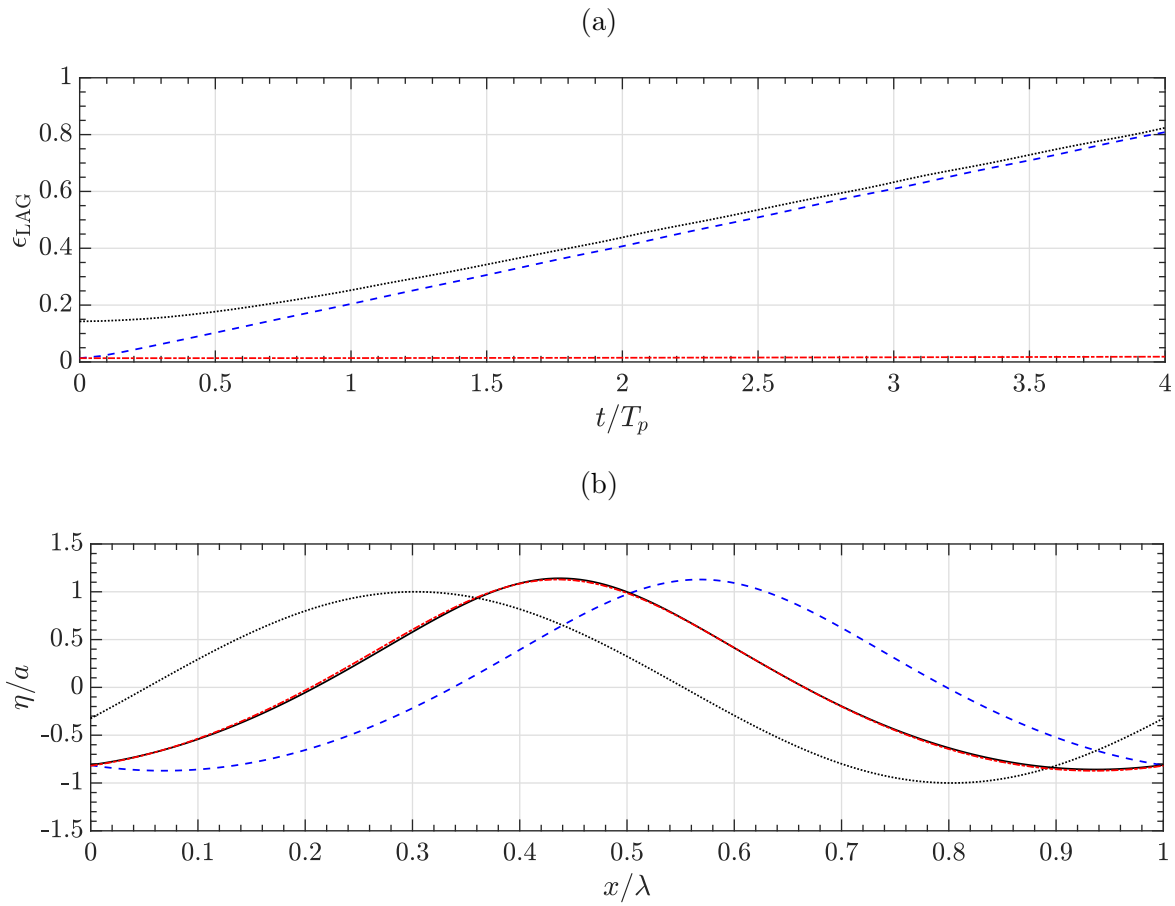

Figure II.9: Comparison of results computed for a periodic wave of steepness $H/\lambda \approx 8\%$, with the HOS (——), linear (·····), CWM2 (- - -), and ICWM (-·-·-) models: (a) error as a function of time with respect to the HOS solution and (b) normalized free surface elevations after approximately four periods of propagation.

the correct value, as discussed in Section II.1.5). By contrast, for ICWM, the improved estimate of the nonlinear phase shift allows for errors to remain very small throughout the four-period propagation time considered here. At this stage, figure II.9b shows a very marked phase shift of the linear and CWM2 model results with respect to the HOS reference solution, whereas no measurable shift is seen for the ICWM solution.

II.3.3 Irregular Waves

The performance of the improved Lagrangian formulation is now assessed for long-crested irregular waves generated on the basis of a specified energy density spectrum $E(\omega)$. Similar to the periodic wave case, we investigate the time evolution of the error of the CWM2, ICWM, and linear solution, with respect to the HOS reference solution. Given a sea state of parameters (H_s, T_p), the influence of energy spreading around the peak spectral frequency on the performance of the various wave models is investigated by considering two different spectra.

II.3.3.1 Spectra Definitions

First, we consider a narrow-banded Gaussian spectrum whose definition follows

$$E(\omega) = \alpha_s \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(\omega-\omega_p)^2}{2\sigma^2}},$$

where the standard deviation is $\sigma/\omega_p \approx 0.08$, in which $\omega_p = 2\pi/T_p$ denotes the peak angular frequency. The second spectrum is a JONSWAP with peakedness parameter $\gamma = 1$, explicitly defined by

$$E(\omega) = \alpha_s \frac{g^2}{\omega^5} \mathrm{e}^{-\frac{5}{4}(\frac{\omega_p}{\omega})^4}.$$

It is similar to the Pierson-Moskowitz spectrum (Pierson Jr. & Moskowitz, 1964) of a fully developed sea. In both cases, parameter α_s is found based on a specified H_s value, defined from the surface elevation standard definition as

$$\int_0^\infty E(\omega)\,\mathrm{d}\omega = \frac{H_s^2}{16}.$$

In the applications, we use $T_p = 10$ s with $H_s = 9$ m for the Gaussian spectrum and $H_s = 6$ m for the JONSWAP spectrum. This yields characteristic deep water steepnesses $H_s/\lambda_p \approx 6\%$ (i.e., very nonlinear, close to wave breaking) and 4%, respectively, with $\lambda_p = 2\pi/k_p$ denoting the peak spectral wavelength. The corresponding spectra are shown in figure II.10.

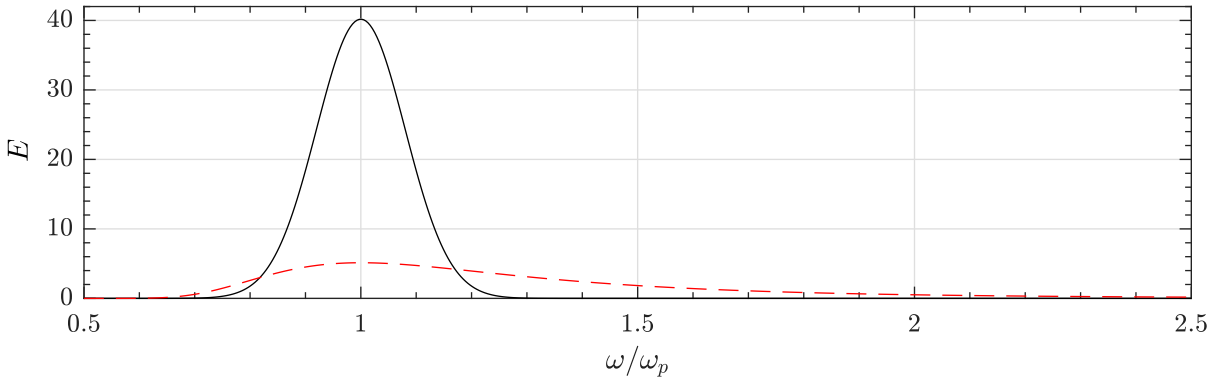

Figure II.10: Energy density spectrum as a function of the dimensionless angular frequency ω/ω_p, for a narrow (Gauss, solid line) or broad-banded (JONSWAP, dashed line) spectrum. The characteristic steepness is $H_s/\lambda_p \approx 6\%$ and 4% for each case, respectively.

For irregular waves, a larger number of dominant waves than in the periodic case must be modeled in the computational domain, in order to achieve convergence of the integral error of equation (II.21) Here we use a spatial domain spanning eight peak wavelengths, i.e., $L = 8\lambda_p$. At $t = 0$, the linear and CWM free surfaces are constructed from the linear spectral information using the same set of random phases. For the HOS model, however, unlike periodic waves, linear information cannot be retrieved (except at initialization time $t = 0$) due to the implicit nonlinear components. Moreover, using linear information to initialize the HOS nonlinear propagation would lead to unstable calculations of higher-order terms (Dommermuth, 2000). Hence, to allow for a deterministic comparison between models and as explained in paragraph (*b*), the HOS was initialized with a nonlinear irregular wave field, here, a second-order Stokes wave solution, following the method detailed by Perignon *et al.* (2010) and Perignon (2011).

II.3.3.2 Propagation Results

Figure II.11 shows the comparison of normalized surface elevations computed in the different models at $t = 4T_p$, for the two types of spectra. For the wave train extracted from a Gaussian spectrum, figure II.11a shows that the ICWM model prediction remains very accurate after $4T_p$. With the other models, for which wave celerity is inaccurate, waves are either delayed (linear solution) or too fast (CWM2), with respect to the HOS reference solution. Similar observations can be made in figure II.11b, although in a less obvious way, for waves extracted from a JONSWAP spectrum: after $4T_p$ of propagation, the overall locations of dominant waves modeled with

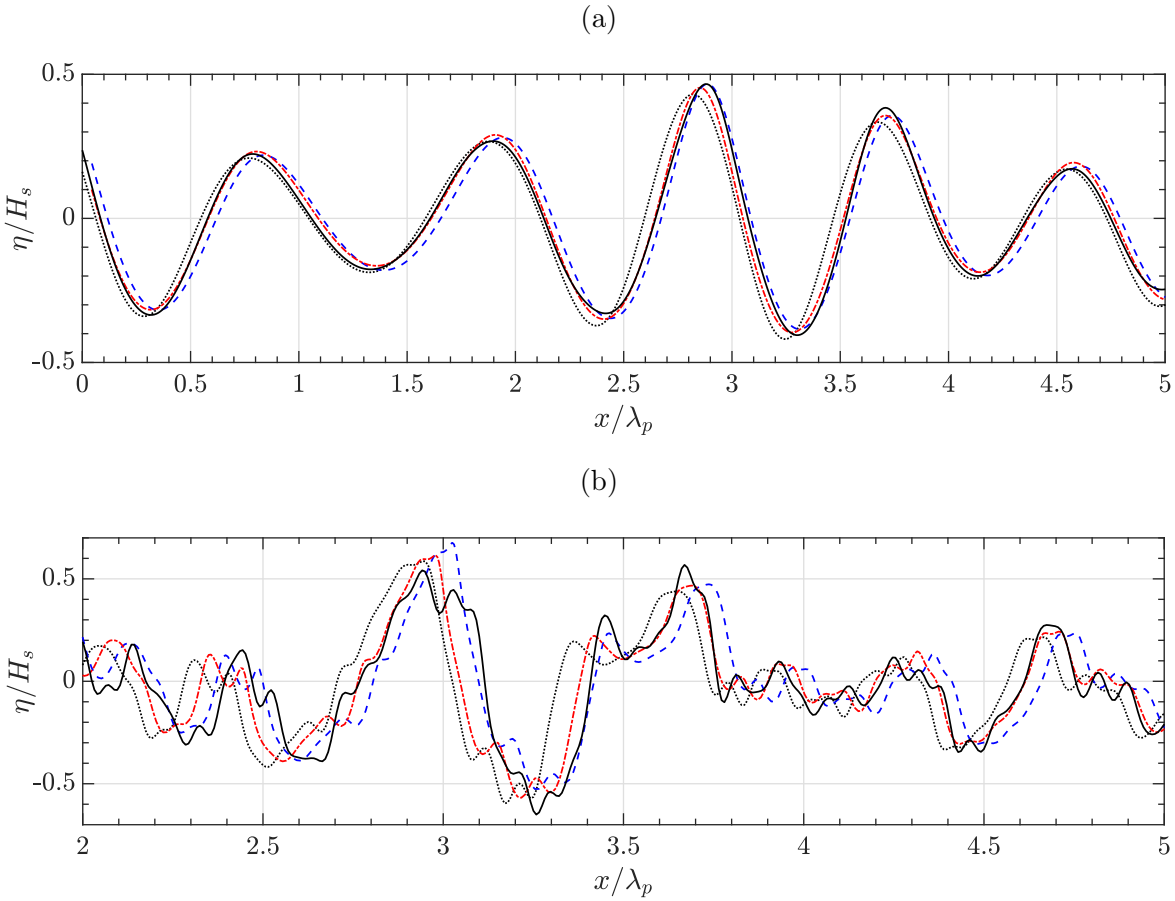

Figure II.11: Normalized surface elevations computed with: HOS (——), linear (·····), CWM2 (- - - -), and ICWM (-·-·-) models, after $t = 4T_p$ of propagation, for irregular wave trains with $T_p = 10$ s extracted from a: (a) Gaussian spectrum with standard deviation $\sigma/\omega_p \approx 0.08$, $H_s = 9$ m and characteristic steepness $H_s/\lambda_p \approx 6\%$, and (b) a JONSWAP spectrum with peakedness parameter $\gamma = 1$ (similar to a PM spectrum), $H_s = 6$ m, and $H_s/\lambda_p \approx 4\%$. See figure II.10 for spectral shapes.

ICWM are consistent with the HOS solution (although not their smaller fluctuations), whereas the linear and CWM2 solutions are already desynchronized with it. This remark is in agreement with results from section II.2.3 showing that ICWM accurately predicts the nonlinear correction of wave phase velocities of energetic waves, that is in a similar way to the Eulerian third-order expansion solution.

More specifically, Figure II.12 shows the time evolution of relative errors computed with equation (II.21) over the first 10 peak periods of propagation, for the two cases of figure II.11. In the Gaussian spectrum case, the ICWM model yields errors reduced by a factor of ~ 2 with respect to the linear solution as well as a significant improvement with respect to results of the CWM2 model. The overall error is larger in the JONSWAP spectrum case, but the same qualitative observations hold true. To quantify the effect of second-order interaction terms, which have been discarded in the ICWM model to obtain a numerically efficient solution, the CWM2 model results with a corrected dispersion relationship (II.20) have also been computed (referred to as "ICWM with interaction terms" in figure II.12). As seen in the figures, these terms only yield significant effects within the first peak period of propagation (particularly in the Gaussian case) and become rapidly negligible with respect to the correction due to the dispersion relationship (ICWM). The black lines in figure II.12 show the relative variation of the wave energy spectrum of the HOS reference surface, which is verified to remain small with respect to the relative er-

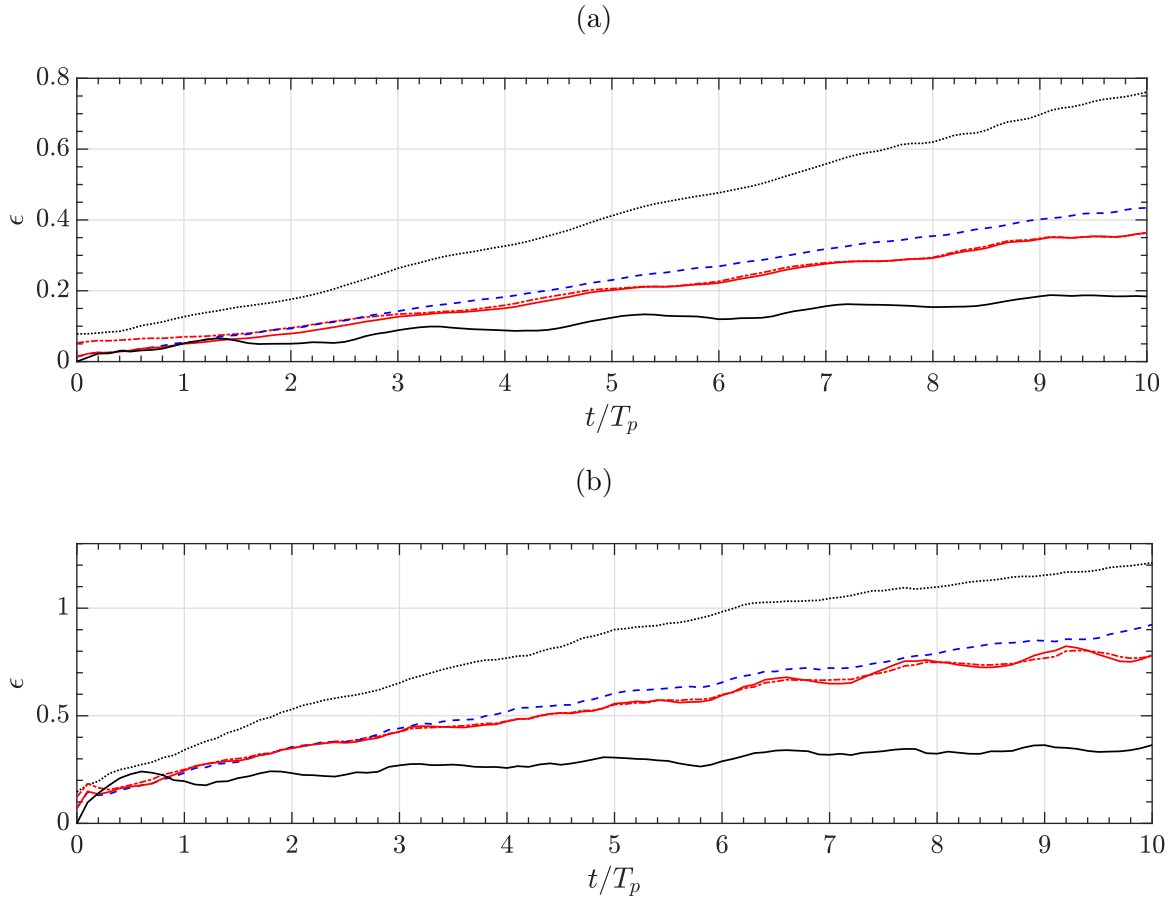

Figure II.12: Time evolution of relative errors (equation (II.21)) for the two cases of figure II.11 (a and b), for the linear (·····), CWM2 (- - - -), ICWM (-·-·-) and ICWM with interaction terms (——) model results. The black solid line shows the relative change (equation (II.22)) of the prescribed spectrum during the HOS time evolution.

ror between methods. This confirms that the improved accuracy of ICWM for wave prediction is essentially due to a corrected wave celerity rather than an accurate description of bound waves.

II.3.3.3 Influence of Surface Shape and Wave Field Steepness

So far, numerical results are shown in figure II.12 for only one single steepness for each spectrum (6% or 4%), which in fact represents the largest value achievable with the HOS model. At larger steepnesses, this model breaks down after a few periods due to the occurrence of wave breaking in the simulated sea-state, which is not explicitly treated in the version of HOS-ocean used in this work. In this section, we perform a more systematic numerical investigation of the effect of wave steepness for the Gaussian spectrum. Also, in order to investigate the relative importance of the horizontal displacement term D_1 with respect to the wave celerity correction, we consider a model following LWT but using the corrected nonlinear dispersion relationship, similar to that of ICWM, and denoted LWT-CDR (Linear Wave Theory with Corrected Dispersion Relation) hereafter.

Hence, in the same way as in the previous section, the time evolution of the RMS error (II.21) is quantified and presented in figure II.13 for the Gaussian spectrum and for three characteristic steepnesses of 2%, 4% and 6%. We see that, at initial time, when it represents a purely linear surface, LWT-CDR yields a significantly larger relative error than ICWM. However, after a few peak periods of propagation, we find that the LWT-CDR and ICWM solutions yield comparable

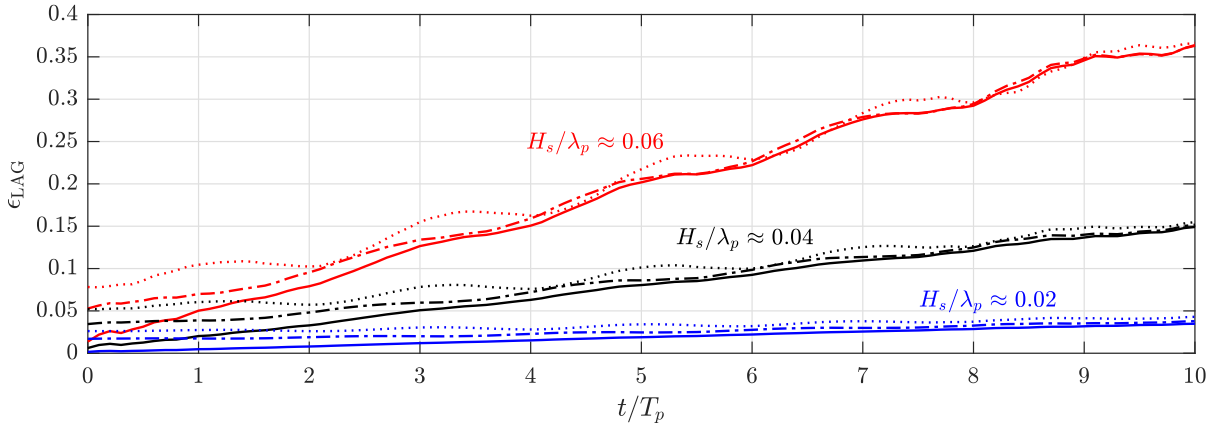

Figure II.13: Time evolution of relative errors (equation (II.21)) for LWT-CDR (·····), ICWM (-·-·-·) and ICWM with interaction terms (——), for the described Gaussian spectrum with three characteristic steepnesses $H_s/\lambda_p \approx 6\%$, 4% and 2%.

errors, showing once again that, as far as predicting surface elevations, the dominant source of error is inaccuracies in wave celerity. As could be expected, increasing the characteristic steepness both increases discrepancies between the linear and nonlinear model results and yields larger overall errors. For all steepnesses, omitting the interaction terms yield lower divergences than using a linear surface shape, and their effects become more rapidly negligible with the propagation time: for the largest steepness, the error of ICWM needs $\sim$ 3–4T_p to reach approximately the same level as ICWM with interaction terms, while LWT-CDR needs $\sim$ 8–9T_p. In contrast, the error corresponding to the standard LWT diverges from the one of ICWM, as shown in figure II.12a.

II.4 Dynamical Aspects

We saw in the previous sections that the ICWM was able to improve the free surface kinematics representation. However, in order to calculate the motion of a structure induced by the predicted waves, the corresponding pressure field has to be solved. We propose in the next section to investigate the properties of the Lagrangian pressure field, then to formulate its improved formulation in a similar way to that of the free surface kinematics and see how it compares to the exact Lagrangian formulation.

II.4.1 Lagrangian Dynamic Pressure Field

We first describe the Lagrangian dynamic pressure field, using the following methodology. We recall that the Lagrangian solution of the fluid particle dynamics is of the form

$$\begin{cases} \boldsymbol{r}\left(\boldsymbol{\xi},\delta,t\right) = \boldsymbol{\xi} + \sum\limits_i \boldsymbol{D}_i\left(\boldsymbol{\xi},\delta,t\right), \\ z\left(\boldsymbol{\xi},\delta,t\right) = \delta + \sum\limits_i Z_i\left(\boldsymbol{\xi},\delta,t\right), \\ p\left(\boldsymbol{\xi},\delta,t\right) = p_{\mathrm{a}} - \rho g\delta + \sum\limits_i P_i\left(\boldsymbol{\xi},\delta,t\right), \end{cases}$$

where $\boldsymbol{r} = (x, y)$, $\boldsymbol{\xi} = (\alpha, \beta)$, $\boldsymbol{D}_i$ and Z_i are respectively the horizontal and vertical particle relative displacements with respect to the particle position at rest $\boldsymbol{\xi}$, and P_i is the pressure field associated with the i^{th}-order of expansion. According to this formalism, we formulate the

dynamic pressure field at order i as

$$p_{\mathrm{d}i} = p - p_{\mathrm{a}} + \rho g z = P_i\left(\boldsymbol{\xi}, \delta, t\right) + \rho g Z_i\left(\boldsymbol{\xi}, \delta, t\right).$$

Note that this formulation of the dynamic pressure includes both the Bernoulli's term (proportional to the squared velocity) and a non-stationary term.

II.4.1.1 First-Order Solution

The Lagrangian first-order solution is

$$\begin{cases} \boldsymbol{D}_1\left(\boldsymbol{\xi}, \delta, t\right) = -A\hat{\boldsymbol{k}}\sin\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega t\right)e^{k\delta}, \\ Z_1\left(\boldsymbol{\xi}, \delta, t\right) = A\cos\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega t\right)e^{k\delta}, \\ P_1\left(\boldsymbol{\xi}, \delta, t\right) = 0, \end{cases}$$

which leads to the dynamic pressure field (see figure II.14a)

$$p_{\mathrm{d}1} = \rho g A\cos\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega t\right)e^{k\delta}.$$

We recognize here a sinusoidal variation that corresponds to the pressure induced by the first-order term in the vertical displacement of the fluid particle. Note that this term is equivalent to the dynamic pressure field derived from the LWT (i.e., term $-\rho\partial\phi/\partial t$). However, we did not have to linearize the free-surface boundary conditions about $z = 0$ to get this result. The pressure we determined is attached to a particle at the location $\left(\boldsymbol{r}\left(\boldsymbol{\xi}, \delta, t\right), z\left(\boldsymbol{\xi}, \delta, t\right)\right)$, and it ensures that the pressure is zero on the free-surface, and not on $z = 0$ as according to the LWT. The first-order Lagrangian pressure field is thus more accurate than the Eulerian result at the same order. This

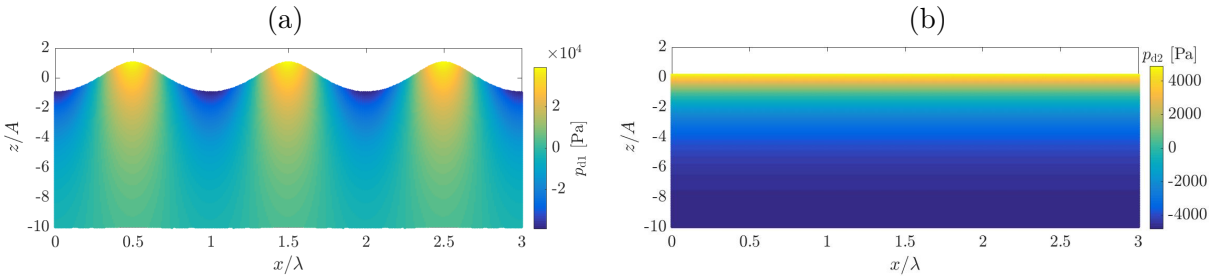

Figure II.14: (a) First- and (b) second-order Lagrangian dynamic pressure field for a wave of steepness $kA = 0.25$ at $t = 0$. In both cases, the particles relative displacement with respect to their reference location at the same order (i.e., first (a) or second (b)) is also plotted.

feature allows to calculate quantities (pressure, orbital velocities) above $z = 0$ without requiring stretching or extrapolation such as Wheeler's method (Wheeler *et al.*, 1970) or delta-stretching method (Rodenbusch & Forristall, 1986), which do not comply with the governing hydrodynamic equations, limiting their reliability. As we will see later, this is true only to a limited extent, which depends on the wave steepness.

II.4.1.2 Second-Order Solution

The Lagrangian second-order solution is

$$\begin{cases} \boldsymbol{D}_2\left(\boldsymbol{\xi}, \delta, t\right) = A^2\omega\boldsymbol{k}te^{2k\delta}, \\ Z_2\left(\boldsymbol{\xi}, \delta, t\right) = \dfrac{1}{2}A^2ke^{2k\delta}, \\ P_2\left(\boldsymbol{\xi}, \delta, t\right) = \rho g\dfrac{1}{2}A^2k\left(e^{2k\delta} - 1\right), \end{cases} \tag{II.23}$$

which leads to the dynamic pressure field (see figure II.14b)

$$p_{\mathrm{d}2} = \rho g \frac{1}{2} A^2 k \left(2e^{2k\delta} - 1\right).$$

The term corresponds to the mean surface level correction that induced a pressure modification that only depends on the particle reference position δ (and not on $\boldsymbol{\xi}$ nor t). The total dynamic pressure field up to the second-order is then (see figure II.15a)

$$p_{\mathrm{d}} = p_{\mathrm{d}1} + p_{\mathrm{d}2} = \rho g A \cos\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega t\right) e^{k\delta} + \rho g \frac{1}{2} A^2 k \left(2e^{2k\delta} - 1\right).$$

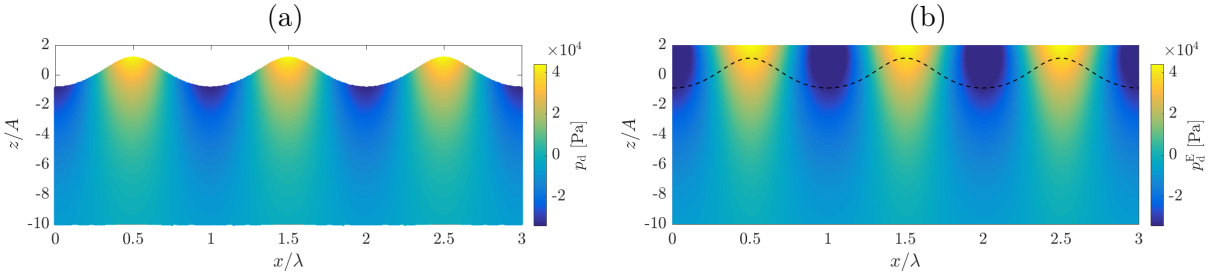

Figure II.15: (a) Lagrangian dynamic pressure field up to the second-order and (b) its approximate in the Eulerian framework for a wave of steepness $kA = 0.25$ at $t = 0$. The dashed line on (b) corresponds to the free surface elevation.

II.4.2 Approximate in the Eulerian Framework

The dynamic pressure field derived in the previous section $p_{\mathrm{d}}\left(\boldsymbol{\xi}, \delta, t\right)$ is related to the particle that is located at coordinates $\left(\boldsymbol{r}\left(\boldsymbol{\xi}, \delta, t\right), z\left(\boldsymbol{\xi}, \delta, t\right)\right)$. Both the pressure and the particle displacements depend on the reference particle locations, but are not explicitly related together. Similar to section II.2.2 in which we look for an approximate Eulerian expression of the surface elevation, we want here to approximate a solution that allows us to determine the pressure field at any location $(\boldsymbol{r}, z)$. First, we express the Lagrangian displacements in a way that the particle shift is implicitly taken into account, following

$$\begin{cases} \boldsymbol{\mathcal{R}}\left(\boldsymbol{\xi}, \delta, t\right) = \boldsymbol{r}\left(\boldsymbol{\xi} - A^2\omega\boldsymbol{k}te^{2k\delta}, \delta, t\right) = \boldsymbol{\xi} + \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}, \delta, t\right), \\ \mathcal{Z}\left(\boldsymbol{\xi}, \delta, t\right) = z\left(\boldsymbol{\xi} - A^2\omega\boldsymbol{k}te^{2k\delta}, \delta, t\right) = \delta + \mathcal{V}\left(\boldsymbol{\xi}, \delta, t\right), \\ \mathcal{P}\left(\boldsymbol{\xi}, \delta, t\right) = p\left(\boldsymbol{\xi} - A^2\omega\boldsymbol{k}te^{2k\delta}, \delta, t\right) = p_{\mathrm{a}} - \rho g\mathcal{Z} + \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{\xi}, \delta, t\right), \end{cases}$$

where

$$\begin{cases} \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}, \delta, t\right) = -A\hat{\boldsymbol{k}}\sin\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega_{\mathrm{L}}t\right) e^{k\delta}, \\ \mathcal{V}\left(\boldsymbol{\xi}, \delta, t\right) = A\cos\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega_{\mathrm{L}}t\right) e^{k\delta} + \dfrac{1}{2}A^2ke^{2k\delta}, \\ \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{\xi}, \delta, t\right) = \rho g A\cos\left(\boldsymbol{k}\cdot\boldsymbol{\xi} - \omega_{\mathrm{L}}t\right) e^{k\delta} + \rho g\dfrac{1}{2}A^2k\left(2e^{2k\delta} - 1\right), \end{cases}$$

and

$$\omega_{\mathrm{L}} = \omega + \boldsymbol{k}\cdot A^2\omega\boldsymbol{k}e^{2k\delta} = \omega\left(1 + A^2k^2e^{2k\delta}\right).$$

The reformulation of the Lagrangian coordinates in terms of a modified angular velocity ω_{L} allows us to go through the following approximation

$$\mathcal{P}\left(\boldsymbol{\xi}, \delta\right) = \mathcal{P}\left(\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi}, \delta\right), \mathcal{Z} - \mathcal{V}\left(\boldsymbol{\xi}, \delta\right)\right) \approx \mathcal{P}\left(\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{\mathcal{R}}, \mathcal{Z}\right), \mathcal{Z} - \mathcal{V}\left(\boldsymbol{\mathcal{R}}, \mathcal{Z}\right)\right) = p^{\mathrm{E}}\left(\boldsymbol{\mathcal{R}}, \mathcal{Z}\right), \tag{II.24}$$

which leads to an explicit formulation of the pressure field in the Eulerian framework

$$p^{\mathrm{E}}(\boldsymbol{r},z,t) = p_{\mathrm{a}} - \rho g z + p_{\mathrm{d}}^{\mathrm{E}}(\boldsymbol{r},z,t),$$

where the Eulerian dynamic pressure (see figure II.15b) is

$$\begin{aligned} p_{\mathrm{d}}^{\mathrm{E}}(\boldsymbol{r},z,t) &= \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{r}-\boldsymbol{\mathcal{D}}(\boldsymbol{r},z,t), z-\mathcal{V}(\boldsymbol{r},z,t), t\right) \\ &= \rho g A \cos\left[\boldsymbol{k}\cdot\left(\boldsymbol{r}-\boldsymbol{\mathcal{D}}(\boldsymbol{r},z,t)\right)-\omega_{\mathrm{L}}t\right] e^{k(z-\mathcal{V}(\boldsymbol{r},z,t))} + \rho g \frac{1}{2}A^2 k\left[2e^{2k(z-\mathcal{V}(\boldsymbol{r},z,t))}-1\right]. \end{aligned}$$

The Lagrangian and Eulerian total pressure fields are plotted on figure II.16. A way to directly assess the accuracy of the Eulerian approximation is to compare the surface elevation and the contour line corresponding to $p = p_{\mathrm{a}}$. According to the Lagrangian equations, the pressure should be equal to the atmospheric pressure on the free surface (here, $p_{\mathrm{a}} = 0$). Figure II.16b shows that the surface elevation and the pressure contour line are very close. Only a small difference is found on the wave crest.

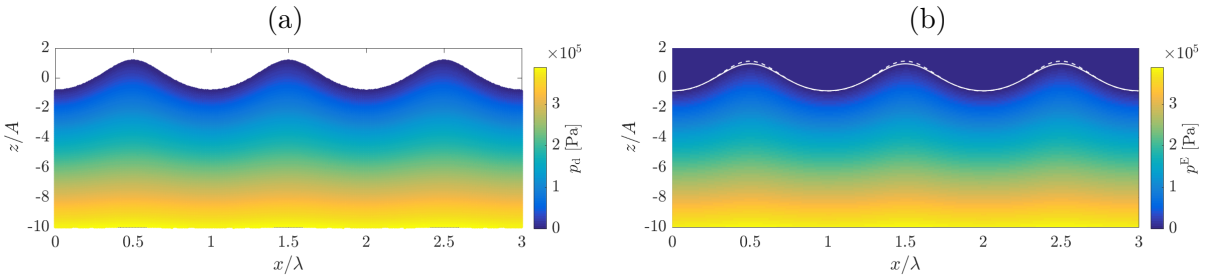

Figure II.16: (a) Lagrangian pressure field up to the second-order and (b) its approximate in the Eulerian framework for a wave of steepness $kA = 0.25$ at $t = 0$. The dashed line on (b) corresponds to the surface elevation, while the continuous line is the pressure contour line $p = p_{\mathrm{a}} = 0$.

II.4.3 Time Propagation of the Pressure Field

As explained by Gjosund (2000, section 4.3), the second-order Lagrangian solution leads to an inconsistency when propagated in time due to the second-order secular term $\boldsymbol{D}_2$ in equation (II.23). This term being a vertically non-uniform current (i.e., Stokes drift exponentially decreasing with depth), a time propagation will lead to a deformation of the initial distribution of particle locations which will eventually break continuity. From a practical point of view, however, a consistent Lagrangian solution can be formulated at every time based on the solution at $t = 0$, that is without time-propagation. The key point is to define a solution that is 'equivalent' (in a sense that there is no particle shift difference along the water column) to the initial solution (i.e., $t = 0$) at every time by considering that the propagated solution is simply a 'phase shift' of wave components, rather than a time propagation of the solution. This solution is the same as the one derived above with a δ-independent phase shift (with $\delta = 0$, i.e., the phase shift of the free surface particles). This simply consists in removing the δ-dependence in the expression of ω_{L}, that is $\omega_{\mathrm{L}} = \omega\left(1 + A^2k^2\right)$. Using this method (which also works for irregular waves), results of the pressure field with a wave steepness $kA = 0.25$ after a propagation of four wave periods $T = 2\pi/\omega$ are shown in figures II.17 and II.18. Consistent results with the case of $t = 0$ are found.

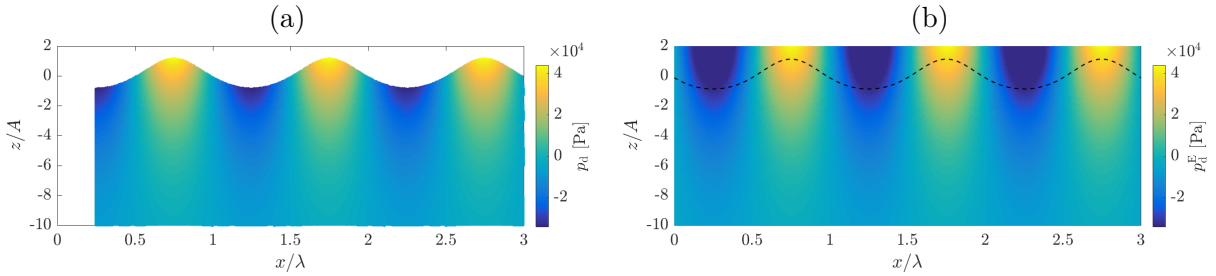

Figure II.17: (a) Lagrangian dynamic pressure field up to the second-order and (b) its approximate in the Eulerian framework for a wave of steepness $kA = 0.25$ at $t = 4T$. The dashed line on (b) corresponds to the surface elevation.

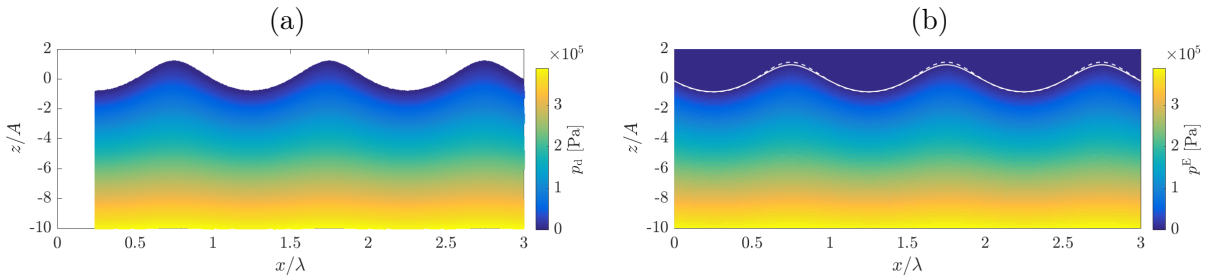

Figure II.18: (a) Lagrangian pressure field up to the second-order and (b) its approximate in the Eulerian framework for a wave of steepness $kA = 0.25$ at $t = 4T$. The dashed line on (b) corresponds to the surface elevation, while the continuous line is the pressure contour line $p = p_{\mathrm{a}} = 0$.

II.4.4 Improved Choppy Irregular Waves

The ICWM Lagrangian components yield

$$\begin{cases} \boldsymbol{r}\left(\boldsymbol{\xi},\delta,t\right) = \boldsymbol{\xi} + \displaystyle\sum_{n=1}^{N} \hat{\boldsymbol{k}}_n \left(-a_n \sin\tilde{\psi}_n + b_n \cos\tilde{\psi}_n\right) e^{k_n\delta} + \boldsymbol{\mathcal{U}}_s t, \\ z\left(\boldsymbol{\xi},\delta,t\right) = \delta + \displaystyle\sum_{n=1}^{N} \left[\left(a_n \cos\tilde{\psi}_n + b_n \sin\tilde{\psi}_n\right) e^{k_n\delta} + \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n e^{2k_n\delta}\right], \\ p\left(\boldsymbol{\xi},\delta,t\right) = p_{\mathrm{a}} - \rho g z + p_{\mathrm{d}}, \end{cases}$$

where $\tilde{\psi}_n = \boldsymbol{k}_n\cdot\boldsymbol{\xi} - \tilde{\omega}_n t$, $\tilde{\omega}_n = \omega_n - 1/2\boldsymbol{k}_n\cdot\boldsymbol{\mathcal{U}}_s$, $\boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N}\left(a_n^2 + b_n^2\right)\omega_n\boldsymbol{k}_n e^{2k_n\delta}$, and the dynamic pressure field is

$$p_{\mathrm{d}} = \rho g \sum_{n=1}^{N}\left[\left(a_n\cos\tilde{\psi}_n + b_n\sin\tilde{\psi}_n\right)e^{k_n\delta} + \frac{1}{2}\left(a_n^2+b_n^2\right)k_n\left(2e^{2k_n\delta}-1\right)\right].$$

The Eulerian approximate of the ICWM Lagrangian results follows from the same procedure as in the previous section. We first implicitly incorporate the particle shift in a modified angular velocity as

$$\begin{cases} \boldsymbol{\mathcal{R}}\left(\boldsymbol{\xi},\delta,t\right) = \boldsymbol{r}\left(\boldsymbol{\xi} - \boldsymbol{\mathcal{U}}_s t,\delta,t\right) = \boldsymbol{\xi} + \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi},\delta,t\right), \\ \mathcal{Z}\left(\boldsymbol{\xi},\delta,t\right) = z\left(\boldsymbol{\xi} - \boldsymbol{\mathcal{U}}_s t,\delta,t\right) = \delta + \mathcal{V}\left(\boldsymbol{\xi},\delta,t\right), \\ \mathcal{P}\left(\boldsymbol{\xi},\delta,t\right) = p\left(\boldsymbol{\xi} - \boldsymbol{\mathcal{U}}_s t,\delta,t\right) = p_{\mathrm{a}} - \rho g\mathcal{Z} + \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{\xi},\delta,t\right), \end{cases}$$

where

$$\begin{cases} \boldsymbol{\mathcal{D}}\left(\boldsymbol{\xi},\delta,t\right) = \displaystyle\sum_{n=1}^{N} \hat{\boldsymbol{k}}_n \left(-a_n \sin\psi_{\mathrm{L}n} + b_n \cos\psi_{\mathrm{L}n}\right) e^{k_n\delta}, \\ \mathcal{V}\left(\boldsymbol{\xi},\delta,t\right) = \displaystyle\sum_{n=1}^{N} \left[\left(a_n \cos\psi_{\mathrm{L}n} + b_n \sin\psi_{\mathrm{L}n}\right) e^{k_n\delta} + \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n e^{2k_n\delta}\right], \\ \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{\xi},\delta,t\right) = \rho g \displaystyle\sum_{n=1}^{N} \left[\left(a_n \cos\psi_{\mathrm{L}n} + b_n \sin\psi_{\mathrm{L}n}\right) e^{k_n\delta} + \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n \left(2e^{2k_n\delta} - 1\right)\right], \end{cases}$$

and $\psi_{\mathrm{L}n} = \boldsymbol{k}_n \cdot \boldsymbol{\xi} - \omega_{\mathrm{L}n} t$ and $\omega_{\mathrm{L}n} = \omega_n + 1/2\boldsymbol{k}_n \cdot \boldsymbol{\mathcal{U}}_s$. Then, applying approximation (II.24) to our ICWM equations, the Eulerian dynamic pressure follows

$$\begin{aligned} p_{\mathrm{d}}^{\mathrm{E}}\left(\boldsymbol{r},z,t\right) &= \mathcal{P}_{\mathrm{d}}\left(\boldsymbol{r} - \boldsymbol{\mathcal{D}}\left(\boldsymbol{r},z,t\right), z - \mathcal{V}\left(\boldsymbol{r},z,t\right), t\right) \\ &= \rho g \sum_{n=1}^{N} \left\{\left(a_n \cos\Psi_n + b_n \sin\Psi_n\right) e^{k_n\left(z-\mathcal{V}(\boldsymbol{r},z,t)\right)} + \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n \left[2e^{2k_n\left(z-\mathcal{V}(\boldsymbol{r},z,t)\right)} - 1\right]\right\}, \end{aligned} \tag{II.25}$$

in which $\Psi_n = \boldsymbol{k}_n \cdot \left[\boldsymbol{r} - \displaystyle\sum_{n=1}^{N} \hat{\boldsymbol{k}}_n \left(-a_n \sin\psi_{\mathrm{L}n} + b_n \cos\psi_{\mathrm{L}n}\right) e^{k_n z}\right] - \omega_{\mathrm{L}n} t.$

Similar to section II.4.3, the Lagrangian inconsistency is solved by removing the δ-dependency of the Stokes drift according to $\boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N}\left(a_n^2 + b_n^2\right)\omega_n \boldsymbol{k}_n$.

II.4.4.1 Results for a Bichromatic Wave Field

A first polychromatic configuration is investigated through bichromatic waves. Two waves are chosen: a short wave of amplitude $A_2 = 0.75$ m and period $T_2 = 5$ s is riding a longer wave of amplitude $A_1 = 3$ m and period $T_1 = 10$ s. Results are presented in figures II.19 and II.20. Even if small differences between the contour line $p = p_a$ of the Eulerian total pressure field and the free surface are visible about the highest crest and lowest trough (figure II.20b), both Lagrangian and Eulerian solutions agree very well.

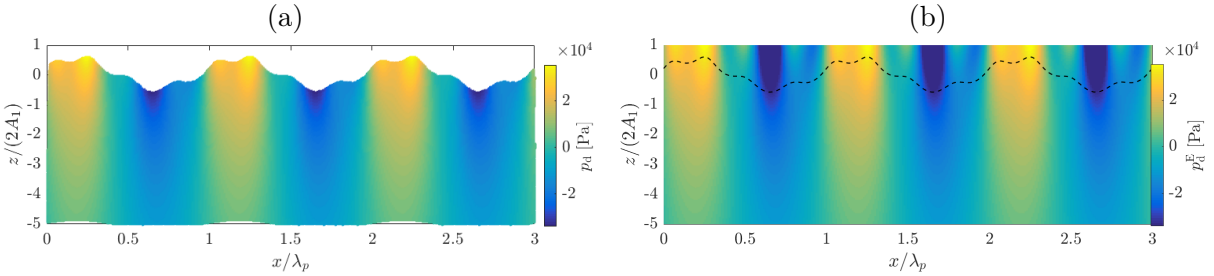

Figure II.19: (a) Lagrangian dynamic pressure field up to the second-order and (b) its approximate in the Eulerian framework for a bichromatic wave field at $t = 0$. The dashed line on (b) corresponds to the surface elevation. ($A_1 = 3$ m, $A_2 = 0.75$ m, $T_1 = 10$ s, $T_2 = 5$ s)

II.4.4.2 Results for a JONSWAP Spectrum

Now, we look at the results obtained for an irregular wave field prescribed by a JONSWAP spectrum (see equation (I.2) for detailed formulation) of peak period $T_p = 10$ s and peak enhancement factor $\gamma = 3.3$. An example of a comparison of total pressure fields is shown in figure II.21 for a wave field of small significant wave height $H_s = 1$ m ($H_s/\lambda_p \approx 0.64\%$). An excellent agreement is found by looking at the correspondence between the Eulerian contour and the Lagrangian free surface.

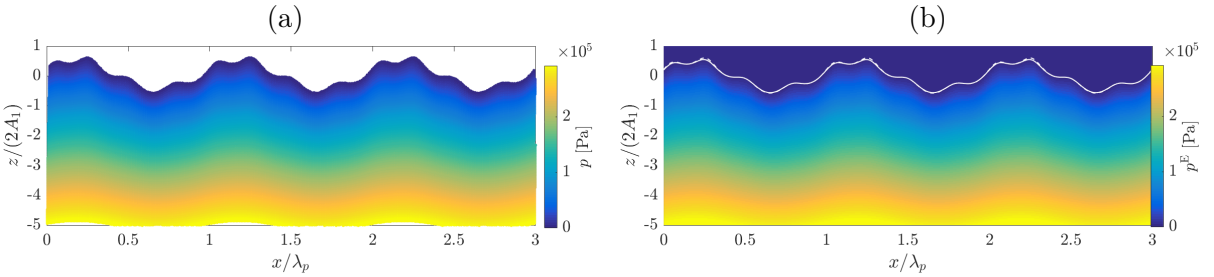

Figure II.20: (a) Lagrangian pressure field up to the second-order and (b) its approximate in the Eulerian framework for a bichromatic wave field at $t = 0$. The dashed line on (b) corresponds to the surface elevation, while the continuous line is the pressure contour line $p = p_a = 0$. ($A_1 = 3$ m, $A_2 = 0.75$ m, $T_1 = 10$ s, $T_2 = 5$ s)

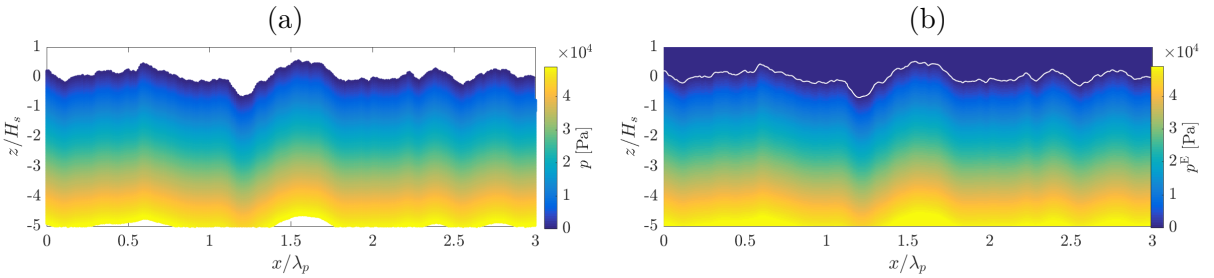

Figure II.21: (a) Lagrangian pressure field up to the second-order and (b) its approximate in the Eulerian framework for a JONSWAP spectrum ($H_s = 1$ m, $T_p = 10$ s, $\gamma = 3.3$) at $t = 0$. The dashed line on (b) corresponds to the surface elevation, while the continuous line is the pressure contour line $p = p_a = 0$.

However, for larger steepness, discrepancies arise in the Eulerian solution, especially near high wave crests, as it can be seen in figure II.22a that shows the Eulerian total pressure for a sea state of significant wave height $H_s = 5$ m ($H_s/\lambda_p \approx 3.2\%$). It looks like from a certain value $z^{\text{lim}} > 0$, the pressure drops to zero. This effect can be reduced by using an iterative scheme, such as describe by Gjosund (2003) and described by equation (II.19), to reduce the discrepancy induced by the approximation (II.24). The converged Eulerian pressure is depicted in figure II.22b. We see that one major correction in the converged solution is the non-null

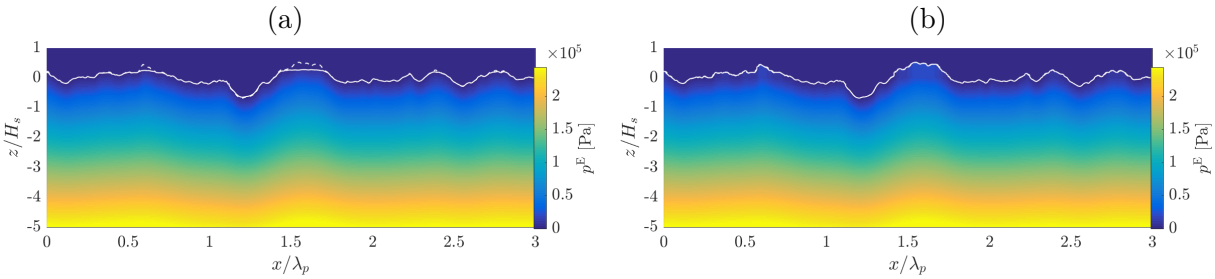

Figure II.22: Eulerian approximate of the second-order Lagrangian pressure field for a JONSWAP spectrum ($H_s = 5$ m, $T_p = 10$ s, $\gamma = 3.3$) at $t = 0$ using (a) equation (II.25) and (b) a convergence scheme. The dashed lines on correspond to the surface elevation, while the continuous line is the pressure contour line $p = p_a = 0$.

pressure above z^{lim}. A closer look at the region near the free surface elevation is shown in figure II.23. The relative improvement of the converged solution is noticeable even below z^{lim}, and we note that the value of the pressure above z^{lim} for the converged solution seems to correspond to a prolongation of the pressure value at z^{lim}.

In order to be able to use this solution as input for a structure response motion solver, further investigation should be done on the influence of such an inaccurate value of the pressure on the

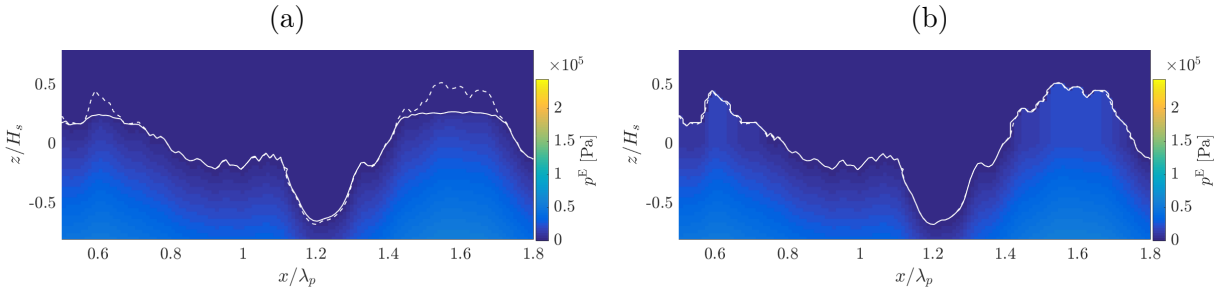

Figure II.23: Similar to figure II.22, with a closer look at the free surface elevation.

structural response calculations. Also, note that a similar formulation of Eulerian estimate of orbital velocities can be formulated for the same purpose of implementation in a wave-structure interaction solver.

Conclusion

In this chapter, we detailed the Lagrangian approach for ocean gravity waves modeling and highlighted its advantages in terms of nonlinear features over its Eulerian counterpart. The properties of a second-order solution for surface elevation kinematics are described, and we assessed the numerical performances of an improved formulation, ICWM, that corrects the nonlinear dispersion relation. In view of implementation in a wave prediction algorithm that is fed with geo-referenced wave measurements, Eulerian approximate solutions of ICWM surface elevation and pressure are derived and their accuracy with respect to the initial Lagrangian solution are investigated.

It was found that the nonlinear correction of the phase velocity is crucial for accurate wave representation after several periods of propagation. The improved formulation allows retrieving the third-order Eulerian solution for regular waves. In the case of irregular waves, it leads to an appropriate correction for energetic wave components but misses the modulation effects of shorter waves. A simple Eulerian approximate that does not necessitate complex mathematical developments nor iterative procedures can be formulated and leads to consistent results. The pressure estimate, however, can lead to spurious values (decrease to zero) near the wave crests for large steepness if no numerical precautions are taken.

Résumé du troisième chapitre

Nous nous concentrons dans le troisième chapitre sur la procédure d'assimilation de données, consistant à accéder aux informations pertinentes contenues dans les mesures de vagues (appelées observations) afin de définir les conditions initiales du modèle physique servant à propager le champ de vagues en vue de sa prédiction. La méthode d'assimilation développée ici se base sur des mesures optiques telles que générées par une caméra lidar. La distribution de telles mesures est d'abord caractérisée au travers d'un approche géométrique, mettant en avant leur forte non-uniformité spatiale due à l'incidence rasante des rayons du capteur à la surface. La reconstruction du champ de vagues (c'est-à-dire l'inversion du modèle physique) est formulée pour le modèle non-linéaire ICWM étudié dans le chapitre précédent, puis nous détaillons le processus itératif de résolution. La formulation analytique de ICWM permet d'écrire le problème inverse sous une forme matricielle explicite, et la solution est trouvée en passant par une procédure de régularisation Tikhonov et une décomposition en valeurs singulières. Une fois le champ de vagues reconstruit, la région spatio-temporelle dans laquelle les informations assimilées restent valides pendant la propagation du modèle de vagues est bornée. À partir de l'étendue et de l'échantillonnage des observations, nous décrivons l'évolution de la zone de prédiction théoriquement accessible. La distribution spatio-temporelle d'une erreur théorique de prédiction est décrite à partir de la quantité d'énergie tronquée lors du processus de mesure, et montre que la taille de la zone de prédiction est fortement dépendante de la précision attendue de la prédiction. Des critères sont ensuite présentés pour la sélection des fréquences et directions de coupure pour le champ de vagues modélisé ainsi que pour l'évolution de la zone de prédiction, nous permettant de fermer le problème inverse et définir clairement une région à l'intérieur de laquelle les performances de prédiction de notre méthode seront évaluées. Enfin, les questionnements relatifs à la détermination de la zone optimale de mesure (menant à l'horizon de prédiction souhaité) sont exposés et en partie traités. Depuis la caractérisation des mesures optiques d'élévation de surface jusqu'à la formulation du problème d'inversion de ICWM et la description de la zone de prédiction théoriquement accessible, ce chapitre termine l'explication des bases théoriques de la méthode de prédiction employée dans cette thèse.

Chapter III

Optical Data Assimilation and Accessible Prediction Region

Introduction

This chapter focuses on the data assimilation process, which consists in getting access to relevant information from the wave measurements (called observations) to properly define the initial conditions of our wave model. As we saw in the first chapter, many methods can be employed to invert wave models, and the methods have to be adapted to both the wave model and the nature of the wave measurements. In the following, the developed assimilation method is based on typical optical measurements such as the one that would be made by a lidar camera observing the ocean surface, namely non-uniformly distributed free surface elevation observations. Also, as our wave model follows an analytical formulation, an explicit inversion system is written and solved. Then, once the wave field is reconstructed, the spatio-temporal zone within which the assimilated information remains valid during the model propagation is bounded. Based on the extent and sampling characteristics of the observations, we detail the evolution of this accessible prediction zone.

We will thus first focus on the characterization of the optical free surface elevation measurements in terms of non-uniformity. Then, the wave field reconstruction (i.e., wave model inversion) problem is formulated for the nonlinear wave model ICWM studied in the previous chapter. Last, the formulation of the theoretically accessible prediction zone is derived, and we will see how it can be used for the determination of the appropriate measurement zone leading to the desired prediction horizon, which is typically on the order of a few wave periods.

III.1 Properties of Optical Measurements

In this work, following the potential of optical sensors to provide accurate free surface elevation measurements (see section I.2.2.2), synthetic lidar measurements are used as observations. The produced observations have distribution properties that we will explicit in this section based on a realistic setup: the sensor is mounted on top of a marine structure, observing the ocean surface at a distance with a grazing angle. After detailing the setup parameters, the resulting spatial sampling properties are described.

III.1.1 Setup Description and Parameters Definitions

We first denote the position and viewing parameters of a lidar camera. On a similar coordinate system to the one described in the previous section for the definition of the wave model, we consider that it is located at a point of coordinates $(x_c(t), y_c(t), z_c(t))$ – the time dependence is relevant in case the structure upon which the camera is mounted is moving (e.g., a ship, a floating wind turbine). The camera is facing the ocean surface with viewing angles $\alpha(t)$ and

$\beta\,(t)$ in the vertical and horizontal directions, respectively. Associated aperture angles $\alpha_a\,(t)$ and $\beta_a\,(t)$ define the extent of the illuminated area, that is the observation zone (or measurement zone). All these parameters and the resulting observation zone are depicted in figure III.1.

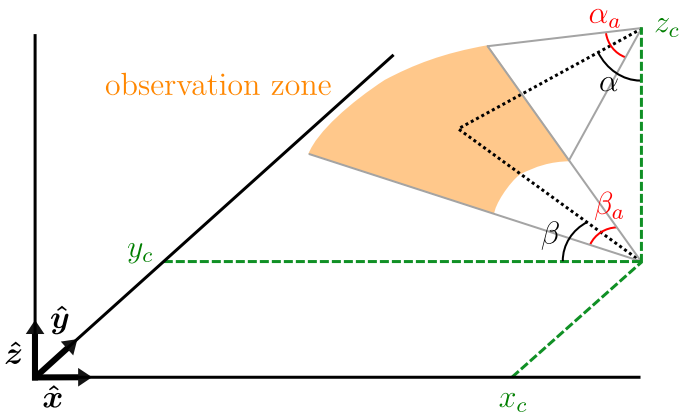

Figure III.1: Geometrical parameters leading to the observation zone.

The described setup leads to data distribution similar to the one originated from a rotating lidar (see figures I.4a, b). This technology is the one employed by Kabel *et al.* (2019) to generate two-dimensional maps of surface elevation from a sensor mounted on a fixed offshore platform in the North Sea. Their light-emission system is based on a source sending simultaneously three laser pulses at a very high rate of 190 kHz, leading to 570000 pulses per second. The associated scanning system is composed of two rotating mirrors oriented in the horizontal and vertical directions. In the horizontal direction, the mirror is a polygon with facets leading to a field of view (or aperture angle) of 60° and rotating at 7270 rpm, while in the vertical direction, a galvanometer mirror oscillates back and forth with a rotating motion of amplitude 10° to generate 10 frames per second in each rotation direction. This setup generates two frames of (theoretically) 28500 free surface elevation observations at a rate of 10 Hz. This high measurement rate with respect to the characteristic time of ocean surface evolution allows to consider that, in each frame, all measurements are evaluated at the same instant, forming a spatial map of the surface. Accordingly, the synthetic observations generated in this work are of the form of *instantaneous* spatial datasets of surface elevation measurements.

III.1.2 Implications of Shallow Incidence Angles

As mentioned previously, the altitude z_c of the sensor being limited by the height above the surface level of the structure upon which it is mounted, the vertical viewing angle α of the camera observing the surface at a distance leads necessary to grazing incidence angles of rays with the surface. The incidence angle α_i is defined as the angle made by the lidar ray and the local surface slope, such as depicted in figure III.2. It follows that, the larger the measurement

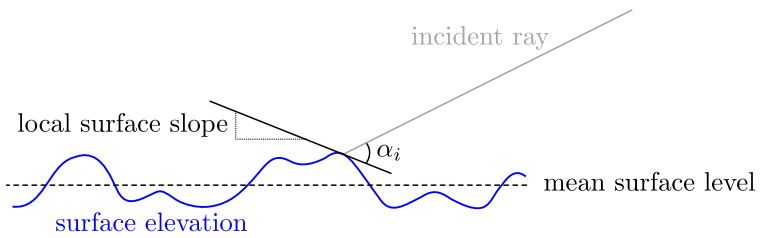

Figure III.2: Definition of the angle of incidence α_i of the incident ray with respect to the sea surface at their crossing location.

distance, the more likely the incidence angle is small. Through a simple geometrical analysis,

we will see that grazing incidence angles have strong implications on the spatial distribution of each dataset.

III.1.2.1 Non-uniform Spatial Sampling

Assuming that the lidar pulse rate is constant and the scanning system relies on mirrors with uniform rotational motions (which is the case for the majority of existing rotating lidar technologies), the angular distribution of rays over the aperture angle is uniform. This angular distribution leads to a geometrical decreasing density of measurement points with the distance from the sensor. Figure III.3 shows an example of the spatial sampling of a one-directional wave field, for which only the altitude z_c, the vertical viewing angle α, aperture α_a and number of laser rays are used for the sensor's definition. The progressive decrease of density of observations

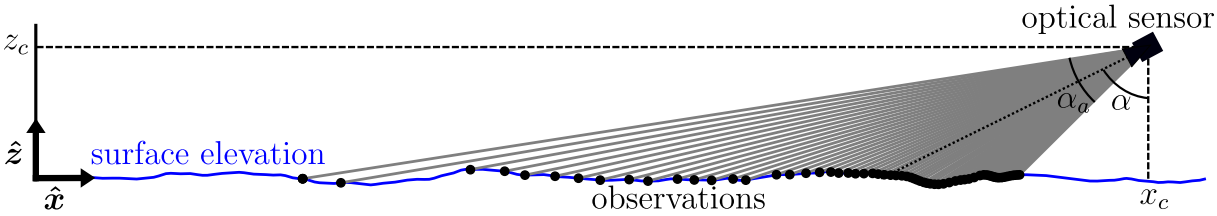

Figure III.3: Spatial sampling of a long-crested wave field by an optical sensor (e.g., lidar camera). Straight lines are ray trajectories.

ahead of the sensor is clear. If the ocean surface were completely flat, the irregularity of the observation grid would be entirely determined by the sensor parameters (z_c, α, α_a, number of beams). However, because of the surface deformation, wave shadowing effects occur, leading, in regions of high surface slope, only wave sides that face the sensor to be illuminated. This has two effects. First, it increases the irregularity in the spatial distribution of the observations, refining the wave crests sampling at the expense of wave troughs. Then, since the wave shadowing phenomenon is related to the local surface elevations, unknown at the time of the measurements, the observations locations cannot be predicted. Observations are thus made at *a priori* unknown locations. As observed in figure III.3 as well as in figure III.4a for a directional wave field, this effect becomes more important at grazing incidence angles, that is for the most distant observation points.

III.1.2.2 Measurements Identification

Optical sensors relying on nadir observations of the surface have already been developed and provide accurate measurement of free surface elevation. However, it becomes technically more challenging to make similar measurements when the incidence angle of the sensor rays gets grazing with respect to the free surface. For example, experimental results obtained by Kabel *et al.* (2019) show that the light intensity of the backscattered signal is significantly higher for small measurement distance, for which the incidence angle remains relatively large, participating further in the decrease observations density with the distance from the lidar system. In this section, we are interested in the modeling of this effect to see how it affects the identification of measurement points. To do this, we assume that when this incidence angle is smaller than a limiting value $\alpha_i^{\lim}$, the power of the backscattered signal is so low that it becomes indistinguishable from the ambient noise and no measurement is possible. This leads to a localized (i.e., non-random) discrimination of observation points. An example of distribution of local incidence angles is shown in figure III.4b. The decreasing probability of obtaining a high incidence angle with the distance from the sensor is clear.

We quantify in figure III.5 the discriminated points over an entire spatial dataset for three different limiting angles of incidence, i.e., $\alpha_i^{\lim} = 3$, 4 and 5°. Besides the fact that the number of

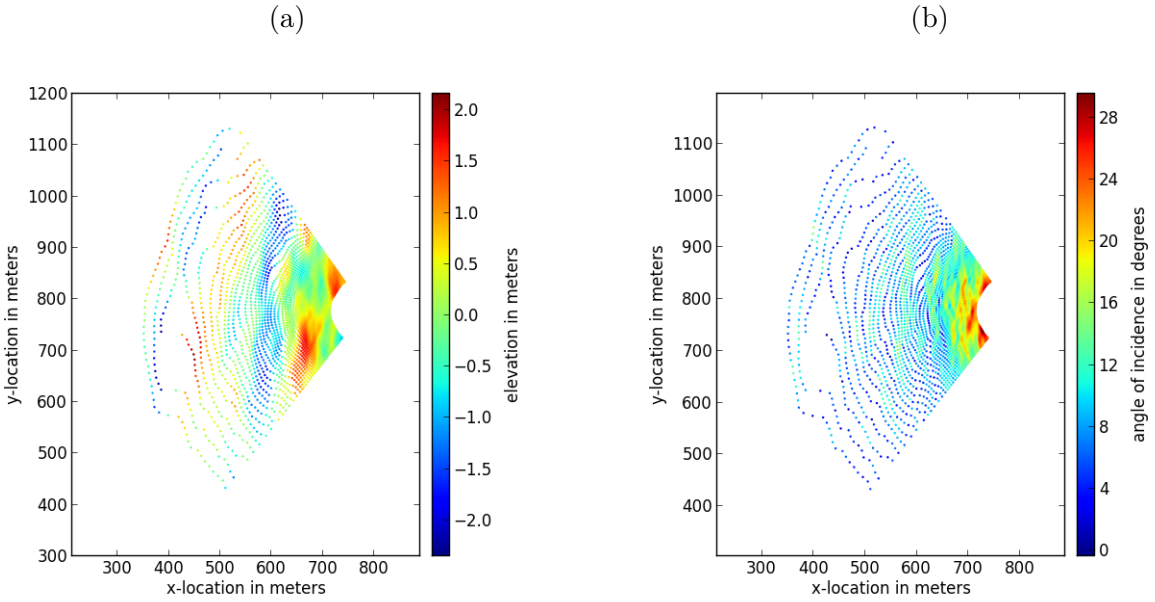

Figure III.4: Example of a spatial distribution of synthetic surface elevation lidar measurements: (a) free surface elevation and (b) incidence angle of the sensor ray with respect to the surface. The underlying sea state is a JONSWAP with $H_s = 3$ m, $T_p = 10$ s, $\gamma = 3.3$ and a main direction of propagation along the x-axis. The optical sensor is located 30 m above the mean surface level and observes the surface with angles $\alpha = 76°$, $\beta = 0°$ and aperture angles $\alpha_a = 20°$, $\beta_a = 110°$, using 64×64 rays. The sensor is located at $(x_c, y_c) = (780, 780)$.

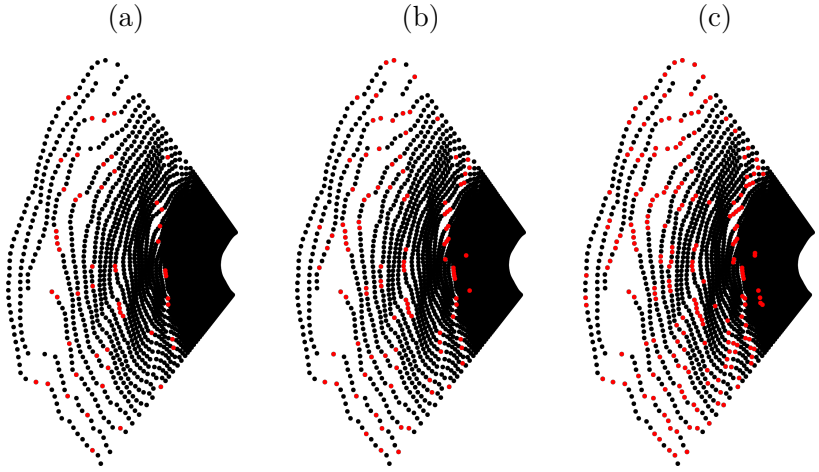

Figure III.5: Distribution of the discriminated points (●) over a set of spatial measurements (●) for three different limiting ray angles of incidence with respect to the surface, $\alpha_i^{\lim}$ = (a) 3°, (b) 4° and (c) 5°. Sea state and optical sensor parameters are the same as in figure III.4.

discriminated points increases with the limiting angle of incidence (i.e., respectively 54, 106 and 200 discriminated points for limiting angles of 3, 4 and 5°), we see that they are not randomly distributed over the spatial dataset, which is in agreement with previous remarks: the points corresponding to large measurement distances (left part of the observation grid) have a larger probability to be discriminated than the other points.

Lidar cameras for free surface elevation measurements are not developed yet, hence no empirical background knowledge is available so far for a high-fidelity modeling of such datasets in terms of distributions, noise level, measurement errors, etc. The present characterization of lidar observations is limited to a purely geometrical approach, and we will not further investigate

their characteristics. Still, the main properties of the expected spatio-temporal distribution of the observations, which, following section I.3, represents the main constrain for the data assimilation method, have been highlighted. The wave model inversion method presented in the next section is thus adapted to the described lidar observations, namely, a selection of instantaneous spatial datasets containing a high number of surface elevations, highly irregularly distributed and whose locations cannot be a priori calculated. Note that, after conversion to surface elevation measurements, radar images give datasets of similar characteristics.

III.2 Variational Assimilation

As detailed in section I.3.4.2, the most common way to assimilate wave elevation data is the variational approach and it is the one that we use to assimilate our lidar observations. The principle is to minimize a quadratic cost function that represents the error between the reconstructed wave field and the observations, through operators that link the wave measurements and the parameters of the wave model. The classical variational cost function is formulated as

$$F(\boldsymbol{p}) = \frac{1}{2}\left(\boldsymbol{p}-\boldsymbol{p}^b\right)^{\mathsf{T}} \mathsf{B}^{-1}\left(\boldsymbol{p}-\boldsymbol{p}^b\right) + \frac{1}{2}\sum_{k=1}^{K}\left(\boldsymbol{\mathcal{Q}}_k\left(\boldsymbol{\mathcal{M}}_k\left(\boldsymbol{p}\right)\right)-\boldsymbol{q}_k\right)^{\mathsf{T}} \mathsf{R}_k^{-1}\left(\boldsymbol{\mathcal{Q}}_k\left(\boldsymbol{\mathcal{M}}_k\left(\boldsymbol{p}\right)\right)-\boldsymbol{q}_k\right),$$

usually written in the more compact form

$$F(\boldsymbol{p}) = \frac{1}{2}\left|\left|\boldsymbol{p}-\boldsymbol{p}^b\right|\right|^2_{\mathsf{B}^{-1}} + \frac{1}{2}\sum_{k=1}^{K}\left|\left|\boldsymbol{\mathcal{Q}}_k\left(\boldsymbol{\mathcal{M}}_k\left(\boldsymbol{p}\right)\right)-\boldsymbol{q}_k\right|\right|^2_{\mathsf{R}_k^{-1}},$$

where $||.||$ denotes the Euclidean norm, and K is total number of instantaneous spatial datasets used in the observations. $\boldsymbol{p}$ is the control vector containing the unknown model parameters, and $\boldsymbol{p}^b$ is a first estimate of model parameters, called background and usually based on the results from previous calculations. $\boldsymbol{\mathcal{M}}_k = \mathcal{M}_{jk}$ ($j = 1, ..., J$ with J the number of spatial observations at each instant) are model operators, mapping between control space and model space. Similarly, $\boldsymbol{\mathcal{Q}}_k = \mathcal{Q}_{jk}$ are observation operators, mapping between model space and observation space. $\boldsymbol{q}_k = q_{jk}$ are the observation vectors containing the measurements. B and R_k are the background and observation error covariance matrices, respectively. These matrices are supposed to provide statistical characterizations of the expected error related to the background estimate solution (for B), and to the measurements and physical model (for R_k).

Note that the formulation of the variational cost function does not require any assumption on the type of observations nor their space/time distribution. Disparate spatio-temporal sets of observations are easily combined in one single minimization problem. Hence, this assimilation method is well adapted to lidar measurements.

III.2.1 Application to Optical Wave Measurements

As a first approximation, we consider that we do not have access to a background and there is no observation error (i.e., no measurement nor wave model errors). Accordingly, the cost function reduces to

$$F(\boldsymbol{p}) = \frac{1}{2}\sum_{k=1}^{K}\left|\left|\boldsymbol{\mathcal{Q}}_k\left(\boldsymbol{\mathcal{M}}_k\left(\boldsymbol{p}\right)\right)-\boldsymbol{q}_k\right|\right|^2. \tag{III.1}$$

For our application, we assume that the lidar measurement system includes the observation operator that links the signal scattered by the surface to the free surface elevations so the observations can then be directly compared to wave models estimates. The observations correspond to the

free surface elevation measurements $\boldsymbol{q}_k = \overline{\boldsymbol{\eta}}_k = \overline{\eta}_{jk}$ at spatial location $\boldsymbol{r}_j$ and time t_k. The model operator corresponds to the wave model. Here, the wave model is ICWM, that is $\boldsymbol{\mathcal{M}}_k = \boldsymbol{\eta}_k$ and $\eta_{jk} = \eta\left(\boldsymbol{r}_j, t_k\right)$ with η following the Eulerian approximation (II.18) of ICWM equation. The wave model parameters are a set of $2N$ unknowns $\boldsymbol{p} = \{a_n, b_n\}$ ($n = 1, ..., N$ with N the number of wave components). In our case, the cost function (III.1) turns into

$$F\left(\boldsymbol{p}\right) = \frac{1}{2}\sum_{k=1}^{K}\left|\left|\boldsymbol{\eta}_k\left(\boldsymbol{p}\right) - \overline{\boldsymbol{\eta}}_k\right|\right|^2 = \frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{J}\left(\eta_{jk}\left(\boldsymbol{p}\right) - \overline{\eta}_{jk}\right)^2 = \frac{1}{2}\sum_{\ell=1}^{L}\left(\eta_\ell\left(\boldsymbol{p}\right) - \overline{\eta}_\ell\right)^2,$$

in which ℓ is a double index that denotes the spatio-temporal location of the free surface elevation, and $L = J \times K$ is the total number of observations. Note that the minimization of the cost function simplifies to a least square minimization, which can be achieved by specifying

$$\boldsymbol{\nabla} F\left(\boldsymbol{p}\right) = 0 \iff \begin{cases} \dfrac{\partial F}{\partial a_m} = 0, \\ \dfrac{\partial F}{\partial b_m} = 0, \end{cases} \qquad m = 1, ..., N. \tag{III.2}$$

In section I.3, we saw that the calculation of the gradient of the cost function is non-trivial for nonlinear wave models. Indeed, the reviewed models involve either mathematical terms that are too complicated for an efficient gradient calculation (e.g., double summations in the expression of the (enhanced) second-order solution), or implicit terms that require the use of numerical integrations (e.g., HOS), both being computationally expensive. The reviewed minimization problems thus rely on optimization procedures (based or not on the gradient calculation) that iteratively converge to the optimal model parameters. Here, since ICWM is analytical and has almost the same mathematical complexity as LWT, we choose to explicitly formulate the gradient of the cost function in the form of a system that is directly inverted. Such as proposed by Grilli *et al.* (2011) and Nouguier *et al.* (2014), a linear approximation followed by nonlinear iterations are used to converge to the nonlinear solution.

III.2.2 System Formulation

Developing the terms from system (III.2) leads to

$$\frac{\partial F}{\partial a_m} = \sum_{\ell=1}^{L}\frac{\partial\left(\eta_\ell - \overline{\eta}_\ell\right)}{\partial a_m}\left(\eta_\ell - \overline{\eta}_\ell\right) = \sum_{\ell=1}^{L}\frac{\partial \eta_\ell}{\partial a_m}\left(\eta_\ell - \overline{\eta}_\ell\right),$$

and similarly $\left(F\right)_{b_m} = \sum_\ell \left(\eta_\ell\right)_{b_m}\left(\eta_\ell - \overline{\eta}_\ell\right)$, where the notation $\left(x^*\right)_{y^*}$ is used to denote the derivation of the quantity x^* with respect to y^*. We thus have to determine the derivative terms $\left(\eta_\ell\right)_{a_m, b_m}$ to explicitly formulate the system that is solved for the wave parameters $\boldsymbol{p}$. Using the approximate Eulerian formulation of ICWM, equation (II.18), we get

$$\begin{aligned} \left(\eta_\ell\right)_{a_m} &= \left(a_m \cos\Psi_{m\ell}\right)_{a_m} + \left(b_m \sin\Psi_{m\ell}\right)_{a_m} + a_m k_m, \\ \left(\eta_\ell\right)_{b_m} &= \left(a_m \cos\Psi_{m\ell}\right)_{b_m} + \left(b_m \sin\Psi_{m\ell}\right)_{b_m} + b_m k_m. \end{aligned}$$

We then calculate

$$\begin{aligned} \left(a_m \cos\Psi_{m\ell}\right)_{a_m} &= \cos\Psi_{m\ell} - a_m \sin\Psi_{m\ell}\left(\Psi_{m\ell}\right)_{a_m}, \\ \left(b_m \sin\Psi_{m\ell}\right)_{a_m} &= b_m \cos\Psi_{m\ell}\left(\Psi_{m\ell}\right)_{a_m}, \\ \left(a_m \cos\Psi_{m\ell}\right)_{b_m} &= -a_m \sin\Psi_{m\ell}\left(\Psi_{m\ell}\right)_{b_m}, \\ \left(b_m \sin\Psi_{m\ell}\right)_{b_m} &= \sin\Psi_{m\ell} + b_m \cos\Psi_{m\ell}\left(\Psi_{m\ell}\right)_{b_m}. \end{aligned}$$

We now have to determine the modal derivatives of the phase term $\Psi_{m\ell}$, for which we recall here the expression

$$\begin{cases} \Psi_{m\ell} = \boldsymbol{k}_m \cdot \left[\boldsymbol{r}_\ell - \sum_{n=1}^{N} \hat{\boldsymbol{k}}_n \left(-a_n \sin\tilde{\phi}_{n\ell} + b_n \cos\tilde{\phi}_{n\ell} \right) \right] - \tilde{\omega}_{\mathrm{L}m} t_\ell, \\ \tilde{\phi}_{n\ell} = \boldsymbol{k}_n \cdot \boldsymbol{r}_\ell - \tilde{\omega}_{\mathrm{L}n} t_\ell, \\ \tilde{\omega}_{\mathrm{L}m} = \omega_m + 1/2 \boldsymbol{k}_m \cdot \boldsymbol{\mathcal{U}}_s, \\ \boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N} \left(a_n^2 + b_n^2 \right) \omega_n \boldsymbol{k}_n. \end{cases}$$

Its derivatives can be written in the form

$$\left(\Psi_{m\ell}\right)_{a_m,b_m} = \left(X_1 + X_2 + X_3\right)_{a_m,b_m},$$

where

$$X_1 = k_m a_m \sin\tilde{\phi}_{m\ell}, \qquad X_2 = -k_m b_m \cos\tilde{\phi}_{m\ell} \qquad \text{and} \qquad X_3 = -\tilde{\omega}_{\mathrm{L}m} t_\ell.$$

for which we calculate

$$\begin{cases} \left(X_1\right)_{a_m} = k_m \left[\sin\tilde{\phi}_{m\ell} + a_m \cos\tilde{\phi}_{m\ell} \left(\tilde{\phi}_{m\ell}\right)_{a_m} \right], \\ \left(X_2\right)_{a_m} = k_m b_m \sin\tilde{\phi}_{m\ell} \left(\tilde{\phi}_{m\ell}\right)_{a_m}, \\ \left(X_3\right)_{a_m} = -a_m \omega_m k_m^2 t_\ell, \end{cases} \qquad \begin{cases} \left(X_1\right)_{b_m} = k_m a_m \cos\tilde{\phi}_{m\ell} \left(\tilde{\phi}_{m\ell}\right)_{b_m}, \\ \left(X_2\right)_{b_m} = -k_m \left[\cos\tilde{\phi}_{m\ell} - b_m \sin\tilde{\phi}_{m\ell} \left(\tilde{\phi}_{m\ell}\right)_{a_m} \right], \\ \left(X_3\right)_{b_m} = -b_m \omega_m k_m^2 t_\ell. \end{cases}$$

Noticing that $\left(\tilde{\phi}_{m\ell}\right)_{a_m,b_m} = \left(X_3\right)_{a_m,b_m}$, we deduce the expressions for the phase terms

$$\begin{aligned} \left(\Psi_{m\ell}\right)_{a_m} &= k_m \left\{ \sin\tilde{\phi}_{m\ell} - \left[k_m \left(a_m \cos\tilde{\phi}_{m\ell} + b_m \sin\tilde{\phi}_{m\ell} \right) + 1 \right] a_m \omega_m k_m t_\ell \right\}, \\ \left(\Psi_{m\ell}\right)_{b_m} &= k_m \left\{ -\cos\tilde{\phi}_{m\ell} - \left[k_m \left(a_m \cos\tilde{\phi}_{m\ell} + b_m \sin\tilde{\phi}_{m\ell} \right) + 1 \right] b_m \omega_m k_m t_\ell \right\}. \end{aligned}$$

It is now possible to retrieve the modal derivatives of the ICWM free surface elevation

$$\begin{aligned} \left(\eta_\ell\right)_{a_m} = {} & \cos\Psi_{m\ell} - k_m \left(a_m \sin\Psi_{m\ell} - b_m \cos\Psi_{m\ell} \right) \\ & \times \left\{ \sin\tilde{\phi}_{m\ell} - \left[k_m \left(a_m \cos\tilde{\phi}_{m\ell} + b_m \sin\tilde{\phi}_{m\ell} \right) + 1 \right] a_m \omega_m k_m t_\ell \right\} + a_m k_m, \\ \left(\eta_\ell\right)_{b_m} = {} & \sin\Psi_{m\ell} - k_m \left(a_m \sin\Psi_{m\ell} - b_m \cos\Psi_{m\ell} \right) \\ & \times \left\{ -\cos\tilde{\phi}_{m\ell} - \left[k_m \left(a_m \cos\tilde{\phi}_{m\ell} + b_m \sin\tilde{\phi}_{m\ell} \right) + 1 \right] b_m \omega_m k_m t_\ell \right\} + b_m k_m. \end{aligned}$$

Denoting $\left(\eta_\ell\right)_{a_m} = P_{m\ell}$ and $\left(\eta_\ell\right)_{b_m} = Q_{m\ell}$, system (III.2) turns into

$$\begin{cases} \sum_{\ell=1}^{L} \sum_{n=1}^{N} a_n \left(\cos\Psi_{n\ell} + \frac{1}{2} a_n k_n \right) P_{m\ell} + b_n \left(\sin\Psi_{n\ell} + \frac{1}{2} b_n k_n \right) P_{m\ell} = \sum_{\ell=1}^{L} \overline{\eta}_\ell P_{m\ell}, \\ \sum_{\ell=1}^{L} \sum_{n=1}^{N} a_n \left(\cos\Psi_{n\ell} + \frac{1}{2} a_n k_n \right) Q_{m\ell} + b_n \left(\sin\Psi_{n\ell} + \frac{1}{2} b_n k_n \right) Q_{m\ell} = \sum_{\ell=1}^{L} \overline{\eta}_\ell Q_{m\ell}. \end{cases} \tag{III.3}$$

This system is then recast in a matrix form $\mathsf{A}\boldsymbol{p} = \mathsf{B}$, where, $\boldsymbol{p}$ is our wave model parameters vector containing $2N$ unknown elements

$$p_n = a_n, \quad p_{N+n} = b_n,$$

$\mathbf{B}$ is a vector containing the observation information

$$\mathsf{B}_m = \sum_{\ell=1}^{L} \overline{\eta}_\ell P_{m\ell}, \qquad \mathsf{B}_{N+m} = \sum_{\ell=1}^{L} \overline{\eta}_\ell Q_{m\ell},$$

and A is a $2N \times 2N$ matrix that follows

$$\mathsf{A}_{mn} = \sum_{\ell=1}^{L} \left(\cos \Psi_{n\ell} + \frac{1}{2} a_n k_n \right) P_{m\ell}, \qquad \mathsf{A}_{m,N+n} = \sum_{\ell=1}^{L} \left(\sin \Psi_{n\ell} + \frac{1}{2} b_n k_n \right) P_{m\ell},$$

$$\mathsf{A}_{N+m,n} = \sum_{\ell=1}^{L} \left(\cos \Psi_{n\ell} + \frac{1}{2} a_n k_n \right) Q_{m\ell}, \qquad \mathsf{A}_{N+m,N+n} = \sum_{\ell=1}^{L} \left(\sin \Psi_{n\ell} + \frac{1}{2} b_n k_n \right) Q_{m\ell}.$$

System (III.3) is finally solved for optimal wave parameters (a_n, b_n). However, since both A and $\mathbf{B}$ depend on wave parameters, we make use of an iterative procedure, detailed in the next section, to find the appropriate nonlinear solution. Note that similar inverse problems can be formulated using CWM1 and LWT-CDR (see appendix A for details).

III.2.3 Linearization and Nonlinear Iterations

To be able to find solutions to the inverse problem (III.3), equations are linearized by computing $\mathsf{A}^{(i)}$ and $\mathbf{B}^{(i)}$ based on wave parameters obtained at the previous iteration i when solving for $\boldsymbol{p}^{(i+1)}$ at iteration $i+1$. The solution is initialized at $i = 0$ using $\mathsf{A}^{(0)}$ and $\mathbf{B}^{(0)}$ based on the reconstruction problem formulated with LWT and detailed below. At each iteration $i+1$, we evaluate relative errors $\mathcal{E}_a$ and $\mathcal{E}_b$ between $\boldsymbol{p}^{(i)} = \left\{ a_n^{(i)}, b_n^{(i)} \right\}$ and $\boldsymbol{p}^{(i+1)} = \left\{ a_n^{(i+1)}, b_n^{(i+1)} \right\}$, which take the form

$$\mathcal{E}_{\mathsf{x}} = \left[\sum_{n}^{N} \left(\mathsf{x}_n^{(i+1)} - \mathsf{x}_n^{(i)} \right)^2 \Big/ \sum_{n}^{N} \left(\mathsf{x}_n^{(i+1)} \right)^2 \right]^{1/2},$$

in which x is refers either to a or b. The solution is considered converged when both $\mathcal{E}_a$ and $\mathcal{E}_b$ are inferior to a tolerance $tol = 10^{-6}$, which is typically achieved within a few to a few dozens iterations depending on the wave steepness. Due to the problem formulation, the iteration process does not necessary converge, requiring to set a maximal number of iterations $N_i^{\max}$. This number is set high enough, such that the value of tol does not influence the number of inversions reaching $N_i^{\max}$. A typical value of 100 was used in the presented numerical and experimental applications of our prediction method.

Nonlinear inversions for which the number of iterations reached $N_i^{\max}$ lead to inconsistent physical representations of the ocean surface, which is why they are discarded from the ensemble of results that is used to calculate the indicators that quantify the prediction accuracy. The apparent randomness of the number of iterations – even if it clearly tends to increase for higher wave steepness, the (nonuniform) spatial distribution of observations also play a significant role – prevented us from properly evaluating the expected proportion of cases for which the number of iterations reached $N_i^{\max}$. However, in operational conditions, this non-convergence is not a major limitation since the linear inversion, which corresponds to the initialization step of the nonlinear inversion (see below), could always be used instead.

The initial solution $\boldsymbol{p}^{(0)}$ is found based on the inversion of LWT. In order to simplify the system formulation, the linear elevation (I.9) is recast in the equivalent form

$$\eta^{\text{lin}}(\boldsymbol{r}, t) = \sum_{n=1}^{N} a_n \cos \psi_n + a_n \sin \psi_n,$$

in which $\psi_n = \boldsymbol{k}_n \cdot \boldsymbol{r} - \omega_n t$ are spatio-temporal phases and $(a_n, b_n) = (A_n \cos\varphi_n, A_n \sin\varphi_n)$ are the wave coefficients describing the ocean surface. Similar to ICWM, we find the free surface elevation derivatives with respect to the wave coefficients, which yield

$$\begin{aligned}\left(\eta_\ell^{\text{lin}}\right)_{a_m} &= \cos\psi_{m\ell},\\ \left(\eta_\ell^{\text{lin}}\right)_{b_m} &= \sin\psi_{m\ell},\end{aligned}$$

and lead to the following linear system

$$\begin{cases}\displaystyle\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n \cos\psi_{n\ell}\cos\psi_{m\ell} + b_n \sin\psi_{n\ell}\cos\psi_{m\ell} = \sum_{\ell=1}^{L}\overline{\eta}_\ell\cos\psi_{m\ell},\\ \displaystyle\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n \cos\psi_{n\ell}\sin\psi_{m\ell} + b_n \sin\psi_{n\ell}\sin\psi_{m\ell} = \sum_{\ell=1}^{L}\overline{\eta}_\ell\sin\psi_{m\ell}.\end{cases}$$

In its matrix form $\mathsf{A}\boldsymbol{p} = \mathbf{B}$, $\mathbf{B}$ is now computed as

$$\mathsf{B}_m = \sum_{\ell=1}^{L}\overline{\eta}_\ell\cos\psi_{m\ell}, \qquad \mathsf{B}_{N+m} = \sum_{\ell=1}^{L}\overline{\eta}_\ell\sin\psi_{m\ell},$$

and A reads

$$\begin{aligned}\mathsf{A}_{mn} &= \sum_{\ell=1}^{L}\cos\psi_{n\ell}\cos\psi_{m\ell}, & \mathsf{A}_{m,N+n} &= \sum_{\ell=1}^{L}\sin\psi_{n\ell}\cos\psi_{m\ell},\\ \mathsf{A}_{N+m,n} &= \sum_{\ell=1}^{L}\cos\psi_{n\ell}\sin\psi_{m\ell}, & \mathsf{A}_{N+m,N+n} &= \sum_{\ell=1}^{L}\sin\psi_{n\ell}\sin\psi_{m\ell}.\end{aligned}$$

Inverting this system allows to get access to the representation of the free surface according to LWT. The solution also serves as initialization of the iteration procedure described above for the ICWM inversion. A similar method is used for the inversion of the other nonlinear wave models, CWM1 and LWT-CDR, that are used in the presented numerical and experimental applications.

III.2.4 Regularization Procedure

In operative applications, the non-uniform spatial distribution of the measurements, the non-periodicity of the assimilated data (while our wave model is intrinsically periodic), as well as the generally overdetermined nature of the problem (i.e., more observations than unknowns) can cause the inverse problem to be ill-conditioned. Nevertheless, consistent results can be achieved, independently of the conditioning of the system matrix to invert (i.e., A), by applying a Tikhonov regularization (Tikhonov, 1963), in which the matrix inversion is replaced by the following minimization problem

$$\min\left\{||\mathsf{A}\boldsymbol{p} - \mathbf{B}||^2 + \mathsf{r}^2\,||\boldsymbol{p}||^2\right\}, \tag{III.4}$$

where r denotes the regularization parameter. The value of r conditions the problem to provide an optimal compromise between minimizing the residual error of the assimilation system and ensuring that the norm of the solution does not become too large, avoiding overfitting issues.

III.2.4.1 L-curve Method for the Selection of the Regularized Parameter

The optimal value of the regularization parameter is found using the 'L-curve' method, which consists in finding the r value corresponding to the point of maximal curvature (i.e., corner) of the parametric curve $(\log||\mathsf{A}\boldsymbol{p}-\mathbf{B}||, \log||\boldsymbol{p}||)$. The L-curve corner can be determined analytically through solving a singular value decomposition (SVD) problem (Calvetti *et al.*, 2004; Hansen, 2000). The SVD consists in factorizing the matrix we want to invert in the form

$$\mathsf{A} = \sum_{i=1}^{2N} \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^{\mathrm{T}},$$

where the singular values σ_i are non-negative quantities appearing in decreasing order, i.e., $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_{2N} \geq 0$. Vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the left and right singular vectors, respectively, and are orthogonal, i.e., $\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{u}_j = \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{v}_j = \delta_{ij}$ (with δ_{ij} the δ-function). The number of non-zero singular values indicates the rank of the matrix, hence, singular values of a well-conditioned matrix are all strictly positive. Injecting this factorization in the least square minimization (III.4), a regularized solution $\boldsymbol{p}_\mathsf{r}$ for a particular regularization parameter r is found as

$$\boldsymbol{p}_\mathsf{r} = \sum_{i=1}^{2N} f_i \frac{\boldsymbol{u}_i^{\mathrm{T}}\mathbf{B}}{\sigma_i}\boldsymbol{v}_i,$$

in which $f_i = \sigma_i^2/\left(\sigma_i^2+\mathsf{r}^2\right)$ are the Tikhonov filter factors varying between zero and one. The role of Tikhonov filters is to give more weight to large singular values as the regularization parameter increases, which stabilizes the inversion of A. For the 'naive' solution (i.e., $\mathsf{r}=0$), all the Tikhonov filter factors are equal to one.

The minimization residual can also be written as a function of the resulting terms of the SVD, and the squared norms of the solution and of the associated residual read

$$||\boldsymbol{p}_\mathsf{r}||^2 = \sum_{i=1}^{2N}\left(f_i\frac{\boldsymbol{u}_i^{\mathrm{T}}\mathbf{B}}{\sigma_i}\boldsymbol{v}_i\right)^2 = u \qquad \text{and} \qquad ||\mathsf{A}\boldsymbol{p}_\mathsf{r}-\mathbf{B}||^2 = \sum_{i=1}^{2N}\left[\left(1-f_i\right)\boldsymbol{u}_i^{\mathrm{T}}\mathbf{B}\right]^2 = v.$$

With these explicit expressions, and denoting u' the derivative of u with respect to r, the curvature of the L-curve yields (Hansen, 2000)

$$\kappa = 2\frac{uv}{u'}\frac{\mathsf{r}^2u'v+2\mathsf{r}uv+\mathsf{r}^4uu'}{\left(\mathsf{r}^2u^2+v^2\right)^{3/2}}. \tag{III.5}$$

The next step consists in sweeping values of r to find the point of maximal curvature. An example of an L-curve and its corresponding curvature is shown in figure III.6. For very low (high) r-values, the L-curve is vertical (horizontal), meaning that the residual (solution) norm is constant and the solution (residual) norm drops logarithmically to zero. Between these two limits of very low and very high r-values, we see that the curve is not clearly shaped as an 'L' with one corner as it would ideally be, but exhibits several corners that represent local optimal regularizations. Regarding the very small computational cost for the calculation of κ, the solution to find the optimal regularization parameter is to sweep several values in a relevant range. From figure III.6, it appears important to consider a relatively large range of r-values to be sure to catch the global optimum, highlighted with a red dot on the presented curves. In our applications, r is swept within the interval $\left[10^{-5}, 10^{5}\right]$, logarithmically sampled with 10^3 values, which was found to be appropriate for our model inversions.

III.2.4.2 Alternative Approach based on Wave Energy

Note that for a linear wave field, this procedure is equivalent to adding a constraint to the minimization problem, physically representing the total energy of the reconstructed wave spectrum,

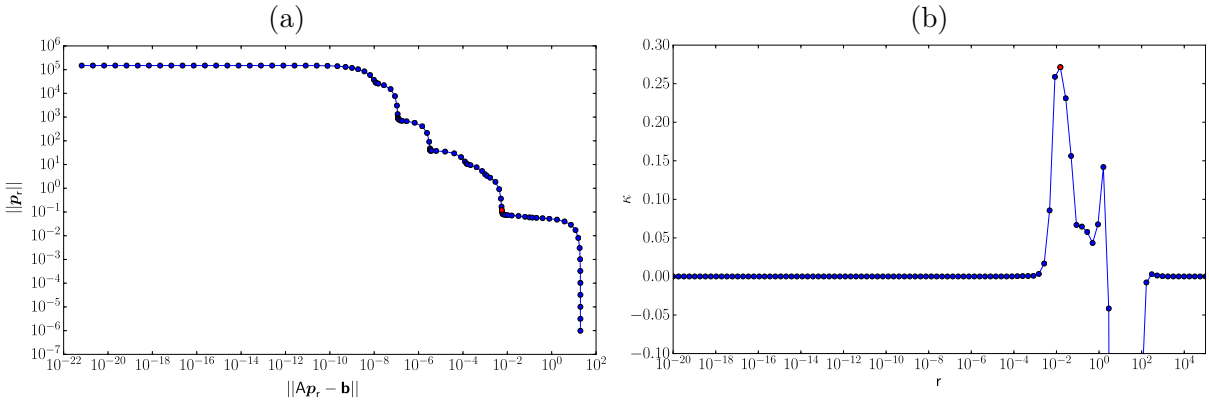

Figure III.6: Example of (a) an L-curve obtained for the regularization of a wave model inversion and (b) its corresponding curvature calculated with equation (III.5). The point of maximal curvature is highlighted in red.

since the latter is proportional to the squared norm of $\boldsymbol{p}$. Indeed, following relation (I.4), we write

$$||\boldsymbol{p}||^2 = \sum_{n}^{N} \left(a_n^2 + b_n^2\right) \approx 2 \iint S\left(k, \theta\right) \mathrm{d}\theta \, \mathrm{d}k = 2\, m_0,$$

and the zeroth-order moment m_0 correspond to the variance σ_η^2 of the free surface elevation, which can be easily estimated from a standard deviation σ_η of set of elevation measurements. By selecting a regularized parameter that leads to $||\boldsymbol{p}_\mathsf{r}|| = \sqrt{2}\sigma_\eta$, the solution would yield a consistent energy quantity with the actual observations, giving a simpler way to determine the optimal r than going through the calculation of the L-curve curvature. However, our inversions using this technique generally led to solutions giving too much weight to high frequency components, that is the regularization parameter was too low, leading to solutions that overfitted the observations. Hence, we use the L-curve curvature method in the following.

III.2.4.3 Note on the Computational Cost of an ICWM Inversion

Following Plassman (2005), the SVD calculation for a dense, non-symmetric matrices is divided in three phases: an orthonormal reduction of A to a bidiagonal form (phase 1), the computation of the SVD of the bidiagonal form (phase 2) and the computation of the SVD of A (phase 3). Phase 1 is typically calculated using Householder reflectors, such as in the Golub-Kahan bidiagonalization algorithm (e.g., Trefethen & Bau, 1997, p. 236), and yields an amount of computations of $\frac{8}{3}\left(2N\right)^3$ for a $2N \times 2N$ matrix (corresponding to the inversion of a wave model with N wave components). Phase 2 relies on an iterative procedure and the computational cost is dependent on the performance of the chosen algorithm. The 'divide and conquer' approach is widely used, and is the one employed by the LAPACK routine `_gesdd` used in this work. It requires, for the worst case, $\mathcal{O}\left(8N^3\right)$ computations. Phase 3 has a typical cost of $\mathcal{O}\left(8N^3\right)$ computations. This amount of operations, multiplied by the number of nonlinear iterations of the described iterative procedure, is not compatible for real-time applications when N gets large, which is the case for directional wave fields.

However, it is to be noted that, in the presented inversion method, efforts are concentrated on accuracy and reliability of the predictions, and the described algorithm based on the SVD calculation does not take full advantage of the computational efficiency of ICWM. Its computational advantage compared to other wave models is the modeling of nonlinear effects through explicit and simple mathematical terms.

In case of observations compatible with Fourier analysis, and in similar way to that of Duncan & Drake (1995), an iterative procedure that aims at decoupling the linear wave components from their nonlinear interactions can be performed with the cost of only three FFTs at each iteration, namely $\mathcal{O}(3N \log N)$ for a wave field modeled with N wave components. Indeed, the ICWM surface elevation at iteration i is calculated through $\eta^{(i)} = F_\eta\left(\hat{\eta}^{(i)}\right)$ where F_η has the cost of 2 IFFTs (according to equation (II.18)) and $\hat{\eta}^{(i)} = \mathrm{FFT}\left(\overline{\eta}^{\mathrm{lin}(i)}\right)$ are the Fourier coefficients of the quantity $\overline{\eta}^{\mathrm{lin}(i)} = \overline{\eta} - C^{\mathrm{nl}(i)}$ that corresponds to the observations $\overline{\eta}$ in which the ICWM nonlinear contributions have been discarded. Nonlinear contributions simply read $C^{\mathrm{nl}(i)} = \eta^{(i-1)} - \mathrm{IFFT}\left(\hat{\eta}^{(i-1)}\right)$ and is equal to zero at initialization step $i = 0$.

The implementation of such a fast ICWM inversion method is direct but necessitates wave measurements compatible with Fourier analysis, which is not directly the case of remote optical measurements. Hence, the computation of Cartesian surface elevation maps from optical measurements would imply interpolation errors that affect the prediction accuracy. The quantification of the errors induced by the interpolation is not provided in this work but should be investigated prior to implementation of the described inversion procedure base on (I)FFTs.

III.3 Prediction Zone

As mentioned earlier, the spatio-temporal region within which the predictions resulting from the analysis of a specific set of wave measurements is bounded. Specifically, when wave measurements are made at a specific sampling rate and over a given observation zone, the amount of accessible wave information contained in the measurements limits the assimilation process, yielding a modeled surface in space/time defined with finite frequency and direction bandwidths. In light of this, the sea-state prediction obtained by propagating the assimilated information is similarly limited to a spatio-temporal region referred to as *prediction zone*. This region can be theoretically calculated on the basis of the spatio-temporal distribution of the observations. Using the fact that the 'observed' (or measured) waves propagate at their respective phase velocity, Morris *et al.* (1998, 1992) introduced the concept of prediction diagrams, defining the accessible prediction zone by the intersection of the regions covered by the propagation of each measured wave component, the limiting ones being the slowest and the fastest components. Even if an elementary wave component propagates at its phase velocity, Wu (2004) showed that the accurate description of a wave field is limited to the knowledge of its wave components energy, which propagates at the wave group velocity, and found consistent results supporting this choice. Abusedra & Belmont (2011) used the stationary phase approximation to justify the choice of the group velocity over the phase velocity, but they also indicate that it only applies under the assumption of large propagation distances. Making use of wave tank experiments and numerical simulations, Naaijen *et al.* (2014) validated the approach based on the group velocity, even for short prediction distances. Finally, Qi *et al.* (2018*b*) provides a theoretical justification not restrained to a short prediction distances and valid even for nonlinear wave models. Through statistical analysis of wave properties, Fucile *et al.* (2018) and Mérigaud & Ringwood (2019) derived a formulation of the prediction error standard deviation and of the prediction mean error, respectively, that agree well with the prediction diagrams relying on group velocities of the waves. Accordingly, we thus consider in the following that the measured information related to wave components propagates at group velocities.

III.3.1 Theoretical Prediction Error

A measure of the prediction error at each point in time and space can be formulated based on the propagation properties of wave information. Indeed, the dispersive behavior of ocean gravity waves makes wave information travels at different speeds depending on their frequency (or equivalently their wavenumber). Hence, they get dispersed when the wave field is propagated, making the accessible wave information dependent on the time and space location of the point at

which the prediction is issued. The amount of accessible wave information can be used to define an amount of wave energy based on the integration of the wave spectrum over a wavenumber band bounded by the wavenumbers associated with the fastest and slowest waves whose information reaches the prediction point.

For a particular set of observations of a unidirectional wave field, we define the ratio of the amount of wave energy described by the accessible information to the total energy as

$$P(x,t) = \int_{k_1}^{k_2} E(k)\,\mathrm{d}k \Big/ m_0, \tag{III.6}$$

in which wavenumbers $k_{1,2}$ correspond to the slowest and fastest wave components whose group velocity allows the point (x,t) to be reached by the measured wave information, $E(k) = \int_0^{2\pi} S(k,\theta)\,\mathrm{d}\theta$ is the surface elevation wavenumber spectrum and m_0 the associated zeroth-order moment. The quantity $1-P$ can be used as a theoretical misfit indicator. k_1 (k_2) is defined as the lowest (highest) measurable wavenumber by the observations.

In the example presented in figure III.7, we consider a unidirectional wave field where all waves are propagating in the same direction and whose energy distribution follows a JONSWAP spectrum of peakedness parameter $\gamma = 3.3$ (see equation (I.2) for detailed formulation). The

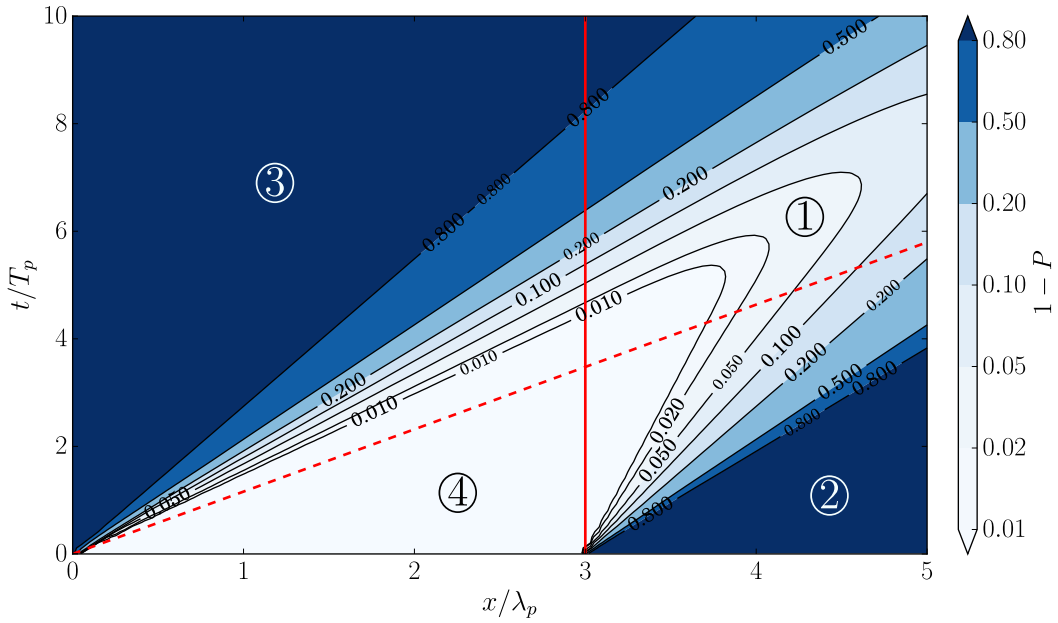

Figure III.7: Spatio-temporal distribution of the theoretical prediction error $1-P$ based on the accessible measured information. The dashed line (- - - -) corresponds to the advection of wave information related to the longest measured wave (i.e., wavelength of $3\lambda_p$), while the continuous line (——) corresponds to the advection of wave information related to an infinitely high frequency.

observations are distributed in space from $x = 0$ to $L_o = 3\lambda_p$, and we choose to set $t = 0$ at the beginning of the wave model propagation. The lowest measurable wavenumber is thus $2\pi/L_o$, and the highest measurable wavenumber is taken equal to $+\infty$ (i.e., we consider that the observations have an infinite sampling rate). The spatio-temporal domain is divided in four distinct subdomains (limited by the two crossing lines) within which group velocities are calculated in different ways. Wave information reaching points located in regions ② and ④ is limited by the group velocity of the longest measurable wave, whose wave number is $2\pi/L_o$. Since every wave component propagates in the positive x-direction, points located in regions ③ and ④ can be reached by wave information that is limited by the static information associated with wavenumber $k \to +\infty$. Wave information reaching points located in region ① propagates at

velocities whose values are lower than the group velocity of the wavenumber $2\pi/L_o$ and higher than the limit of zero (i.e., group velocity of a wave of infinite wavenumber). For the sake of simplicity, we use the linear dispersion relation in deep water to relate group velocities to corresponding wavenumbers, but its finite-depth version could also be used. Formally, the limits for the integral term in the numerator of P yield

$$\textcircled{1}\begin{vmatrix} k_1 = g/\left(2c_{g1}\right)^2, \\ k_2 = g/\left(2c_{g2}\right)^2, \end{vmatrix} \quad \textcircled{2}\begin{vmatrix} k_1 = 2\pi/L_o, \\ k_2 = g/\left(2c_{g2}\right)^2, \end{vmatrix} \quad \textcircled{3}\begin{vmatrix} k_1 = g/\left(2c_{g1}\right)^2, \\ k_2 = +\infty, \end{vmatrix} \quad \textcircled{4}\begin{vmatrix} k_1 = 2\pi/L_o, \\ k_2 = +\infty, \end{vmatrix}$$

where the fastest group velocity is calculated as $c_{g1} = x/t$ and the slowest is $c_{g2} = (x - L_o)/t$.

This same indicator has been employed by Naaijen *et al.* (2014) and was verified to give consistent results with numerical simulations and wave tank experiments. More than the just a binary indicator that states if a point (x, t) is in/out a prediction zone, this indicator allows to get access to an a priori estimate of the prediction error. For the example presented in figure III.7, the region in which $1 - P$ is lower than 1% is bounded by a prediction time of $\sim 5T_p$ for a downstream point located $\sim \frac{3}{4}\lambda_p$ past the last observation. The region in which $1 - P$ is lower than 10% extents to $t > 8T_p$ and $x > 5\lambda_p$ (i.e., $> 2\lambda_p$ past the last observation). The extent of the accessible prediction zone is thus strongly dependent on the desired prediction accuracy. From equation (III.6), we notice that it is also dependent on the wave spectrum: for a similar spectral shape (e.g., JONSWAP with different γ values), the larger the spectral bandwidth of the spectrum, the higher the value of theoretical misfit $1 - P$.

III.3.2 Accessible Prediction Zone Evolution

We saw in the previous section that the limits of the prediction zone depends on the expected prediction accuracy: the higher the prediction accuracy, the smaller the extent of the accessible prediction zone. In this work, the performance of the predictions are quantified on a properly limited prediction zone, hence, clear definitions of boundaries need to be formulated. Assuming that limiting frequency and direction bandwidths of the wave model are known, we describe in this section the spatio-temporal evolution of the prediction zone boundaries in both cases of long- and short-crested waves when wave measurements are made using an optical sensor such as a lidar camera.

III.3.2.1 Spatio-Temporal Evolution for Long-Crested Waves

As detailed in the introduction of this section, both theoretical and experimental studies have shown that disturbances associated with wave components in a dispersive wave field, of given amplitude, frequency, and phase, travel at the associated group velocity along their direction of propagation. The intersection of the slowest and fastest components thus determines the boundary of the spatio-temporal region within which information is available and a prediction can be issued. The choice of the limiting wave frequencies is done regarding an "acceptable" level of error (see section III.3.1 for details), and is discussed in section III.3.4. Consequently, as time increases, the accessible prediction zone, denoted $\mathcal{P}$, shrinks, to eventually disappear when the assimilated information is completely dispersed in space. Figure III.8 illustrates this phenomenon for a unidirectional wave field propagating in the x-direction. The last time used for the assimilation corresponds to the reconstruction time t_r. If only spatial data is used in the assimilation (figure III.8a), the prediction zone at reconstruction time $\mathcal{P}(t_r)$ is the spatial area where observations were made. However, when spatio-temporal datasets are used (over an assimilation time T_a, figure III.8b), $\mathcal{P}(t_r)$ expands due to the advection of wave information during T_a. A point $(x, t \geq t_r)$ is included in the prediction zone if

$$x_o^{\mathrm{b}} + c_{g1}\left(t - t_r\right) \leq x \leq x_o^{\mathrm{e}} + c_{g2}\left(t - t_r\right), \tag{III.7}$$

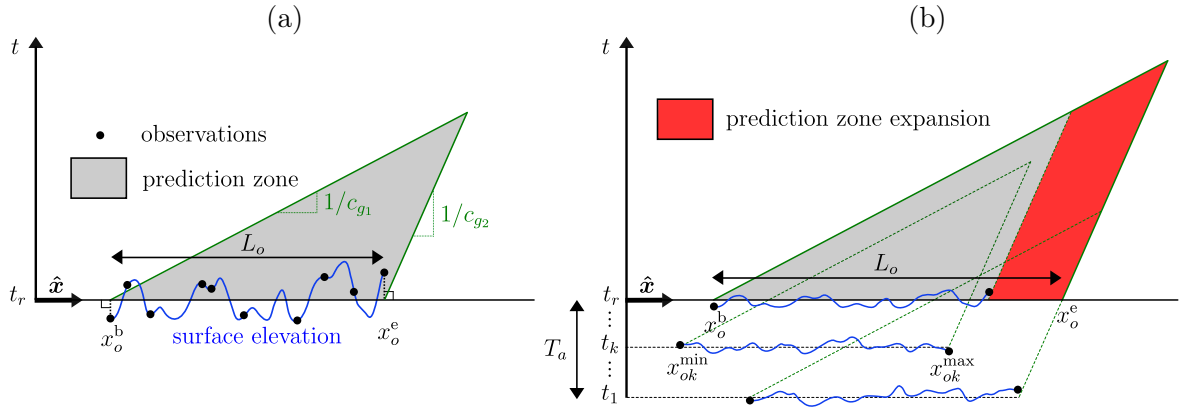

Figure III.8: Evolution of the wave prediction zone in time and space for the assimilation of: (a) spatial data and (b) spatio-temporal data (dash lines are prediction zones boundaries at time t_k). The increase in the prediction zone relative to that of spatial only observations is highlighted in red. c_{g1} and c_{g2} denote the fastest and slowest group velocities, respectively.

where c_{g1} and c_{g2} are the fastest and slowest group velocities, respectively, and x_o^{b} and x_o^{e} define the beginning and the end of $\mathcal{P}(t_r)$ as

$$\begin{cases} x_o^{\mathrm{b}} = \min\limits_k \left\{ x_{ok}^{\min} + c_{g1}\left(t_r - t_k\right) \right\}, \\ x_o^{\mathrm{e}} = \max\limits_k \left\{ x_{ok}^{\max} + c_{g2}\left(t_r - t_k\right) \right\}, \end{cases} \tag{III.8}$$

where index $k \in \{1, ..., K\}$ with K the number of observation times.

In order to avoid confusion, we clarify here the different terms related to the notions of observation, reconstruction and prediction:

- the **prediction zone** is defined by equations (III.7) and (III.8), and is denoted $\mathcal{P}(t)$ with $t \geq t_r$;
- the **observation zone** (or **measurement zone**) corresponds to the spatio-temporal region covered by the observations (or wave measurements);
- the **predicted solution** (or **prediction**) refers to the wave model solution included in the prediction zone;
- the **reconstructed solution** (or **reconstruction**) refers to the wave model solution included in the observation zone.

III.3.2.2 Extension to Short-Crested Waves

In case of short-crested waves, the wave field is decomposed in individual wave components propagating in direction $\theta \in \left[\theta^{\min}, \theta^{\max}\right]$. The corresponding prediction zone can be calculated for each direction using equations (III.7) and (III.8), as if it were a one-directional case, by replacing x_o by $d_o = \boldsymbol{r}_o \cdot \hat{\boldsymbol{k}} = x_o \cos\theta + y_o \sin\theta$, i.e., the distance along the considered direction. The intersection of each zone forms the two-dimensional prediction zone. An exhaustive and detailed formulation of the prediction zone boundaries is provided by Qi *et al.* (2018*b*). Here, the prediction zone area is approximated based on two extreme directions $\theta^{\min}$ and $\theta^{\max}$, as shown in figure III.9.

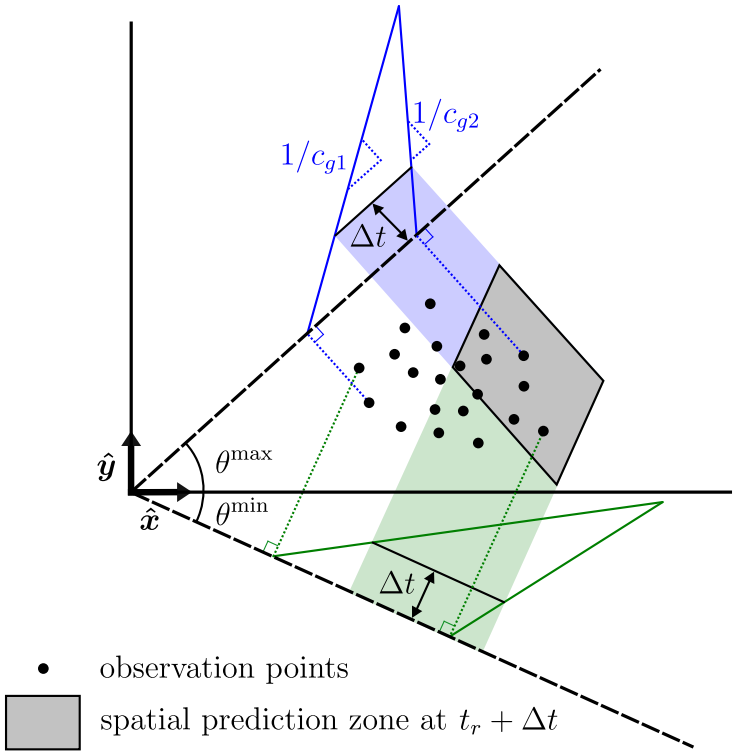

Figure III.9: Location of the prediction zone for a two-dimensional set of observations after a propagation time of $t = t_r + \Delta t$. It is approximated by the intersection of the one-dimensional prediction zones calculated at two extreme directions.

III.3.3 Bandwidths of the Modeled Wave Field

The dynamics of a wave field is dominated by the propagation of its most energetic components. Since ocean wave generation processes yield energy spectra with fairly narrow frequency bandwidth in their main direction of propagation, the dynamics of the wave field can be accurately represented by finite frequency and direction bandwidths having appropriate cutoff limits $\left(k^{\min,\max}, \theta^{\min,\max}\right)$. Appropriate definition of these limits is important for the assimilation procedure that parameterizes the wave model on the basis of prescribed values of wave frequencies and directions.

III.3.3.1 Limitations Due to the Observation Grid

As mentioned above, the spatio-temporal characteristics of the observation grid constrain the wave information that is accessible. The smallest wavenumber measurable in this grid $k^{\min} = 2\pi/L_o$ is function of the largest distance $L_o = x_o^{\mathrm{e}} - x_o^{\mathrm{b}}$ between two observation points at reconstruction time t_r (see figure III.8b). However, L_o itself depends on the minimum and maximum group velocities, thus on the cutoff frequencies. This paradox is resolved by conservatively calculating $k^{\min} = 2\pi/L_c$, with $L_c = x_o^{\mathrm{e}} - x_{o(k=K)}^{\min} \leq L_o$. L_c only depends on the spatio-temporal location of the observations and on the minimum group velocity, which is related to the high cutoff wavenumber $k^{\max}$.

When reconstructing a signal over a regular observation grid (i.e., with a constant spatial sampling), the maximum high cutoff frequency must satisfy Shannon's condition $k^{\max} \leq 2\pi/\left(2\ell_o\right)$ where ℓ_o is the distance between two observation points. Since the observation grid is highly irregular when using an optical method (section III.1.2.1) and the prediction error depends on the amount of energy included in the wave model (section III.3.1), $k^{\max}$ is set so that the spectral energy truncated at higher frequencies is negligible for the dynamic description of the wave

field (details in the next section). Also, since optical measurements provide observations at many spatial locations, there is no constraint on the cutoff directions from the observation grid. Hence, likewise $k^{\max}$, $\theta^{\min,\max}$ are calculated based on the truncation of a negligible amount of energy.

III.3.3.2 Wave Energy as Criterion

If an estimate of the underlying wave field is known, one way to choose these parameters is to evaluate their influence on the quality of the reconstructed or predicted wave field, since a convergence of the quality of the reconstruction/prediction is expected as $k^{\max}$ increases and $\left[\theta^{\min}, \theta^{\max}\right]$ gets larger.

In operational conditions, such information is not typically available and one must rely on an estimate of the wave spectrum $S(k,\theta)$, for instance based on measurements of the motion of the structure of interest, or from earlier measurements of free surface elevations. In this case, relevant values of $k^{\max}$ and $\theta^{\min,\max}$ can be selected by ensuring that

$$\int_{k^{\min}}^{k^{\max}} \int_{\theta^{\min}}^{\theta^{\max}} S(k,\theta)\,\mathrm{d}\theta\,\mathrm{d}k \leq (1-\mu) \int_{0}^{+\infty} \int_{0}^{2\pi} S(k,\theta)\,\mathrm{d}\theta\,\mathrm{d}k = (1-\mu)\, m_0, \tag{III.9}$$

where $\mu \ll 1$ is the fraction of total energy that can be considered as negligible on free surface dynamics.

III.3.4 Group Velocities for the Determination of the Prediction Zone

In practice, the low cutoff wavenumber $k^{\min}$ calculated for the wave model parameterization is not adapted for the determination of the evolution of the prediction zone. Indeed, $k^{\min}$ being defined with respect to L_c (section III.3.3.1), which is calculated from a particular set of observations, the comparison of predictions made from different observations would not be relevant since the prediction zone evolution would be based on a different value of $k^{\min}$, thus on different theoretical prediction errors, as explained in section III.3.1.

The determination of $c_{g1,2}$ that drive the evolution of $\mathcal{P}$ and limit its extent could then be based on the amount of neglected energy beyond the boundaries of the interval formed by the associated wavenumbers $[k_1, k_2]$, i.e., $\int_0^{k_1} E(k)\,\mathrm{d}k = \int_{k_2}^{+\infty} E(k)\,\mathrm{d}k \ll m_0$, where $E(k)$ is the wavenumber spectrum. However, the asymptotic behavior of the wave spectrum at large wavenumbers influences the estimate of the integral of $E(k)$ for $k \to +\infty$ in a way that leads the wave amplitude associated with k_1 to be larger than the one associated with k_2, i.e., $E(k_1) > E(k_2)$. Since one intended application of deterministic wave prediction is the motion prediction of floating structures (whose response amplitude operators generally do not filter out low-frequency excitations), long waves are important to take into account in the estimate of the prediction accuracy. Hence, it appears important to impose a constraint on the amplitude of waves associated to the cutoff wavenumbers.

Here, we choose to calculate the limiting wave amplitudes with respect to the peak of the wave spectrum. Accordingly, the wave group velocities $c_{g1,2}$ that govern the evolution of the prediction zone boundaries are defined based on two angular velocities $\omega_{1,2}$ such that

$$E(\omega_1) = E(\omega_2) = \mu E(\omega_p), \tag{III.10}$$

where $E(\omega) = E(k)/c_g$ is the surface elevation angular frequency spectrum, ω_p is the angular frequency based on its peak wavelength, i.e., that of maximum energy, and μ is a small fraction, here, of the peak spectral energy. ω_1 (ω_2) is the smallest (largest) angular frequency to respect condition (III.10). In the following, we will use the linear deep-water dispersion relationship to estimate the group velocities from $\omega_{1,2}$. An example is given in figure III.10 that shows the location of the angular frequencies $\omega_{1,2}$ with respect to the wave spectrum for $\mu = 0.05$.

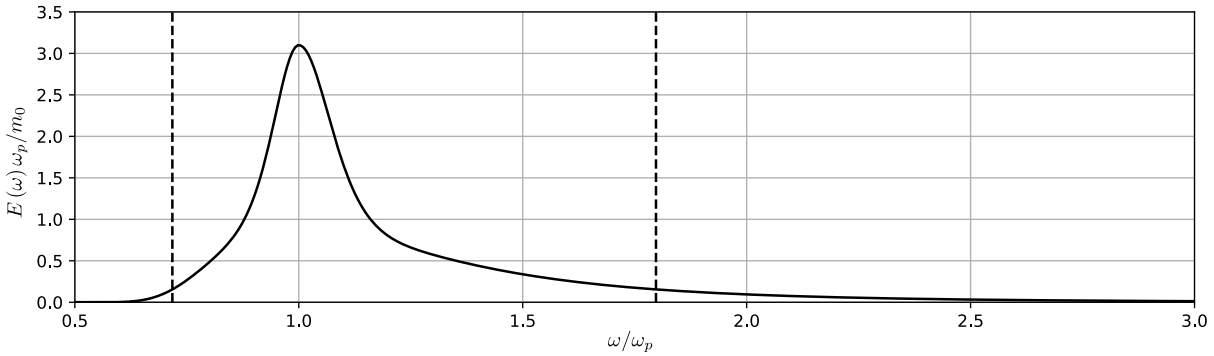

Figure III.10: Wave spectrum (——) bounded by angular frequencies $\omega_{1,2}$ (- - - -) calculated according to equation (III.10) with $\mu = 0.05$.

III.4 Measurement Zone Determination

In operations, a target observation zone has to be defined for the parametrization of the sensor's viewing and aperture angles. Then, the problem can be formulated as follows: for any wave field of spectrum $S(k,\theta)$, find the appropriate spatio-temporal observation zone that leads the target horizon of prediction $\mathcal{H}(t)$ to be included in the accessible prediction zone $\mathcal{P}(t)$. The target horizon of prediction is a spatio-temporal region within which the wave field has to be predicted for the purpose of the application. To solve this problem, the relationship between the observations and the accessible prediction zone described previously has to be inverted. With no more information than the horizon of prediction, however, this inversion problem is ill posed: many sets of observations can lead to a prediction zone that includes the same horizon of prediction (but with different prediction accuracy). The definition of the optimal inversion (i.e., observations that lead to the most accurate prediction) involves many sea state and measurement properties and could be the object of an independent study. Nevertheless, under specific constraining assumptions, it is possible to relate the horizon of prediction and the sensor's parameters, as we will see below.

In this section, we first investigate the implication of the extent of the measurement zone on the time/space location of the farthest predictable point in case of a unidirectional wave field. Then, in a restrictive configuration and still for unidirectional waves, we give the expression of the camera's viewing and aperture angles as a function of an arbitrary horizon of prediction.

III.4.1 Observation Footprint Versus High-Frequency Resolution

We saw in section III.3.1 that the extent of the prediction zone is strongly dependent on the desired value of prediction error. Here, we try to address the following question: for a prescribed theoretical prediction error (i.e., $1-P$, equation (III.6)), what setup allows to predict the farthest? Two strategies are used, one based on the truncation of low-frequency components by reducing the footprint of the observations, the other based on the truncation of high-frequency components. To do this, we consider a long-crested wave field and determine the value of the farthest point included in the prediction zone, that is the point $(x^{\max}, t^{\max})$ where $\mathcal{P}(t)$ ends (see figure III.8), as a function of the cutoff wavenumbers $k_{1,2}$ that are used to calculate P.

Formally, $x^{\max}$ and $t^{\max}$, respectively the farthest predictable location in the downstream direction and the latest predictable time, are calculated as

$$t^{\max} = L_o/(c_{g1} - c_{g2}) \quad \text{where} \quad L_o = 2\pi/k_1, \qquad \text{and} \qquad x^{\max} = t^{\max} c_{g2}. \tag{III.11}$$

Results are plotted in figure III.11, showing the evolution of P, $x^{\max}$, $t^{\max}$ and L_o as a function of k_1 with a fixed k_2 (case (1), dashed green line) or k_2 with a fixed k_1 (case (2), continuous blue line). In both cases (1) and (2), the value of the fixed wavenumber is taken sufficiently low or high (depending on the case) so the evaluation of P is considered independent on it. In

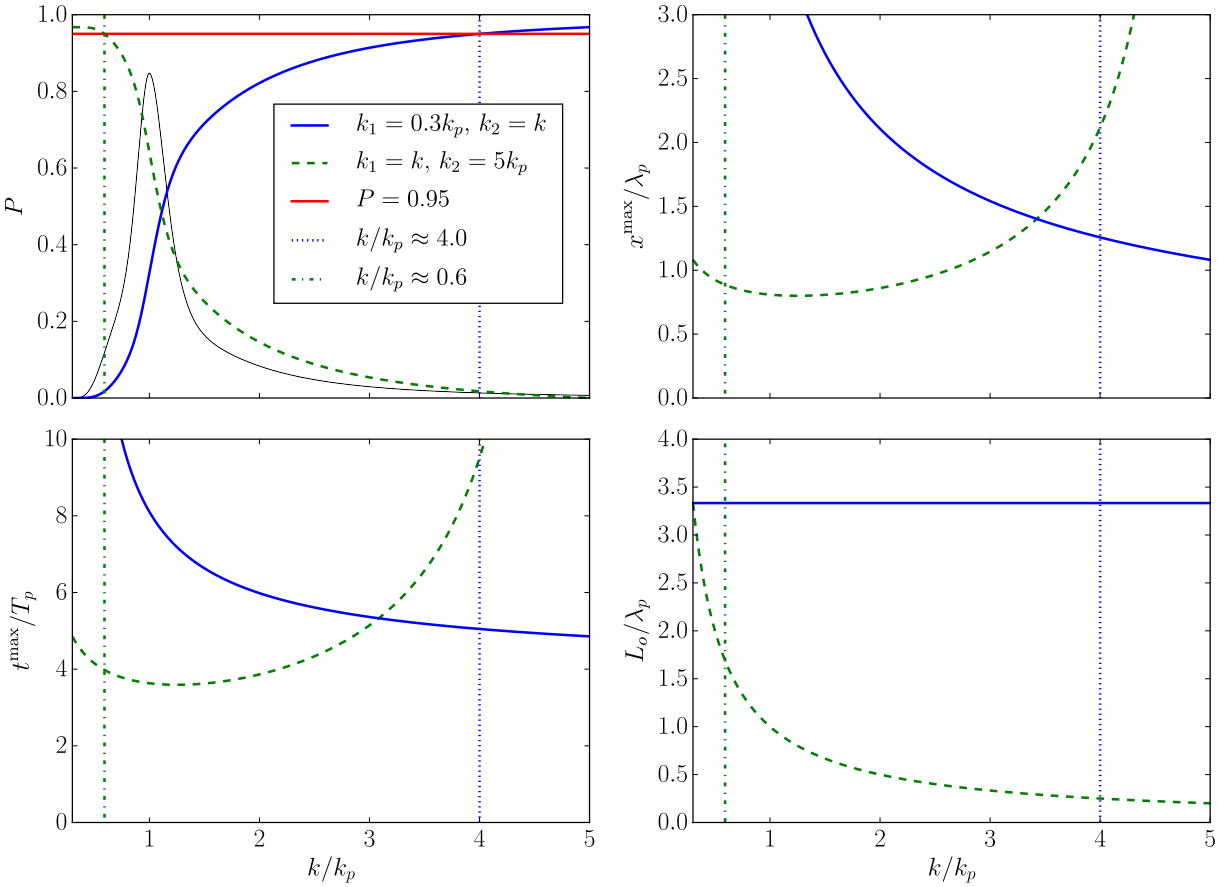

Figure III.11: Evolution of P (top left), $x^{\max}$ (top right), $t^{\max}$ (bottom left) and L_o (bottom right) as a function of k_1 with a fixed k_2 (case (1), ----) or k_2 with a fixed k_1 (case (2), ——). In both cases, the cutoff wavenumbers leading to a theoretical prediction error of 5% are marked (case (1): -·-·-; case (2): ·····).

practice, we take $k_2 = 5k_p$ in case (1) and $k_1 = 0.3k_p$ (leading to a realistic value of $L_o \approx 3.3\lambda_p$) in case (2). The wave spectrum (JONSWAP, $\gamma = 3.3$) is also given in the top-left plot as a visual help for interpretation. We see that the truncation of the low frequency by reducing the size of L_o do not directly lead to increase $x^{\max}$ and $t^{\max}$, which both decrease until $k_1 \gtrsim k_p$ and $P \lesssim 0.4$. It should then be preferred to truncate the spectrum at high frequencies which directly contributes to increase both $x^{\max}$ and $t^{\max}$. For example, for a theoretical prediction error of 5% (i.e., $P = 0.95$), $x^{\max}$ attains approximately $1.25\lambda_p$ and $t^{\max}$ is on the order of $5T_p$ in case (2). On the other hand, in case (1), $x^{\max}$ and $t^{\max}$ respectively drop approximately to $0.8\lambda_p$ and $4T_p$.

III.4.2 Derivation in a Simple Configuration

We detail here the method to determine the measurement zone that leads to the desired prediction zone. The problem being quite complex to formulate in the general case, we consider the simpler configuration of long-crested waves whose cutoff frequencies beyond which the wave energy is negligible are known.

The determination of the measurement zone starts from the horizon of prediction, defined as a spatio-temporal region that has to be included in the accessible prediction zone. It is defined by both a time interval $[t_a, t_b]$ and a space interval $x_c(t) \pm d$, in which d is a distance surrounding the mobile location $x_c(t)$ of the center of the mobile structure of interest. It yields $\mathcal{H}(t) = x_c(t) \pm d$ for $t \in [t_a, t_b]$. The relation between the accessible prediction zone and the measurement zone is not bijective: multiple sets of measurements can be used to get access to the desired prediction zone. In the presented example, we limit the observations to a fixed spatial zone during the

duration of the measurements and the assimilation time T_a is predefined.

To ensure that $\mathcal{H}(t)$ is included in the prediction zone, we determine the points $(x_1^{\mathrm{p}}, t_1^{\mathrm{p}})$ and $(x_2^{\mathrm{p}}, t_2^{\mathrm{p}})$ constraining the size of the prediction zone as shown in figure III.12 and whose expressions yield

$$\begin{aligned} t_1^{\mathrm{p}} &= \arg\min_t \left\{ x_c(t) - d + c_{g1}(t - t_a) \right\}, \\ t_2^{\mathrm{p}} &= \arg\min_t \left\{ x_c(t) + d + c_{g2}(t - t_a) \right\}, \end{aligned} \quad \text{for} \quad t \in [t_a, t_b], \quad \text{and} \quad \begin{aligned} x_1^{\mathrm{p}} &= x_c(t_1^{\mathrm{p}}) - d, \\ x_2^{\mathrm{p}} &= x_c(t_2^{\mathrm{p}}) + d. \end{aligned}$$

From these points, it is possible to determine the size and localization of the measurement zone using the group velocities corresponding to the cutoff frequencies, the assimilation time T_a and the time instant t_r at which the prediction is supposed to be issued. It follows

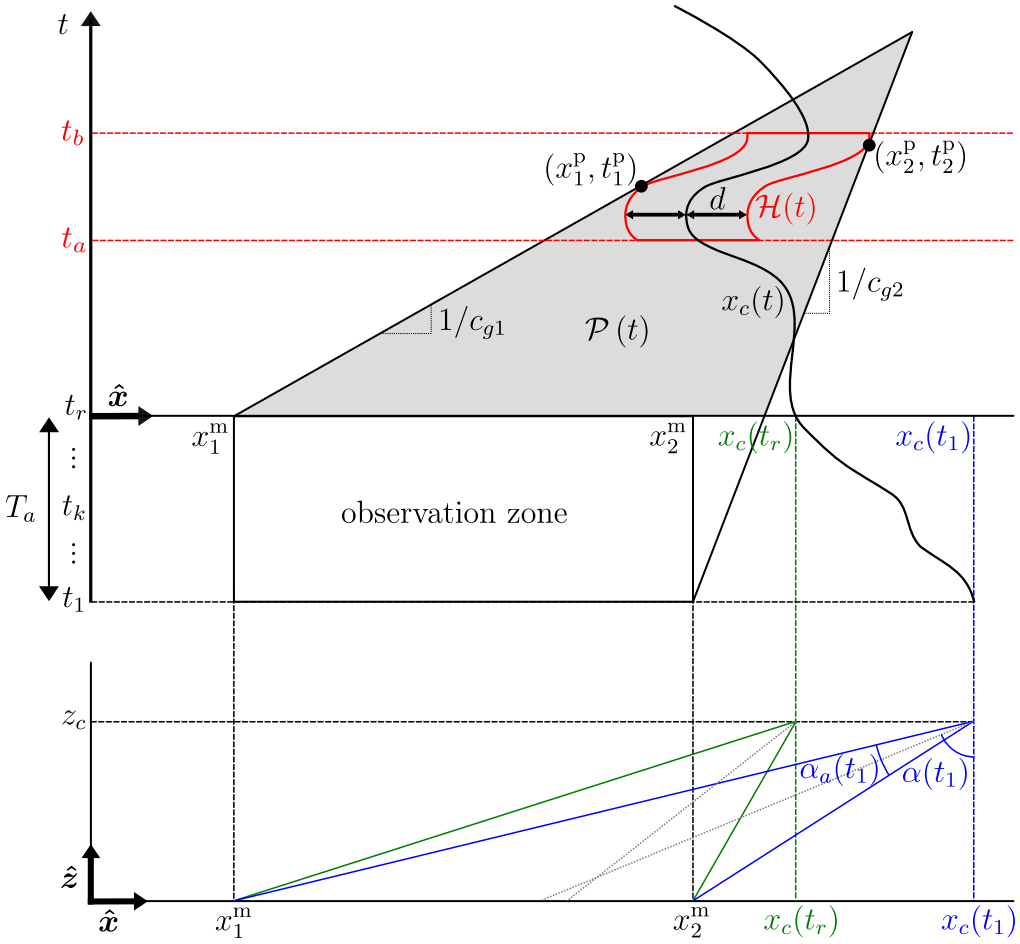

Figure III.12: Space/time diagram for the determination of the camera's parameters (α, α_a) from a prediction horizon $\mathcal{H}(t)$, an assimilation time T_a and a mobile structure trajectory $x_c(t)$.

$$\begin{aligned} x_1^{\mathrm{m}} &= x_1^{\mathrm{p}} - c_{g1}(t_1^{\mathrm{p}} - t_r), \\ x_2^{\mathrm{m}} &= x_2^{\mathrm{p}} - c_{g2}(t_2^{\mathrm{p}} - t_r + T_a). \end{aligned}$$

Finally, from these measurement zone characteristics, we calculate the parameters of the sensor as a function of time during T_a, as

$$\begin{aligned} \alpha(t) &= \frac{1}{2}\left[\tan^{-1}\left(\frac{x_1^{\mathrm{m}} - x_c(t)}{z_c} \right) + \tan^{-1}\left(\frac{x_2^{\mathrm{m}} - x_c(t)}{z_c} \right) \right], \\ \alpha_a(t) &= \tan^{-1}\left(\frac{x_2^{\mathrm{m}} - x_c(t)}{z_c} \right) - \tan^{-1}\left(\frac{x_1^{\mathrm{m}} - x_c(t)}{z_c} \right), \end{aligned} \quad \text{for} \quad t \in [t_r - T_a, t_r].$$

At any time during the measurements, we are able to unambiguously determine the values of the sensor parameters, being the viewing and aperture angles. To do this, we supposed that: the wave field is unidirectional, we know the cutoff frequencies that govern the evolution of the prediction zone, the assimilation time is predefined as well as the reconstruction instant, and the spatial extent of the measurements is set constant during the whole assimilation time. The above analysis is thus very limited, and issues related to the generalization of the solution are addressed in the next section.

Quantification for a practical case

In order to have in mind some orders of magnitude related to the time and space scales involved in the measurements, we consider in this paragraph the case of a ship that is 142 m long, navigating in a unidirectional JONSWAP sea state of peak period $T_p = 10$ s and peakedness parameter $\gamma = 3.3$. With similar assumptions to those in the configuration depicted above, we consider a horizon of prediction that is 200 m long and centered on the ship location (i.e., $d = 100$ m), and that spans a 10 s time, from 20 s to 30 s after the first measured waves enter the region (i.e., from $t_a = t_r + 2T_p$ to $t_b = t_r + 3T_p$). Assuming that we limit the assimilation time T_a to 10 s ($= T_p$), we calculate the distance over which the lidar has to measure surface elevations (i.e., $x_2^{\mathrm{m}} - x_1^{\mathrm{m}}$) for the horizon of prediction to be included in the accessible prediction zone, for two configurations that are with and without a ship forward speed. In both cases, the group velocities for the prediction zone calculation are determined for the wave spectrum using condition (III.10) with $\mu = 0.05$ (i.e., $c_{g_1} \approx 10.8$ m/s and $c_{g_2} \approx 4.3$ m/s).

For the first configuration in which the ship has no speed, we find a measurement distance of 395 m ($\approx 2.5\lambda_p$). For the second configuration in which the ship is moving in the opposite wave direction with a forward speed of 15 m/s, we find a measurement distance of 545 m ($\approx 3.5\lambda_p$). Considering a longer assimilation time of 20 s (i.e., $T_a = 2T_p$), we find 352 m ($\approx 2.3\lambda_p$) and 502 m ($\approx 3.2\lambda_p$) with and without the forward speed, respectively. In all cases, the measurement distance is on the same order of magnitude than the theoretical possibilities of lidar sensors, which should be able to measure over distances of several peak wavelengths when mounted on a ship (example of such configuration in figure III.4a).

III.4.3 Difficulties for Generalization

As mentioned above, the principal difficulty to provide a reliable method for the determination of the measurement zone in all configurations comes from the surjective relationship between the measurement zone and the resulting prediction zone. Beyond the restrictions dictated by physical description of the wave propagation processes (i.e., advection of the wave information at group velocity) and the technical limitations of the sensor (e.g., altitude, grazing incidence, resolution), it is necessary to formulate restrictions on the measurement strategy to close the problem and be able to find a solution. Such restrictions can be formulated based on the relations between some properties of the observations and how they condition the data assimilation procedure:

- The *size* in space/time of the observations determines the lowest measurable frequency. It is thus important to either cover a surface (parameter L_o in figure III.8) large enough to be equal to the longest wavelength of interest, or make sure that the measurements are made during a time (T_a) that is long enough to be equal to the longest wave period of interest;

- The *sampling rate* in space/time of the observations determine the highest measurable frequency. Nyquist theorem applies only in case of uniformly sampled data, which is not the case of remote optical measurements. However, it is possible to say that, at least, the minimal spatial (temporal) interval between two observations has to be lower than

half of the lowest wavelength (period) of interest. Instead of the minimal interval, global indicators that generalize the Nyquist theorem to non-uniformly sampled data can be used (e.g., Belmont, 1995);

- The *level of non-uniformity* of the observations has an impact on the performances of the assimilation process, thus influences the quality of the predictions. A limiting value of a measure of the degree of non-uniformity such as the one proposed by Belmont & Morris (1994) could serve as an additional restriction;
- The *density* in space/time of the observations determines the quantity of physics that can be assimilated for a specific size of observations, and the larger the assimilated quantity of physics, the more accurate the prediction;
- The *angle of incidence* of the sensor's beams with the ocean surface, beyond the creation shadowing effects and irregularity in the observations, increases the measurement error. Since the angle of incidence increases with the measurement distance from the sensor, it seems reasonable to impose a limitation of measurement distance.

The properties of the algorithm that is used to invert the wave model as well as the properties of the considered optical sensor have not been investigated enough to provide clear restrictions in every wave configuration. Moreover, the presented methodology is sensible to the complexity of the sea state underlying the observations. In case of multi-modal sea state in which the directional energy distribution is very spread out (e.g., superposition of two sea states with opposite main directions of propagation) the described approach requires to cover a zone that is large enough to measure every wave component that is supposed to interfere with the structure trajectory.

Conclusion

The distribution of lidar surface elevation measurements were characterized using a geometrical approach, highlighting the strong spatial non-uniformity. An assimilation procedure adapted to such data were developed to accurately parameterize the nonlinear wave model detailed in chapter 2, namely ICWM. We then explained the relationship between the accessible prediction zone and the spatio-temporal location of the wave measurements. The space/time distribution of the theoretical prediction error was derived from the amount of truncated energy by the accessible measured information, showing that the extent of the prediction zone strongly depends on the expected prediction accuracy. Criteria are presented for the selection of the frequency and direction bandwidths for the modeled wave field as well as for the prediction zone evolution, allowing to close the wave-model inverse problem and to define a clear region within which the prediction capabilities will be estimated. Finally, issues related to the determination of the optimal measurement zone were exposed and partly addressed.

From the characterization of remote optical free surface observations to the derivation of the ICWM inversion and the description of the accessible prediction region, this chapter completes the theoretical basis of the prediction method employed in this thesis.

Résumé du quatrième chapitre

Dans le quatrième chapitre, les propriétés et performances du modèle non-linéaire et de la méthode de prédiction développés sont étudiés à partir de jeux de données réalistes – bien que synthétiques et supposés sans erreur de mesure – de mesures optiques. Des simulations haute-fidélité de champs de vagues sont utilisés pour fournir des surfaces océaniques de références, à partir desquelles les observations sont générées. Après une description de la méthode de génération des mesures lidar synthétiques, nous étudions la dépendance de la précision des reconstructions/prédictions à divers paramètres d'assimilation, à savoir le nombre de composantes dans le modèle pour la représentation du champ de vagues, la fréquence de coupure haute de l'intervalle utilisé, la non-uniformité de la distribution des observations, la durée d'assimilation et l'étendue directionnelle des observations dans le cas d'un champ de vagues directionnel. Nous montrons que la précision de la reconstruction converge rapidement avec le nombre de composantes dans le modèle de vagues (quelques dizaines suffisent pour un champ de vagues unidirectionnel), ainsi qu'avec la fréquence de coupure haute. La caractéristique principale des mesures optiques, qui est de produire des données distribuées spatialement de manière très irrégulière, rend la reconstruction spatiale de champ de vagues particulièrement difficile. En revanche, inclure plusieurs jeux de mesures acquis à des instants différents dans l'inversion du modèle de vagues permet de contrer cette difficulté. Avec une procédure d'assimilation correctement paramétrée, l'effet de la prise en compte des propriétés non-linéaires de ICWM sur la qualité de la prédiction est analysé pour des champs de vagues de différentes cambrures caractéristiques. Il est montré que, malgré sa formulation analytique relativement simple, ICWM est capable de modéliser des effets non-linéaires menant à une amélioration non-négligeable de la prédiction. Comparée à celle de LWT, l'utilisation de ICWM dans l'algorithme de prédiction permet une réduction relative de l'erreur de prédiction de l'ordre de 25% pour la cambrure étudiée la plus élevée, soit $H_s/\lambda_p \approx 3.8\%$. Une large partie des résultats de ce chapitre est également présentée dans Desmars *et al.* (2018*a*,*b*).

Chapter IV

Numerical Investigation of the Prediction Algorithm

Introduction

In this chapter, we investigate the capabilities of the developed nonlinear wave model and prediction method from synthetic data. High-fidelity numerical simulations of wave fields are used to provide reference ocean surfaces, from which lidar-like observations are generated through a geometrical approach.

After a brief description of the method for the generation of synthetic lidar observations, we study the dependence of the proposed prediction algorithm, in terms of reconstruction/prediction accuracy, on some assimilation parameters, namely the number of wave components in the wave models for the wave field representation, the high cutoff boundary of the frequency bandwidth, the non-uniformity of the observations' distribution, the assimilation time, and the directional extent of spatial observations in case of short-crested waves. For a properly parameterized assimilation procedure, the impact of the inclusion of the nonlinear properties of ICWM on the prediction accuracy is then investigated for different characteristic wave steepnesses.

IV.1 Generation of Synthetic Lidar Measurements

In this section, we first go through the numerical settings of a High-Order Spectral (HOS) model to properly generate a realistic nonlinear surface, then we describe the process to calculate the locations of lidar rays intersections with that reference surface.

IV.1.1 Reference Surface Generation

The reference surface is generated by means of HOS simulations through the open-source solver HOS-ocean (Bonnefoy *et al.*, 2010; Ducrozet *et al.*, 2007, 2016), developed at ECN. The HOS method allows performing deterministic and completely nonlinear (according to the potential flow theory) simulations. It is thus adapted for the accurate and computationally efficient propagation of complex sea states, even in the realistic configuration of a directional irregular waves. A detailed description of the HOS framework is given in section I.3.4.1.

The numerical procedure to run HOS simulations is as follows. First, a linear surface is generated through an IFFT using a specified wave energy spectrum and random phases to obtain the initial elevation $\eta_{\mathrm{r}}(\boldsymbol{r}, t = 0)$ that is propagated in time. The size of the spatial domain is chosen so that it includes the spatial extent of both the observation and prediction zones. The spatial discretization (directly linked to the number of modes for the surface representation) is chosen high enough to ensure converged simulations. In the following, and according to previous numerical studies (e.g., Ducrozet *et al.*, 2016, 2012), we chose to have a number of nodes per peak wavelength of 32 along the main direction of propagation. In case of short-crested waves,

16 modes are used in the transverse direction. Periodic boundary conditions are used in order to run simulations over a long period of time without boundary influence. Directly propagating a linear surface with a nonlinear HOS procedure creates non-physical high-frequency standing waves that critically affect the validity of the solution. In order to prevent the creation such numerical perturbations, HOS-ocean uses an adjustment procedure developed by Dommermuth (2000) to allow nonlinearities to develop smoothly during a relaxation time, here fixed to 10 peak periods. A conservative time of 20 peak periods is taken before the surface is considered with nonlinearities fully developed and used for the generation of the observations. In the presented cases, the order M of the HOS expansion is set to 1 or 5, to generate a linear or a nonlinear reference surface, respectively. Simulations result in complex modal amplitudes that serve to directly reconstruct the free surface elevation on spatial grids. In all cases, the water depth is taken so that the deep water assumption applies.

IV.1.2 Surface-Rays Intersection Calculation

From the nonlinear surface previously described and a supposed set of lidar camera parameters (refer to section III.1.1 for details), we calculate the instantaneous intersections between lidar rays and the free surface. In order to determine the intersection between a specific ray and the surface, a purely geometrical approach is used: given a spatial set of free surface elevation values η_r, we look for the surface point that minimizes its distance d_r with the lidar ray trajectory. In directional waves, this distance is

$$d_r = ||\boldsymbol{u}_s \times \boldsymbol{u}_r||\,,$$

where $\boldsymbol{u}_s = (x_\mathrm{r} - x_c, y_\mathrm{r} - y_c, \eta_\mathrm{r} - z_c)$ is a vector defined by two points, the lidar location and the free surface point of interest, and $\boldsymbol{u}_r = (\sin\alpha_r \cos\beta_r, \sin\alpha_r \sin\beta_r, -\cos\alpha_r)$ is a unit ray vector. α_r and β_r are the vertical and horizontal ray angles and are swept within intervals $[\alpha - \alpha_a/2, \alpha + \alpha_a/2]$ and $[\beta - \beta_a/2, \beta + \beta_a/2]$ (where α_a and β_a are the lidar aperture angles, see figure III.1), respectively. All these parameters are depicted in figure IV.1. In case of

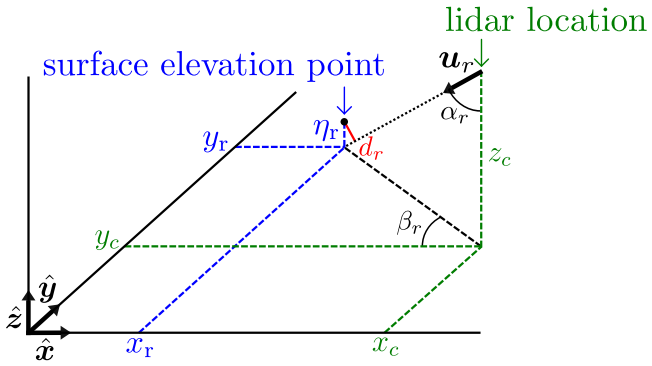

Figure IV.1: Parameters for the calculation of the distance between synthetic lidar rays and free surface elevation points.

unidirectional waves, this distance is calculated in a slightly different way:

$$d_r = ||-\boldsymbol{u}_s + (\boldsymbol{u}_s.\boldsymbol{u}_r)\,\boldsymbol{u}_r||\,,$$

in which $\boldsymbol{u}_s = (x_\mathrm{r} - x_c, \eta_\mathrm{r} - z_c)$ and $\boldsymbol{u}_r = (\sin\alpha_r, -\cos\alpha_r)$.

After a first calculation on a relatively rough grid of surface points, the calculation of intersection is done using a refined grid around the initially found location. This ensures that each ray leads to a unique observation, even in spatial zones of very dense measurements (i.e., close to the lidar). In addition, we make sure to select the first crossing point of the surface by the lidar ray, so that shadowing effects are accurately modeled. An example of a space series of observations obtained using this method for a unidirectional wave field is depicted in figure III.3.

IV.2 Sensitivity Analysis

In the following, we illustrate and discuss key characteristics of the proposed ocean wave prediction algorithm through its application to reference synthetic surface elevation datasets generated as explained in the previous section. For a unidirectional wave field, the influence of assimilation parameters on the agreement between the reconstruction/prediction and the reference data is studied, in particular, the number of wave components, the high cutoff frequency, the non-uniformity of the distribution of the observations, and the number of assimilated space series, while highlighting differences between linear and nonlinear prediction results. We then apply the algorithm to predict a directional wave field and study the influence of the horizontal aperture angle of the lidar camera.

IV.2.1 Synthetic Dataset Parameters

We use a JONSWAP spectrum with parameters $H_s = 3$ m and $T_p = 10$ s leading to the characteristic steepness $H_s/\lambda_p \approx 2\%$, and a peak enhancement factor $\gamma = 3.3$. In case of short-crested waves, the directional wave spectrum is obtained by multiplying equation (I.2) with the directional spreading function (I.3) (using the spreading parameter $\nu = \pi/4$ and main direction of propagation $\theta^{\text{dir}} = 0°$, that is along the x-direction). The synthetic lidar is fixed, located at an elevation $z_c = 30$ m above the mean surface level, observing the surface with viewing angles $(\alpha, \beta) = (76°, 0°)$ and aperture angles (α_a, β_a). α_a is fixed to 20°, while the influence of β_a is investigated in section IV.2.4.1.

Following the formalism introduced in section III.2, a number J of observations are created to form each spatial series, which has to be multiplied by a number K of observation times to get the total number of observations $L = J \times K$. The time sampling rate of the observations is constant and set to 1 Hz $(= 10 f_p)$.

IV.2.2 Prediction Error Definitions

The following normalized root mean square (rms) error is used to assess the prediction accuracy

$$\mathcal{E}^{\text{rms}}(t) = \frac{1}{N_s} \sum_{i=1}^{N_s} \left[\int_{\mathcal{P}(t)} \left(\eta_i(\boldsymbol{r}, t) - \eta_{\text{r}i}(\boldsymbol{r}, t) \right)^2 \mathrm{d}\boldsymbol{r} \Big/ \int_{\mathcal{P}(t)} \eta_{\text{r}i}^2 \, \mathrm{d}\boldsymbol{r} \right]^{1/2}. \tag{IV.1}$$

It quantifies the scaled mean squared difference between the reconstructed $(\eta(\boldsymbol{r}, t))$ and reference (η_{r}) surface elevations over the prediction zone $\mathcal{P}(t)$. Since the quality of the prediction depends on the measured local surface geometry at observation times, the misfit indicator is computed by averaging error values for N_s surface predictions. An unbiased estimate can only be obtained for a large number of samples from independent wave field realizations (i.e., of different set of random wave phases using the same initial wave spectrum) with, to the limit, $N_s \to \infty$. In the following numerical investigation, $N_s = 50$ realizations are generated (except where specified) with HOS-ocean, which allows the presented results to be converged.

In order to locally assess the spatial agreement between the predicted and reference surfaces, the following normalized misfit indicator is computed

$$\mathcal{E}(\boldsymbol{r}, t) = \frac{1}{N_s} \sum_{i=1}^{N_s} \left| \eta_i(\boldsymbol{r}, t) - \eta_{\text{r}i}(\boldsymbol{r}, t) \right| / H_s. \tag{IV.2}$$

IV.2.3 Prediction of Long-Crested Waves

The quantification of the influence of the model and reconstruction parameters is first done for long crested waves. The reconstruction is performed using N wave components whose wavenum-

bers are linearly distributed in an interval $\left[k^{\min}, k^{\max}\right]$. The low cutoff wavenumber is calculated from the spatial extent of the observation zone at the first observation time t_1, i.e., $k^{\min} = 2\pi/\left(x_{o1}^{\max} - x_{o1}^{\min}\right)$ (see figure III.8b). Both the number of wave components N and high cutoff wavenumber $k^{\max}$ are investigated and chosen with respect to the convergence of the reconstruction error, that is the rms prediction error at reconstruction time $\mathcal{E}^{\text{rms}}(t_r)$. In this section, the wavenumbers $k_{1,2}$ that are used to calculate the group velocities $c_{g1,2}$ that decides the prediction zone evolution are similar to the cutoff wavenumbers of the reconstructed wave field, i.e., $k_1 = k^{\min}$ and $k_2 = k^{\max}$.

IV.2.3.1 Convergence with Frequency Parameters

Linear long-crested wave fields are first generated and then reconstructed based on data acquired at a single time (i.e., $K = 1$), with $L = 64$ observation points uniformly distributed over the observation area. Figure IV.2 shows the prediction error $\mathcal{E}^{\text{rms}}$ at reconstruction time t_r, as a function of the number of wave components N used in the reconstructed wave field, for various high cutoff frequencies. Note that Shannon's criterion yields $k^{\max} \lessapprox 13.9k_p$, and the HOS

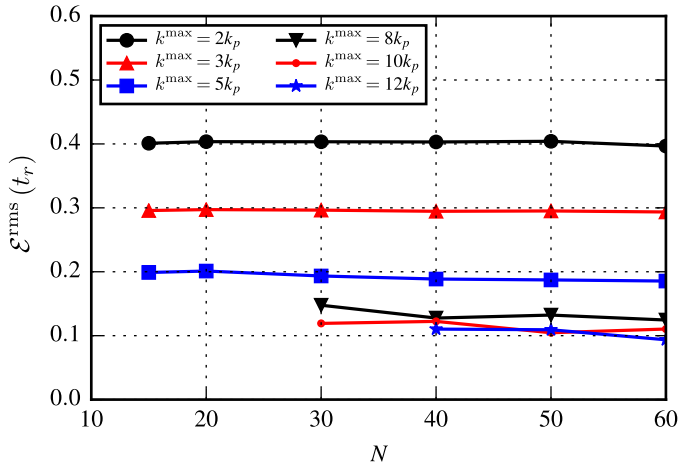

Figure IV.2: Reconstruction error of a linear surface as a function of the number of wave components, for different high cutoff frequencies, using a uniform observation grid for $K = 1$ (i.e., only spatial observations).

surfaces were generated using a cutoff frequency $k^{\max} = 16k_p$. As expected, results show that the reconstruction error decreases as the high cutoff frequency grows, and eventually converges to a minimum, here approximately 10%, for $k^{\max} = 10k_p$. In all cases, the number of wave components ($N = 15$ to 60) appears to have a very limited influence on the accuracy of the reconstruction. In the following we use $k^{\max} = 10k_p$ and $N = 50$.

IV.2.3.2 Modeling of Wave Nonlinearities

Next, nonlinear reference surfaces are generated using the same parameters and the rms error over the prediction zone is computed as a function of time $t = t_r$ to $t_r + 3T_p \approx 0.8t^{\max}$ (with $t^{\max}$ calculated according to equation (III.11)), for the same $L = 64$ uniformly distributed observation points, using either the LWT, CWM1, or ICWM model in the reconstruction (see appendix A.1 for details about CWM1 inversion).

For each model, the ocean surface is first reconstructed at t_r and then propagated to a later time when its prediction error (IV.1) with respect to the reference HOS solution is evaluated. For comparison, the case of figure IV.2 propagated in time is also marked on the figure (i.e., prediction using the LWT of a linear reference surface). Figure IV.3 shows that at $t = t_r$ all models give the approximately same results, whereas, as time increases, each wave model behaves

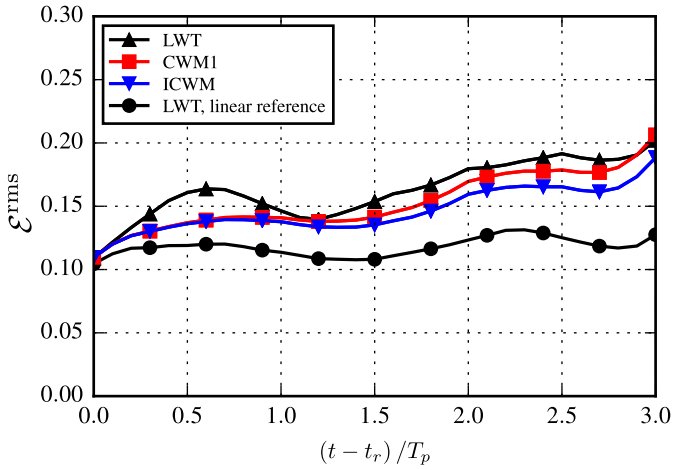

Figure IV.3: Prediction error of a nonlinear ocean surface as a function of time, for different wave models, using a uniform observation grid for $K = 1$ (i.e., only spatial observations).

according to its own properties. While the CWM1 model provides a better prediction than the LWT model, the ICWM further improves the reconstruction accuracy as time increases, due to the additional nonlinear phase velocity correction. Here, the improvement of the nonlinear over the linear prediction is on the order of a few percent. The linear prediction of a linear reference surface, shown on the same figure, is an indication of the best reconstruction accuracy that could be achieved for this problem and dataset, if all nonlinear effects were properly accounted for in the nonlinear wave models. Hence, the difference between this linear and nonlinear forecasts using ICWM increases with time due to the incomplete representation of time-dependent nonlinear effects in the latter model, to approximately reach 5% after $3T_p$ of propagation.

IV.2.3.3 Non-Uniformity of Observations and Inclusion of Time Information

In this section, the observations are no longer uniformly spread over the lidar footprint, but follows the sampling procedure explained in section IV.1 that leads to non-uniform spatial distributions. First, considering a single observation time (i.e., $K = 1$), figure IV.4 shows that the prediction error greatly increases (by more than a factor of 2) as compared to a uniform dataset. As suggested in earlier work (Nouguier *et al.*, 2014), this deterioration of the ocean reconstruc-

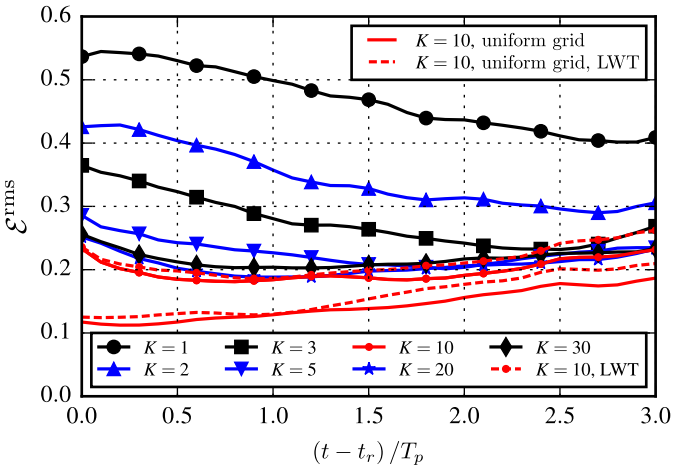

Figure IV.4: Prediction error of a nonlinear ocean surface as a function of time, for different numbers of observation times K, using ICWM over a non-uniform grid of observation points. Some LWT and/or uniform grid results are plotted for reference.

tion and prediction can be significantly reduced by increasing the number of observation times.

Indeed, this has the effect of incorporating some surface kinematic properties (e.g., wave velocity) into the observation dataset, helping the wave model parameterization for an accurate wave field representation. Thus, for $K = 2$, figure IV.4 shows that the error already drops by 10% as compared to $K = 1$, and then further drops when increasing K to reach a minimum of $\sim 20\%$ for $K = 10$. The latter, however, is still $\sim 5\%$ larger than the error achieved when using uniform observations, for the same number of observation times (also shown in the figure). Comparing results in figures IV.3 and IV.4 we also see that, when using a uniform grid of observation points, increasing K has no significant effect on the prediction error. For both the LWT and ICWM models, maximum errors indeed remain on the order of 15 to 20% for $K = 1$ and 10, even if ICWM leads to a consistently better prediction.

Figure IV.4 also shows that the error decreases when K increases form 1 to 10, but then gradually increases for larger K values. This is likely to result from increasing numerical errors in the inverse problem solution, as the observed dataset becomes larger. The number of wave components, here fixed to 50, might be the reason of the progressive increase of the prediction error for $K = 20$ and 30: the larger the assimilation time, the larger the need for wave components to accurately fit the data. Increasing the number of wave components would have lead to continue the convergence of the prediction error to a minimal value, as it is shown later in section V.3.2.

In general, increasing the assimilation time T_a results both in a decrease of the prediction error and in an increase of the size of the prediction zone. Both effects are illustrated in figure IV.5, which shows the reference and reconstructed surface elevations, for $K = 1$ and 30, at reconstruction time t_r ($= t_1$ for $K = 1$), together with the observation points at this specific instant. With $K = 30$, the reconstructed surface is much closer to the reference data than for

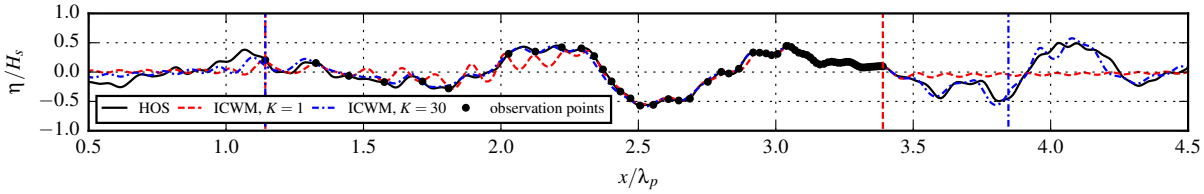

Figure IV.5: Elevation of the reference (HOS) and reconstructed (ICWM) surfaces, for different numbers of observation times. Vertical lines define the spatial limits of the accessible prediction zones.

$K = 1$, and the prediction zone is larger. The figure shows that, for $K = 30$, the reconstructed surface remains quite close to the reference data even beyond the prediction zone ($x/\lambda_p \gtrsim 3.8$). This suggests that the estimate for the upper limit of the prediction zone, which is calculated based on the distance traveled by the slowest wave component during the assimilation time, is conservative (equation (III.7)). As explained in section III.3.1 that presents a theoretical approach to determine the prediction error, since longer waves travel faster, the prediction of low frequency components can still be accurate beyond the upper limit (here, defined with respect to $k^{\max}$). A similar effect affects the high frequency components at distances below the lower limit, when time increases ($t > t_r$). This validates the presence of a "gray area" surrounding the prediction zone and growing with time, over which the amount of measured information could still be enough for an accurate prediction, but with decreasing accuracy as distance increases from the prediction zone boundary.

In figure IV.5, when the density of observation points decreases ($K = 1$, $1.2 \lesssim x/\lambda_p \lesssim 2.3$), the reconstructed surface tends not to fit well the reference surface elevation. As Shannon's condition is not met over this part of the observation zone, an aliasing phenomenon leads to overestimating high frequency component amplitudes, which causes the observed high frequency surface oscillations. However, since this poorly reconstructed part of the observation zone is

located close to the lower limit of the prediction zone, and the inaccuracy only affects high frequency components, the growing "gray area" effect tends to remove that part of the surface from the prediction zone, while the other part remains within it. This explains the decreasing prediction error as a function of time for small K values observed in figure IV.4. As indicated before, this phenomenon disappears with several observation times, due to an incorporation of wave kinematic properties into the observations: wave propagation's effects on the free surface elevation are intrinsically captured in observations made at different times.

These effects are further shown in figure IV.6 which presents the spatio-temporal evolution of the misfit $\mathcal{E}$ in case of the prediction of a linear surface. As explained above, for $K = 1$, the

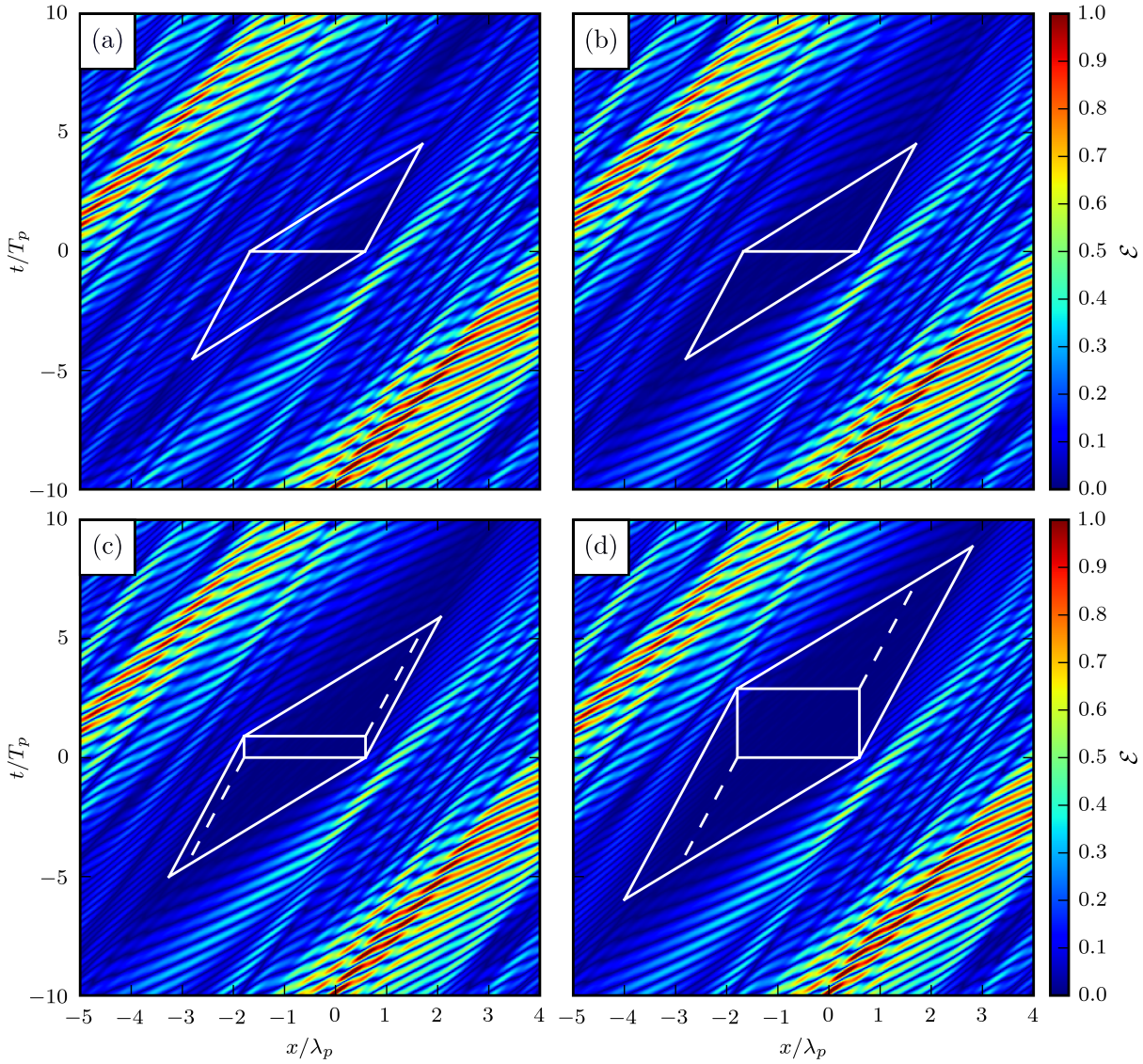

Figure IV.6: Spatio-temporal evolution of the misfit $\mathcal{E}$ ($N_s = 1$) for $K = 1$ in both cases of (a) non-uniform and (b) uniform distributions of observations, and for (c) $K = 10$ and (d) $K = 30$ in case of non-uniform observations. White lines define the theoretical boundaries of the spatio-temporal region of low error. The reference surface is linear and waves propagate in the positive x-direction. Observation zones correspond to horizontal lines in case of $K = 1$ (a,b) or rectangles for $K = 10$ and 30 (c,d).

prediction inaccuracies are mainly located where the measurement density is poor (i.e., for low x values along the horizontal line, figure IV.6a), allowing for high-frequency perturbations, which are progressively removed from the prediction zone as time increases. This does not happen in

case of uniform observations, for which respecting Shannon's condition prevents aliasing (figure IV.6b). The consequences of increasing K are depicted in figures IV.6c,d, showing the results for $K = 10$ and $K = 30$ in case of non-uniform observations. They consist in removing the initial high-frequency perturbations and increasing the size of the prediction zone. Note that, similar to the prediction zone, the boundaries of a "retrospection zone" is also plotted in figure IV.6 for $t < 0$, showing the extent of the region within which assimilated wave information are enough to accurately estimate the wave field hindcast.

IV.2.4 Prediction of Short-Crested Waves

We now generate a nonlinear short-crested wave field with HOS-ocean, for the same input JONSWAP spectrum as before, multiplied by the directional spreading function (I.3) (recalling $\nu = \pi/4$ and $\theta^{\text{dir}} = 0°$) to obtain the directional spectrum depicted in figure IV.7. The

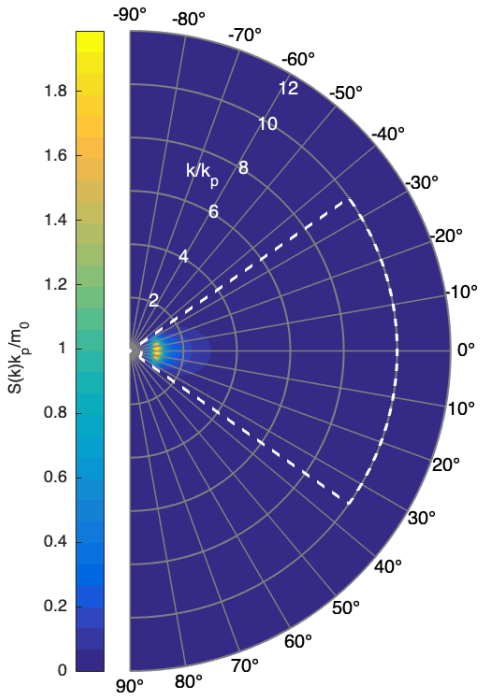

Figure IV.7: Directional wavenumber spectrum of the generated short-crested wave field. The dashed lines mark the spectral boundaries used in the wave model to perform the predictions.

wave field reconstruction is done using an observation grid of 64×64 spatial points ($J = 4096$) with $k^{\max} = 10k_p$, $K = 10$, and 50 wave frequencies distributed over 30 directions, yielding $N = 1500$ wave components. Lidar parameters detailed in section IV.2.1 are used. The low cutoff wave number $k^{\min}$ is similarly defined as for long-crested waves, using the spatial extent of the measurements at the first observation time along the main wave direction θ^{dir}. The limits of the direction bandwidth $\left[\theta^{\min}, \theta^{\max}\right]$ of the reconstructed wave field are chosen with respect to a level of energy in the wave spectrum, and in a way that the direction bandwidth is centered on θ^{dir}. The criterion defined by equation (III.9) is used to truncate the directional wave spectrum by an amount corresponding to 1% of the energy contained between $k^{\min}$ and $k^{\max}$, leading to $\theta^{\min,\max} \approx \pm 35°$. These frequency and directional boundaries are marked on top of the directional wave spectrum in figure IV.7.

Since the number of spatial observations is significantly larger than before, calculating the exact intersection of the L optical rays with time-varying reference surfaces becomes numerically demanding. Thus, we estimate the horizontal location of observation points $\boldsymbol{r}_\ell$ as if the reference

surface were plane, neglecting shadowing effects. Results are comparable since the major contribution to the spatial spreading of observation points, i.e., the increasingly grazing incidence of lidar rays at large distances, is accounted for. Also, the prediction error is now calculated for one surface and not averaged over 50 surface realizations as before. However, even if the uncertainty related to the calculation of the prediction error is complex to quantify, the estimate of the mean rms error $\mathcal{E}^{\text{rms}}$ is less dependent on the number of surface realizations for short-crested waves than for long-crested waves, since the integral term in equation IV.1 is evaluated over an extended (i.e., two-dimensional) spatial region $\mathcal{P}(t)$. Finally, the results presented here are limited to the linear wave prediction (although the reference oceanic surface is nonlinear), but the comments pertaining to the two following sections deal with aspects that should be independent of the wave model used.

IV.2.4.1 Directional Spreading of the Observation Datasets

In the following, we estimate the effect of the directional spreading of the observation datasets on the ocean reconstruction, which is function of the horizontal aperture angle β_a of the lidar (figure III.1). Figure IV.8 shows the rms prediction error calculated as a function of time over the two-dimensional prediction zone (see section III.3.2.2 for definition). With $\beta_a = 10°$, the error is

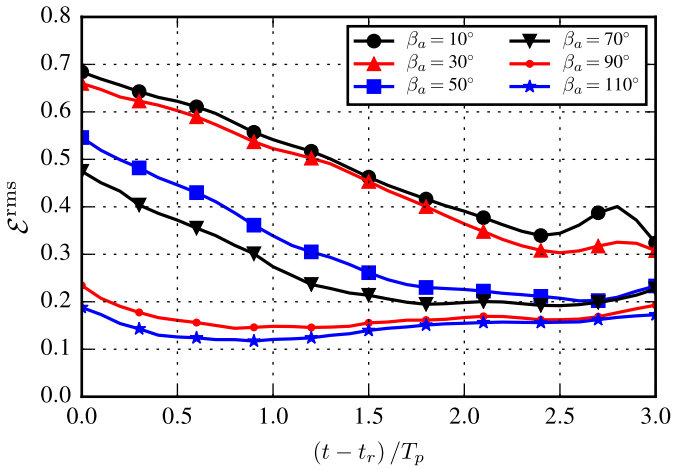

Figure IV.8: Prediction error of a nonlinear ocean surface ($N_s = 1$) as a function of time, using LWT for different horizontal aperture angles.

$\sim 70\%$ at reconstruction time and almost monotonically decreases as time increases up to $3T_p$ of propagation. Similar to long-crested waves, this decrease in error results from the growing "gray area" surrounding the prediction zone boundary. Poorly reconstructed parts of the ocean surface close to the initial prediction zone boundary are gradually removed as time increases. However, with the decreasing size of the prediction zone, the error estimate becomes more subject to slight variations, depending on which waves are crossing the prediction zone at the considered time, which explains the increasing error at $t = t_r + 2.5T_p$. Such effects did not occur in the long-crested case since errors were averaged over 50 different surface realizations. Figure IV.8 further shows that increasing β_a results in a significantly reduced error, which converges to a minimum on the order of 15% for $\beta_a = 110°$, which is on the same order of magnitude as long-crested cases. This large error reduction is related to the area of the ocean surface covered by the observation zone (i.e., lidar footprint), as compared to the area of the prediction zone at reconstruction time $\mathcal{P}(t_r)$. As shown in figure III.9, the two-dimensional prediction zone depends on the range of directions of propagation of wave components in the reconstructed wave field. Here, as $\theta^{\text{min,max}} \approx \pm 35°$ and the lidar is facing the main wave direction, the optimal aperture angle that leads to a maximal prediction zone coverage is $180 - \left(|\theta^{\text{max}}| + \left|\theta^{\text{min}}\right|\right) \approx 110°$.

IV.2.4.2 Spatial Distribution of the Error

Figure IV.9 shows the reference (HOS-ocean) and predicted (LWT) surface elevation for the previous case with $\beta_a = 110°$ after $2T_p$ of propagation. Within the prediction zone, whose

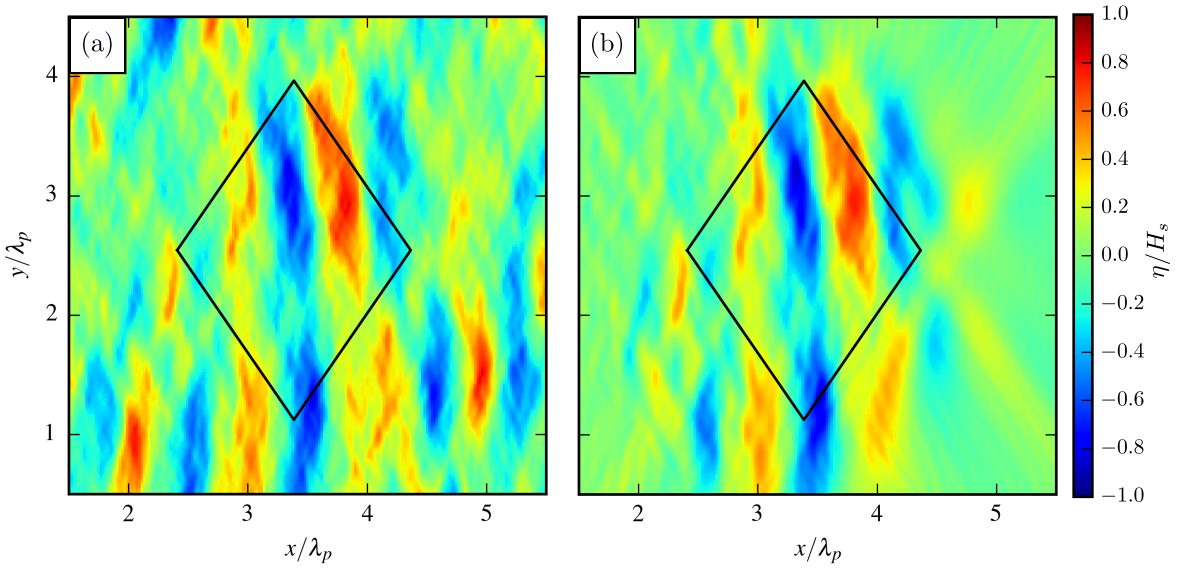

Figure IV.9: Elevation of (a) the reference surface η_r (HOS-ocean) and (b) the predicted surface η (LWT) after $2T_p$ of propagation. Black lines mark the spatial boundaries of the prediction zone.

boundaries are depicted by the straight lines, the agreement between the two surfaces is visually good. In addition, due to wave dispersion effects, we notice that, further away the prediction zone boundaries in the positive x-direction (i.e., main direction of propagation of the wave field), the predicted surface is dominated by low-frequency components (long waves). In contrast, in the spatial region located further away the prediction zone in the opposite direction, the predicted surface elevation is dominated by high-frequency components (short waves). Implications on the distribution of the prediction error are addressed below.

For the same case, figure IV.10 shows the horizontal distribution of the misfit $\mathcal{E}$ at reconstruction time and after two peak periods of propagation. As expected, the misfit is clearly much smaller within the prediction zone than outside of it. At both selected times, the "gray area" with reduced local error, which surrounds the prediction zone boundary, can clearly be identified. At reconstruction time, the "gray area" is close to the upper limit of the prediction zone (right side of the boundary in figure IV.10a) because the reconstructed longer waves propagate faster, and thus reach beyond this limit during assimilation time. After $2T_p$, the "gray area" has expanded and is now completely surrounding the prediction zone. This results in a different spatial distribution of the prediction error: the right side of the "gray area" is dominated by perturbations pertaining to short waves, while they pertain to long waves on the left side.

The presence of a growing low-error area surrounding the prediction zone confirms that the method used here for estimating its boundary, i.e., considering $k_1 = k^{\min}$ and $k_2 = k^{\max}$ for the determination of the group velocities $c_{g1,2}$, is too conservative. As a result, for future prediction applications, we choose to use the criterion based on the spectral amplitude of limiting wave components (described in section III.3.4) for the determination of the limiting group velocities $c_{g1,2}$. As it will be seen later, this criterion leads the theoretical prediction zone boundaries to better agree with the evolution of the actual region of low prediction error, using both numerical and experimental input data.

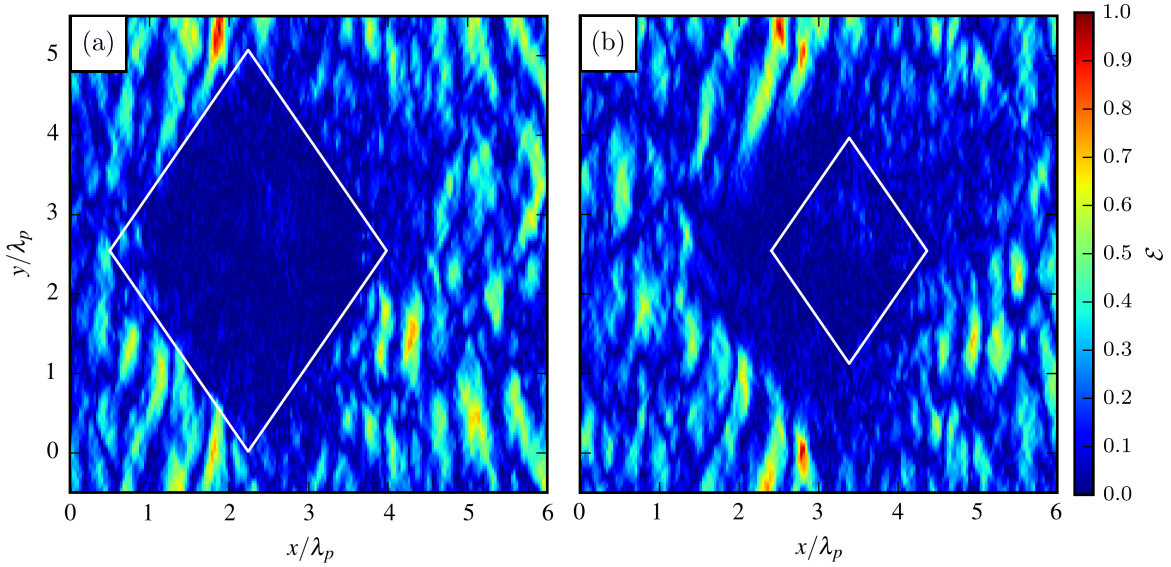

Figure IV.10: Spatial evolution of the misfit $\mathcal{E}$ ($N_s = 1$) at (a) reconstruction time and (b) after $2T_p$ of propagation. White lines mark the spatial boundaries of the prediction zone.

IV.3 Influence of Wave Steepness

In this application, we focus on the underlying physics of the sea state and its complexity which is characterized, among other parameters, by its wave steepness. Similar to the previous section, one-dimensional synthetic optical data is geometrically created for irregular sea states described by their wave energy spectrum and generated using a nonlinear wave model (HOS-ocean). The ocean surface is then reconstructed and predicted based on this dataset, using LWT and ICWM algorithms, and compared to the reference HOS surface. This is repeated for sea states of various characteristic wave steepnesses.

IV.3.1 Setup and Error Definitions

Sea state parameters and optical sensor configuration are similar to those used in the previous section, except the significant wave height H_s that varies between 1.5 and 6 m, leading the characteristic wave steepness to vary from $H_s/\lambda_p \approx 1\%$ to 3.8% (i.e., close to the wave breaking limit). Converged results were achieved using $N = 50$ frequency components, a high cutoff wavenumber $k^{\max} = 10k_p$, and a number of observation times $K = 10$ with a sampling rate of 1 Hz. Here, the low cutoff frequency is calculated according the theoretical lowest measurable wavelength, i.e., $k^{\min} = 2\pi/L_o$ (see details in section III.3.3.1). Moreover, in operational conditions, the predictions are only desired over a specific horizon of prediction corresponding to the space and time scales of the operation. Accordingly, we choose here to evaluate the prediction error over a spatio-temporal target zone $\mathcal{T}(t)$ that is 100 m long, centered on the lidar location x_c, and spans a 10 s time, from 20 s to 30 s after the first measured waves enter the region. Hence, a point (x, t) is in the target zone if $x_c - 50 \leq x \leq x_c + 50$ and $t_r + 20 \leq t \leq t_r + 30$.

The two indicators described in section IV.2.2 are used, except that the integration for the rms prediction error (IV.1) is now performed over the target zone $\mathcal{T}(t)$ instead of the entire prediction zone $\mathcal{P}(t)$, and the resulting error is averaged over 1000 surface realizations. Compared to 50, this high number of realizations allows reducing the uncertainty related to the estimate of the mean prediction misfits by an factor $\sqrt{50/1000} \approx 4.5$ (according to the central limit theorem). This is particularly important for the evaluation of the time evolution of the mean values of $\mathcal{E}(x,t)$ presented in figure IV.11b, which are more sensible to each particular realization than $\mathcal{E}^{\text{rms}}(t)$ that is already averaged over a spatial region.

IV.3.2 Prediction Results

Figure IV.11a shows $\mathcal{E}$ computed over a large spatio-temporal domain using ICWM, for a specific realization with $H_s = 3$ m. Boundaries of the prediction zone are calculated according to

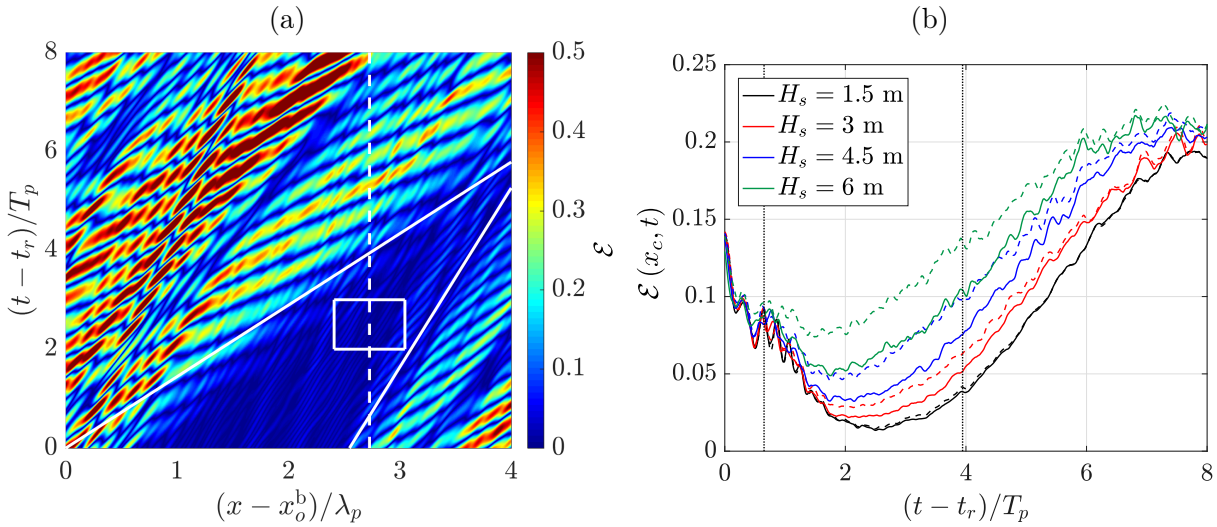

Figure IV.11: Numerical error of ocean prediction: (a) $\mathcal{E}$ ($N_s = 1$) using ICWM with $H_s = 3$ m; white box marks the spatio-temporal target zone $\mathcal{T}(t)$, within the prediction zone $\mathcal{P}(t)$ marked by oblique lines; dashed line at $x = x_c$ marks sensor location; (b) $\mathcal{E}$ ($N_s = 1000$) at $x = x_c$, as a function of time, for different steepnesses using LWT (- - - -) or ICWM (——); vertical dotted lines mark the prediction zone boundaries.

equation (III.7) and group velocities $c_{g1,2}$, based on equation (III.10) for $\omega_{1,2}$, with $\mu = 5\%$. For this realization, the prediction misfit is very low in the spatio-temporal target zone $\mathcal{T}(t)$, within the prediction zone $\mathcal{P}(t)$. Note that the error is higher at $t = t_r$ for low x-values than in the rest of the prediction zone. As explained in section IV.2.3.3, this is due to the poor spatial sampling in this region which induces high-frequency perturbations in the reconstructed surface. Since the badly resolved high-frequency components propagate slower than the rest of the wave components, they progressively leave the prediction zone as time increases. Figure IV.11b shows the time evolution of the same error averaged over 1,000 realizations at the sensor location x_c, for 4 different characteristic steepnesses (i.e., significant wave heights), using LWT and ICWM. For each steepness, the error decreases from a relatively large value at $t = t_r$, as waves enter the prediction zone (located between the vertical dotted lines in the figure) where it reaches its lowest value, and then increases to reach ~ 0.2 after $8T_p$ of propagation. Within the prediction zone, ICWM does reduce $\mathcal{E}$ in all cases as compared to LWT, but most significantly for the largest steepness. Also, probably due to nonlinear phase shifts for steep waves, the location of the minimum error shifts towards lower times as steepness increases.

Figure IV.12a further quantifies the prediction algorithm performance over the target zone (box in figure IV.11a), by plotting the time evolution of $\mathcal{E}^{\text{rms}}$ integrated within it, for LWT and ICWM. As observed earlier, this error significantly reduces using ICWM for average to high steepness. Moreover, $\mathcal{E}^{\text{rms}}$ only slowly varies with time and, hence, its time-average $\langle \mathcal{E}^{\text{rms}} \rangle$ shown in figure IV.12b can provide a global estimate of the prediction model performance, as a function of the characteristic steepness H_s/λ_p. In all cases this average error reduces using ICWM rather than LWT, with the largest reduction, 13% (from 0.52 to 0.39), occurring for the steepest wave field. This confirms the importance of accurately representing nonlinear effects, with a model such as ICWM, when performing an ocean surface prediction based on observations, and in particular wave asymmetry and phase shifts, which increase as the sea state becomes more severe.

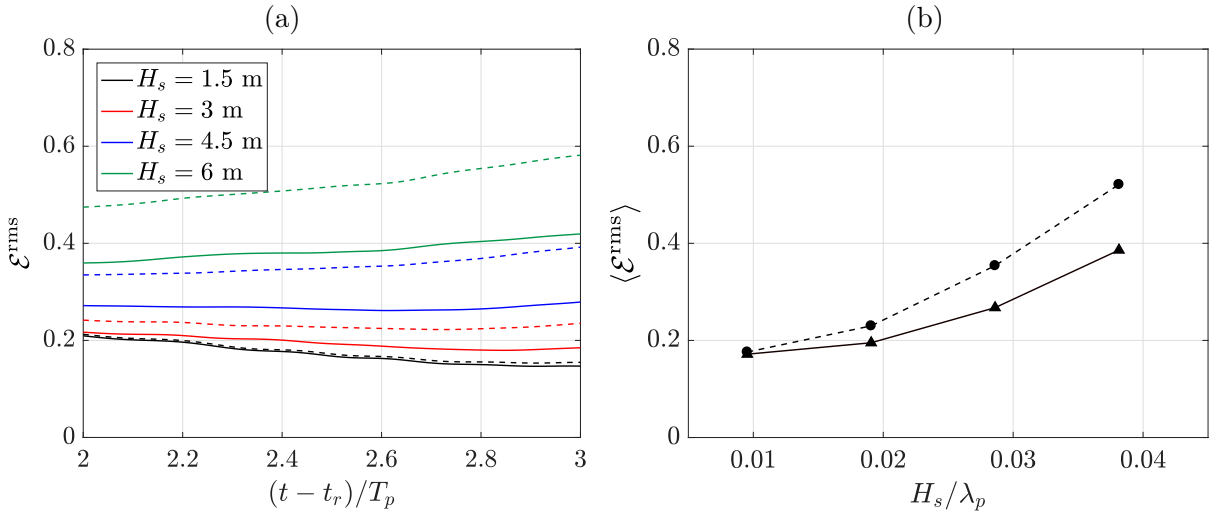

Figure IV.12: Numerical error of ocean reconstruction $\mathcal{E}^{\mathrm{rms}}$: (a) instantaneous; and (b) time-averaged over the target zone, for different steepnesses, using LWT (- - - -) or ICWM (——).

Conclusion

In this chapter, we characterized the proposed ocean surface elevation prediction method and showed that both long- and short-crested waves can be accurately predicted on the basis of non-uniform datasets of realistic – but synthetic and error free – optical measurements. Algorithms based on linear and various nonlinear wave models were applied and validated. We quantified the influence of several wave model parameters which are shown to be of significant importance to the wave field reconstruction. We found that the reconstruction accuracy converges quickly with the number of wave components (a few dozens are enough for long-crested waves), as well as with the high cutoff frequency, for which converged predictions are found for $k^{\mathrm{max}} = 10k_p$. We also quantified effects of the main characteristic of optical measurements, which is to yield highly sparsely distributed datasets, making for a particularly challenging spatial reconstruction. However, we showed that using spatio-temporal information allows overcoming this issue.

Despite its simple analytical formulation, ICWM is able to model relevant nonlinear effects in ocean wave fields that impact the accuracy of the wave prediction. The investigation of the prediction performance with respect to the wave steepness showed that wave predictions were significantly improved using ICWM, as compared to LWT, for moderate to relatively high steepness of approximately 3.8%. The larger the wave steepness, the higher the relative improvement, reaching a relative error reduction of 25% for the largest steepness.

In order to further assess the capabilities of the proposed prediction method in configurations that are closer to the real operational conditions, predictions based on experimental data, supplemented with digital twin simulations, are investigated in the next chapter.

Résumé du cinquième chapitre

Dans le cinquième et dernier chapitre, nous appliquons notre algorithme de prédiction à des données expérimentales issues d'une campagne menée dans le bassin océanique de l'École Centrale de Nantes. Des champs de vagues de différentes cambrures caractéristiques et/ou différents étalements directionnels ont été générés physiquement en bassin de houle, ainsi qu'à l'aide d'un bassin numérique basé sur le formalisme HOS, fournissant un jumeau numérique ("digital twin") du dispositif expérimental. Les sondes (physiques ou numériques, mesurant des séries temporelles d'élévation de surface en un point de l'espace) ont été non-uniformément distribuées en espace, afin de représenter l'échantillonnage typique d'une surface par un capteur lidar observant la surface avec un angle d'incidence rasant. Dans le but de valider la méthodologie de génération de données, l'adéquation des états de mer générés avec leurs caractéristiques théoriques (hauteurs significatives cibles) est d'abord vérifiée. Ensuite, le jeu de données correspondant aux cas de houle unidirectionnelle est analysé plus finement pour la quantification et la qualification physique des perturbations relatives aux conditions expérimentales. Au travers de l'examination des oscillations résiduelles mesurées par les sondes dans le bassin de houle réel, ces perturbations semblent être la conséquence de l'excitation par le batteur de modes propres transverses dans le bassin. Elles sont également identifiées comme la cause principale des différences observées entre les données expérimentales et numériques. Nous évaluons par la suite, la sensibilité de l'algorithme de prédiction à des paramètres d'assimilation. Notre étude montre que la précision de la prédiction converge avec la durée des séries temporelles d'élévation de surface assimilées, c'est-à-dire la quantité de données utilisées pour effectuer la prédiction. Aussi, plus le pas de temps des données assimilée est faible, plus la convergence est rapide, et plus la cambrure caractéristique est élevée, plus l'erreur de prédiction augmente, sous la forme d'un déphasage non-linéaire des vagues prédites avec la solution de référence et d'une divergence de forme. Ces deux types de désaccord sont explicités au travers de la comparaison entre des élévations de surface pour un groupe de vagues spécifique. Les performances de ICWM relatives à LWT et LWT-CDR (LWT incluant une correction de phase similaire à celle de ICWM) sont ensuite quantifiées pour différentes configurations d'état de mer. Nous montrons que, pour l'estimateur d'erreur définie dans le cadre de cette étude, ICWM améliore les prédictions à l'intérieur de la zone de prédiction accessible, comparé à LWT, avec une réduction maximale de l'erreur de prédiction de $\sim 35\%$ pour une cambrure intermédiaire de $H_s/\lambda_p \approx 3.18\%$ (résultats numériques). Dans la région spatio-temporelle couverte par les observations, la correction non-linéaire de forme incluse dans ICWM améliore la représentation de l'élévation de surface. Pour les vagues qui se sont propagées (en espace ou en temps), le facteur principal de réduction d'erreur de prédiction est la modélisation correcte des vitesses de propagation des vagues. En revanche, en ce qui concerne la prédiction de pente, les propriétés de représentation de forme asymétrique des vagues de ICWM permet d'obtenir une réduction de l'erreur moyenne de 20% sur la zone de prédiction, pour une cambrure de $\sim$ 1–2% et des données numériques, en comparaison à une modélisation linéaire avec correction de la vitesse de propagation (LWT-CDR). Au travers d'une procédure entièrement validée expérimentalement, ces résultats démontrent la pertinence de l'utilisation de ICWM en vue de la prédiction de champs déterministes de la surface océanique. Les résultats présentés dans ce chapitre ont été publiés dans Desmars *et al.* (2020).

Chapter V

Experimental Investigation of the Prediction Algorithm

Introduction

We further apply our wave reconstruction and prediction algorithm to experimental data generated in the ECN ocean wave tank. With the help of a digital twin provided by a fully nonlinear HOS numerical framework, we investigate the performances and relative improvements of ICWM compared to LWT and LWT-CDR under different sea state configurations.

We first describe the methodology applied to generate relevant free surface elevation data in the wave tank and characterize the generated sea states regarding their adequacy with theoretical characteristic quantities. Then, the dataset pertaining to cases of long-crested waves is further analyzed through the quantification and physical qualification of the perturbations belonging to the experimental conditions. After the definitions of the prediction misfit indicators used in this study, we evaluate the sensitivity of the proposed prediction algorithms to assimilation parameters. Their accuracy is finally discussed in the last section of this chapter. A significant part of the following work is also presented in Desmars *et al.* (2020).

V.1 General Presentation of Experiments

This first section aims at describing the experimental campaign that will be used to perform predictions from real wave data. After a general presentation of the methodology to record free surface elevation fields, we quantify characteristic quantities of the generated sea states in order to properly characterize the data and validate their usability.

V.1.1 Description of the Experimental Campaign

Long- and short-crested wave trains are generated in the ocean wave tank of ECN, which is 50 m long, 30 m wide, and 5 m deep. Datasets are referred to full-scale wave parameters, but they are performed at a $\ell^* = 1{:}50$ geometric scale (corresponding time scale is $t^* = \sqrt{\ell^*} \approx 7.06$ under Froude scaling). Waves are generated at one side of the tank by 48 individual rotating flaps, and absorbed by a beach at the other extremity. In the following, we detail the experimental plan (i.e., the targeted sea states generated in the wave tank) and the experimental setup used to acquire surface elevation data.

V.1.1.1 Specified Sea States

Two characteristic parameters of the sea states are investigated in our experimental campaign: the directional spreading and the characteristic wave steepness. In each case, we use a JONSWAP

wave frequency spectrum (see equation I.2), along with a directional spreading function defined by

$$G(\theta) = \cos^{2s}\left(\frac{\theta - \theta^{\text{dir}}}{2}\right), \tag{V.1}$$

in which s is the parameter that defines the directional spreading of the wave field and θ^{dir} is the main wave direction of propagation, to obtain a directional wave spectrum according to equation (I.1). The directional spreading function for the three s values that are used later is depicted in figure V.1. The peak period $T_p = 10$ s (full scale) is kept constant for every generated sea states,

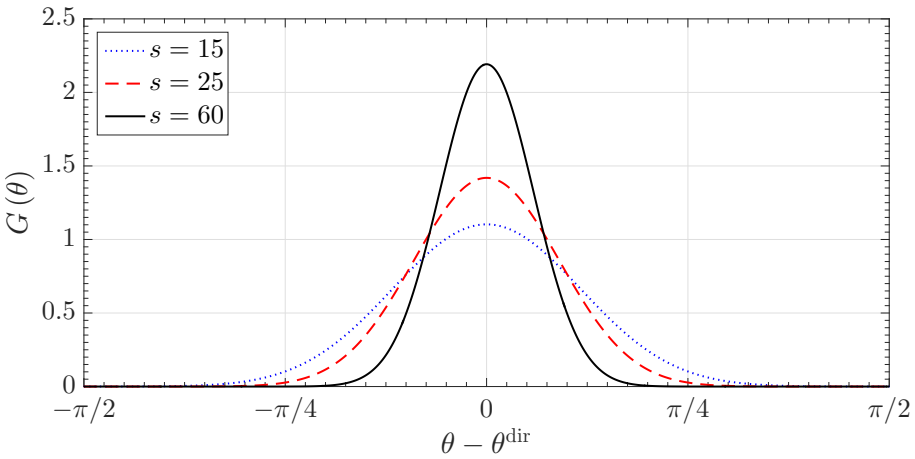

Figure V.1: Directional spreading function (V.1) for different s values.

while the significant wave height H_s varies such that the characteristic wave steepness ranges from $H_s/\lambda_p \approx 0.64\%$ (virtually linear) to 7.05% (highly nonlinear).

As summarized in table V.1, four spreading parameters $s = 15$, 25, 60 and $+\infty$ are chosen, defining four sets of measurements. The steepness is swept for two of these sets, $s = 25$ and $+\infty$,

set #	H_s [m]	T_p [s]	H_s/λ_p [%]	s
1	5.00	10	3.20	15
2	2.00	10	1.28	25
	3.00	10	1.92	25
	4.00	10	2.56	25
	5.00	10	3.20	25
	6.00	10	3.84	25
	7.00	10	4.48	25
	9.00	10	5.76	25
	11.00	10	7.05	25
3	5.00	10	3.20	60
4	1.00	10	0.64	$+\infty$
	2.00	10	1.28	$+\infty$
	3.00	10	1.92	$+\infty$
	4.00	10	2.56	$+\infty$
	5.00	10	3.20	$+\infty$
	6.00	10	3.84	$+\infty$
	7.00	10	4.48	$+\infty$
	9.00	10	5.76	$+\infty$

Table V.1: Summary of the full-scale sea states parameters used in the experiments.

the latter corresponding to the limiting case of long-crested waves. For one specific steepness

of $H_s/\lambda_p \approx 3.20\%$, sea states are generated using the four directional spreading values. Note that, at wave-tank scale, the peak wavelength λ_p is about 3.12 m, which is lower that the water depth, indicating that the bottom of the wave tank does not play a significant role in the wave propagation processes and that deep-water approximation is applicable. Hence, the wave models described in chapter II under the deep-water assumption are used in the following without further consideration about the influence of the water depth.

V.1.1.2 Methodology and Setup Configuration

As discussed in section I.2.2.2, no off-the-shelf operational lidar camera for ocean surface profiling is available so far. In our experiments, the surface elevation data are thus generated through measurements of surface elevation time series using wave probes. For each sea state, the wave fields are generated during a time period of 300 s at wave-tank scale (i.e., approximately $212T_p$). In order to partly reflect the influence of the nonuniform spatial distribution of optical measurements, the wave probes disposition is decided with respect to what a realistic lidar camera footprint would have on the surface. The space sampling rate is dictated by the number of available wave probes for our experimental campaign. We have at our disposal 22 resistive plus 2 capacitive wave probes, leading to a total of 24 probes. Since this number is very low compared to the number of spatial observations made by a real lidar sensor (refer to section III.1.1 for details), we increase, in case of short-crested waves (i.e., $s = 15, 25, 60$), the total number of spatial observations by relying on the wave tank ability to replicate sea state realizations. More specifically, the same wavemaker motion is executed multiple times to generate wave fields of similar deterministic characteristics (i.e., similar distributions of waves' amplitudes, directions of propagation *and* phases), while, at each time, the wave probes are positioned at different locations.

Long-Crested Waves

In case of long-crested waves ($s = +\infty$), the locations of the wave probes are determined as follows. A virtual "one-directional" lidar camera is located approximately halfway through the wave tank, at an elevation $z_c = 30$ m (0.6 m in tank scale, $z_c/\lambda_p \approx 0.19$), and aims at the water surface with an angle $\alpha = 76°$ and 20 virtual beams which are uniformly spread over an aperture angle of $\alpha_a = 20°$ (see figure III.3 for a representation of α and α_a). We consider the illumination of a flat surface, which allowed us to position the probes vertically into the water. This way, wave-shadowing effects are not reproduced but only the geometrically decreasing density of observations, which is the prominent source of non-uniformity in the measurement points locations. In order to assess the prediction performances, two additional probes are disposed further away, generating reference surface elevations that can be compared to the predicted elevations. The probes are distributed along the x-axis such that the lidar is facing the wavemaker, as depicted in figure V.2. Every probe is labeled according to its x-location, from 1 for that closest to the

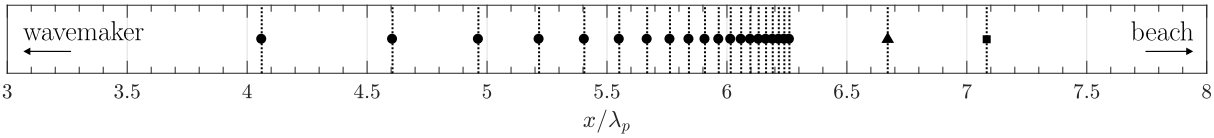

Figure V.2: Location of observation wave probes 1 to 20 (●) and of two additional downstream probes 21 (▲) and 22 (■). The wavemaker is located at $x/\lambda_p = 0$ and the beach at $x/\lambda_p \approx 14.86$.

wavemaker to 22 for that furthest away. In this case, the number of spatial observations is constant at $J = 20$.

Short-Crested Waves

As mentioned above, the number of spatial observations in case of short-crested waves is increased by generating similar wave fields for which the locations of the wave probes are different. Practically, the aligned wave probes distribution described previously is mounted on a ladder, and oriented for each sea state realization according to an angle χ with respect to the opposite x-direction. As depicted in figure V.3, the rotational axis of the ladder coincides with the location of the probe 21. The chosen χ values correspond to the horizontal viewing and aperture angles

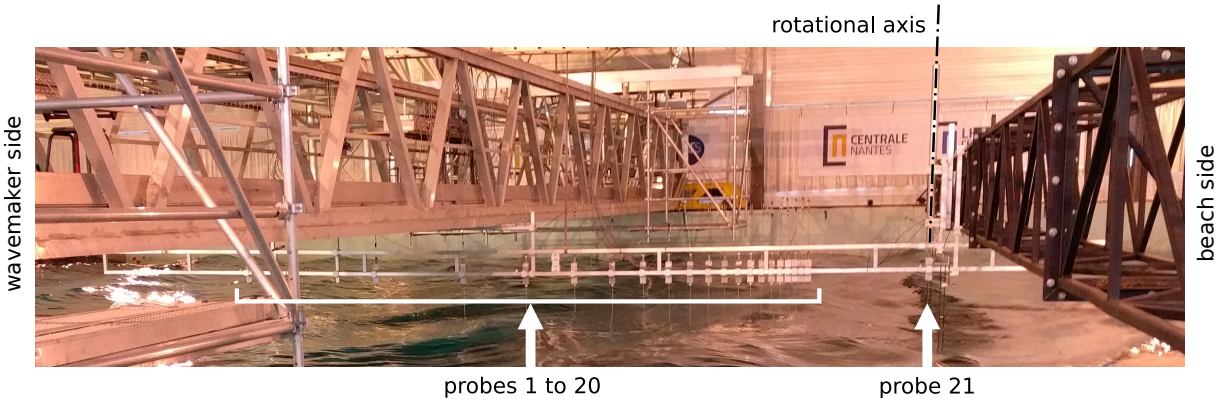

Figure V.3: Photo of the ladder upon which the wave probes are mounted. The whole ladder can be rotated around a vertical axis located at the horizontal location of the probe 21.

of our virtual lidar camera (figure III.3), set to $\beta = \theta^{\text{dir}} = 0°$ and $\beta_a = 40°$, respectively. A reasonable number of 9 wave-field replicates is chosen, fixing the directional sampling to the following distribution of angles $\chi = \{-20°, -15°, -10°, -5°, 0°, 5°, 10°, 15°, 20°\}$, which yields the geometrical distribution of the wave probe locations depicted in figure V.4. Hence, a total

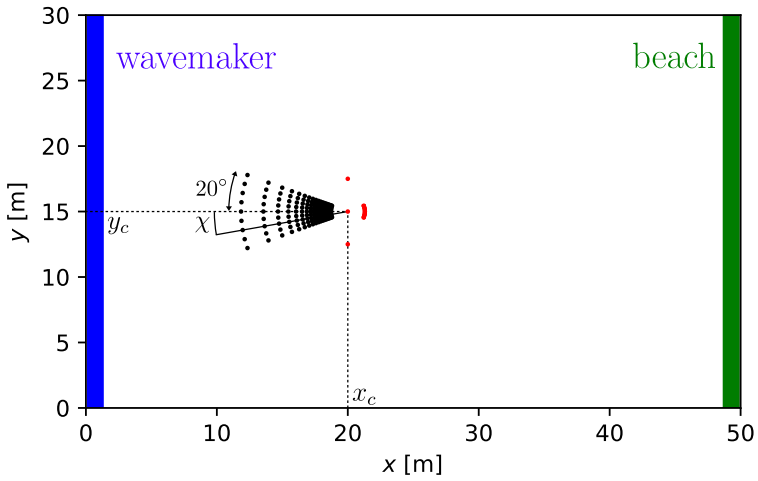

Figure V.4: Locations of the wave probes in the ECN wave tank. Two sets of wave probes are depicted: one is used to generate the observations (●), and the other to compare with predictions (●). Waves propagate from the wavemaker to the beach, i.e., along the x-axis direction.

number of $J = 20 \times 9 = 180$ spatial observations are used in the case of short-crested waves.

Figure V.4 also shows the location of the downstream reference probes (in red), corresponding to the probes 21 (fixed) and 22 (moving according to the rotation of the ladder), and to the two capacitive probes that are positioned sideways, providing two additional surface elevation references. The location of these reference probes is chosen with respect to the extent of the accessible prediction zone for a typical horizon of prediction. For instance, we see on figure V.5 that these reference probes are centered in the prediction zone for a prediction time of $2.5T_p$

after the reconstruction and calculated with an assimilation time $T_a = T_p$, a direction bandwidth $\left[\theta^{\min}, \theta^{\max}\right]$ (see figure III.9) centered on θ^{dir} and set to truncate 1% of the energy beyond the limits (i.e., similar to section IV.2.4), and a spreading parameter $s = 25$.

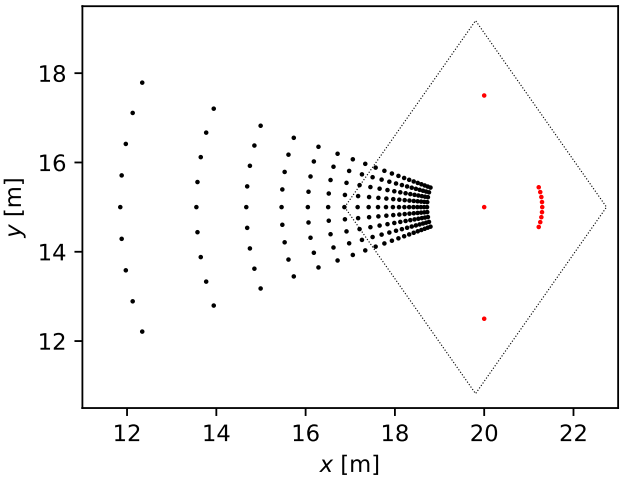

Figure V.5: Similar to figure V.4 with a closer look at the location of the wave probes. The parallelogram (·····) defines the spatial boundaries of the prediction zone after $2.5T_p$ of propagation using and assimilation time $T_a = T_p$, a direction bandwidth that truncates 1% of the wave energy and a spreading parameter $s = 25$.

The obtained spatial distribution is similar to what a rotating lidar would have yield, neglecting wave shadowing effects, and with fewer spatial observations due to the limited number of available wave probes for our experimental campaign.

V.1.2 Sea States Characterization

The wavemaker motion that is set for the wave field realizations is deduced by applying a transfer function based on the finite-depth linear wavemaker theory and on the wavemaker geometry, which is a rotating flap that is hinged three meters below the mean surface level. The amplitude of the wavemaker deflection is adjusted according to the target H_s values. The Dalrymple method (Dean & Dalrymple, 1991) for the generation of short-crested waves is employed to properly manage the extent of the spatial region within which the directional properties of the generated wave field are consistent with the specified sea state. Without further consideration, the obtained theoretical motion serves as input for our physical wavemaker. In order to (i) verify that the wave fields physically generated in the wave tank correspond to the theoretically specified sea states, and to (ii) investigate the wave maker capacity to reproduce similar wave field realizations, we quantify the significant wave height for different experimental configurations. Then, beyond the evaluation of a statistical quantity, we check the agreement of the actual surface elevations obtained for different wave-field replicates.

V.1.2.1 Significant Wave Height Definition

The estimation of the significant wave height is done through the calculation of the standard deviation σ_η of the surface elevation over the steady part of the signal (i.e., when every generated wave component has reach the furthest wave probes from the wavemaker) as

$$H_s = 4\sigma_\eta. \tag{V.2}$$

As explained later in section V.2.3, the signals used to calculate σ_η contain ~ 173 waves, which is enough to yield converged results. Multiple mean values for H_s are calculated. First, the estimates for every wave probes are averaged, yielding one mean value $\langle H_s \rangle_{\text{p}}$ for every wave

probe network direction χ, which allows us to evaluate the wave tank ability to replicate wave fields in terms of input energy. We also define a global mean $\langle H_s \rangle$ averaging all estimates that should theoretically lead to the same value of H_s. Along with the mean values, we determine their associated standard deviations ($\mathrm{std}\,(H_s)_\mathrm{p}$ and $\mathrm{std}\,(H_s)$) in order to check the variability in the H_s estimation in each configuration.

Wave-field replicates

We first evaluate the significant wave height for each wave-field replicates (i.e., sea-state realizations with an identical set of wave phases), and for various directional spreading values. Figure V.6a shows the estimates $\langle H_s \rangle_\mathrm{p}$ normalized by the target value $H_s^{\mathrm{target}} = 5$ m that corresponds to the specified value reported in table V.1. We first see that, for every sea-state

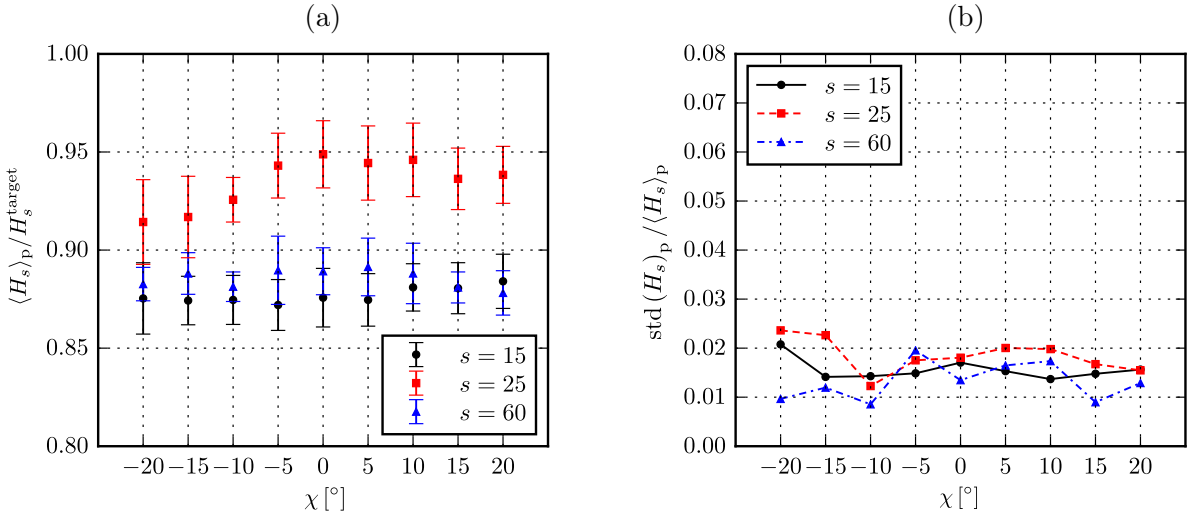

Figure V.6: (a) Normalized global mean estimate of H_s and (b) associated standard deviation for each similar wave field realization.

realization, the generated wave field has a significant height that is lower than the specified one, but still higher than 87% of the target. As also explained later in section V.2.2.1, this phenomenon is likely to come from properties of the physical wavemaker transfer function. While for $s = 15$ and 60 the mean estimates $\langle H_s \rangle_\mathrm{p}$ have similar values ranging from 87% to 89% of the target, cases corresponding to $s = 25$ yield normalized values fluctuating between 91% and 95%. The overall high ratios validate the capacity of the method (relying on the deep-water linear wavemaker theory and the Dalrymple's method) to generate wave fields of energy that is close to the specified one. Additionally, variations of only $\sim$ 1–2% for $s = 15$ and 60, and of $\sim$ 3–4% for $s = 25$, are observed between every wave-field replicate, which confirms that the wave-field replicates are of comparable amount of energy.

Figure V.6b presents the corresponding standard deviations normalized with the associated mean value. We see that, proportionally to the mean significant wave height, similar values of standard deviations are obtained ($\sim$ 1–2% of the measured mean significant wave height) for every wave-field replicate as well as for every directional spreading. This indicates that, within the spatial region over which the probes are deployed, the spatial variation of H_s is limited to only $\sim$ 2%.

Characteristic wave steepness

For directional spreading values $s = 25$ and $+\infty$, sea states of various wave steepnesses are generated in the wave tank. For the case $s = 25$, we use data from the 9 wave-field replicates to calculate a global mean significant wave height $\langle H_s \rangle$ (calculated at each probe using equation

(V.2), then averaged over the $24 \times 9 = 216$ recorded time series) and its associated standard deviation $\mathrm{std}\,(H_s)$. For the case $s = +\infty$, only one wave field realization was generated for each steepness, meaning that we use 24 time series to calculate $\langle H_s \rangle$ and $\mathrm{std}\,(H_s)$.

Figure V.7a shows the measured significant wave heights as a function of the target wave heights reported in table V.1. A clear trend in the evolution of the normalized H_s estimates is

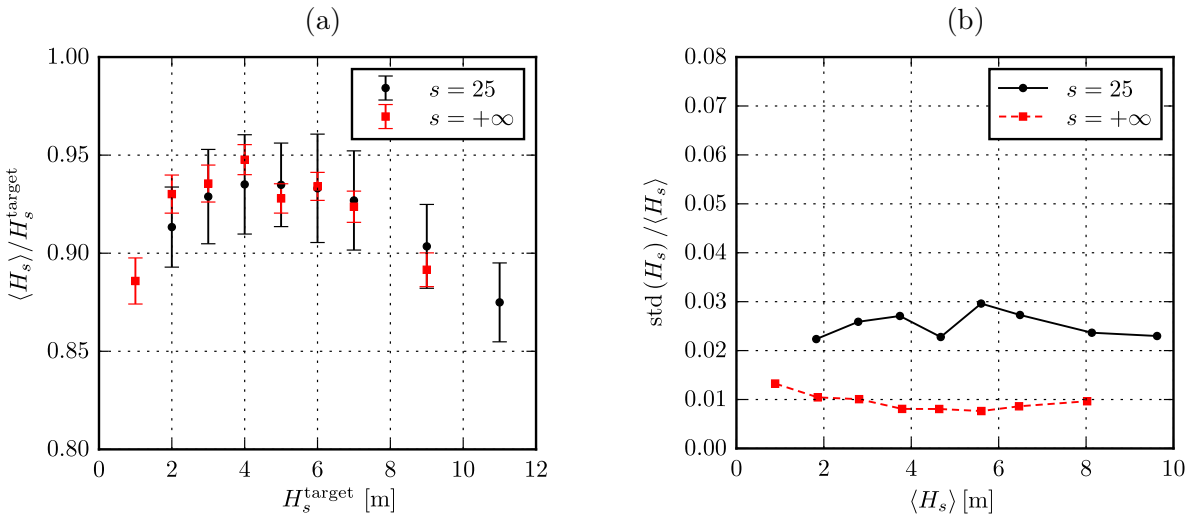

Figure V.7: (a) Normalized global mean estimate of H_s and (b) associated standard deviation as a function of the characteristic wave steepness for both short-crested waves ($s = 25$) and long-crested waves ($s = +\infty$).

observable for both s values: the ratios increase to reach maximal values of $\sim 95\%$ and $\sim 93\%$ for $s = 25$ and $+\infty$, respectively, for a target value of 4 m, before deceasing for higher H_s^{target}. Two reasons can explain this behavior. First and as already mentioned in the previous paragraph, the wavemaker does not behave exactly as what the linear wavemaker theory predicts. Indeed, due to the generation of local disturbances as well as to possible inaccuracies in the wavemaker motion amplitude (see section V.2.2.1 for details), the transmitted energy, quantified through the estimate of H_s, is lower than the specified value. Second, for large steepness (typically for $H_s^{\mathrm{target}} \gtrsim 4$ m), local wave breaking events start appearing, leading to the dissipation of wave energy. The larger the steepness, the larger the dissipation through wave breaking, which could explain the decrease of the ratio $\langle H_s \rangle / H_s^{\mathrm{target}}$ for large steepness. Even if the mean significant wave heights are similar for both directional spreading values, standard deviations (see figure V.7b) are lower for the long-crested than for the short-crested waves, for which they yield $\sim 1\%$ and $\sim$ 2–3% of the significant wave height, respectively. These values are almost constant for each wave steepness, indicating that the variation of H_s linearly depends on the actual H_s. The larger values for $s = 25$ is explained by the fact that $\langle H_s \rangle$ is calculated using different wave-field replicates, which contributes to enlarge the spreading of the measured H_s values around the mean.

Directional spreading

For each directional spreading, figure V.8a shows the normalized global mean estimate of H_s for a steepness corresponding to $H_s^{\mathrm{target}} = 5$ m. No clear trend is observed as all values range from 87% to 93%, the highest value corresponding to the moderate directional spreading $s = 25$. For the same reason as that mentioned in the previous paragraph, the standard deviation is larger for short-crested waves (i.e., for $s = 15$, 25, 60) than for long-crested ones (figure V.8b).

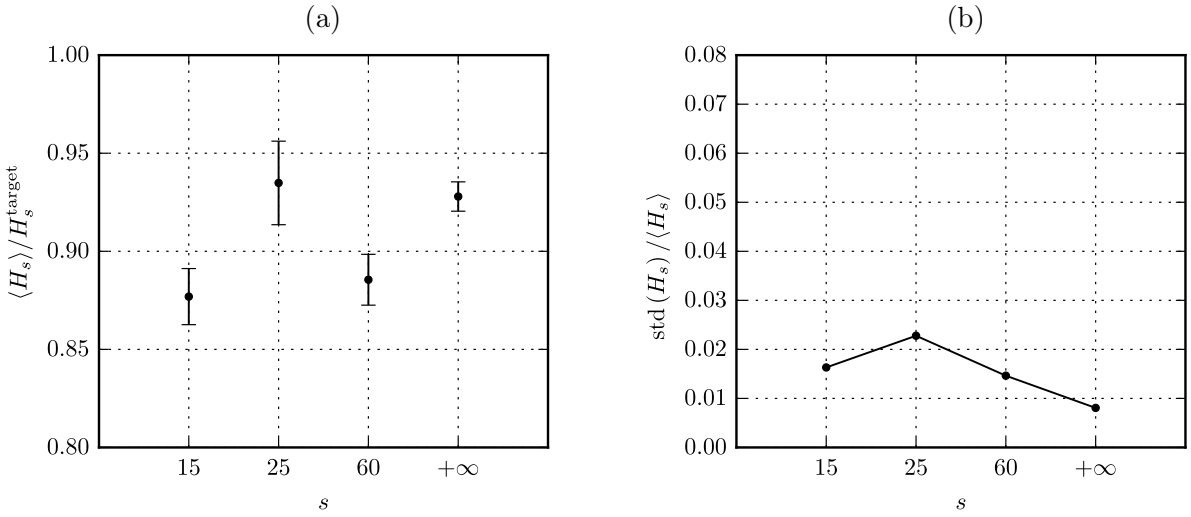

Figure V.8: (a) Normalized global mean estimate of H_s and (b) associated standard deviation as a function of the directional spreading of the generated sea states. Data corresponding to a target $H_s^{\text{target}} = 5$ m are used for $s = 25$ and $+\infty$.

V.1.2.2 Free-Surface Elevation

In order to check the repeatability of the generated wave fields in terms of deterministic quantities, we quantify here the standard deviation of the free surface elevation at the location of a fixed probe. For instance, at the location of one of the capacitive probe (the furthest along the y-axis), figure V.9 presents a sample of the recorded time series of free surface elevation for the sea state corresponding to $s = 25$ and $H_s^{\text{target}} = 5$ m. The mean surface elevation $\langle \eta \rangle$

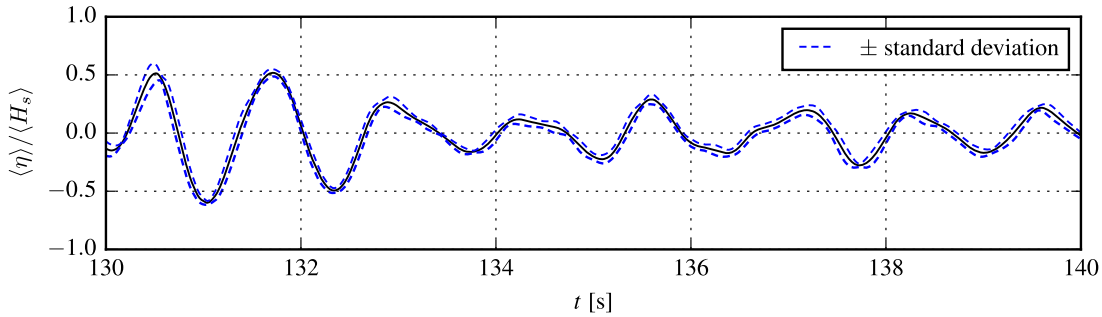

Figure V.9: Normalized mean free surface elevation with instantaneous standard deviation. The directional spreading is $s = 25$ and the measured significant wave height is $\langle H_s \rangle \approx 4.7$ m.

(averaged over the nine wave field realizations) is plotted as well as the instantaneous standard deviation. On average, the obtained standard deviation is about 4.5% of the measured significant wave height for the depicted case. For the other cases, the mean standard deviation is about 3% to 5% depending on characteristic wave steepness. This quantification allows defining an uncertainty that could help for the assessment of the performances of predictions based on multiple wave-field replicates. Since the case of $s = 25$ yields the largest standard deviation for the estimate of H_s, lower values are expected for the cases of other directional spreading.

Last, figure V.10 shows the mean surface elevations obtained at the same location for cases of different characteristic wave steepness. We recall that, for all the presented cases, we keep the same set of initial wave phases, such that only the amplitude of the wavemaker motion changes. The two main nonlinear effects on surface wave kinematics (that are, at least partly, modeled in ICWM) are clearly observable: waves get shifted towards lower times (i.e., waves propagate with

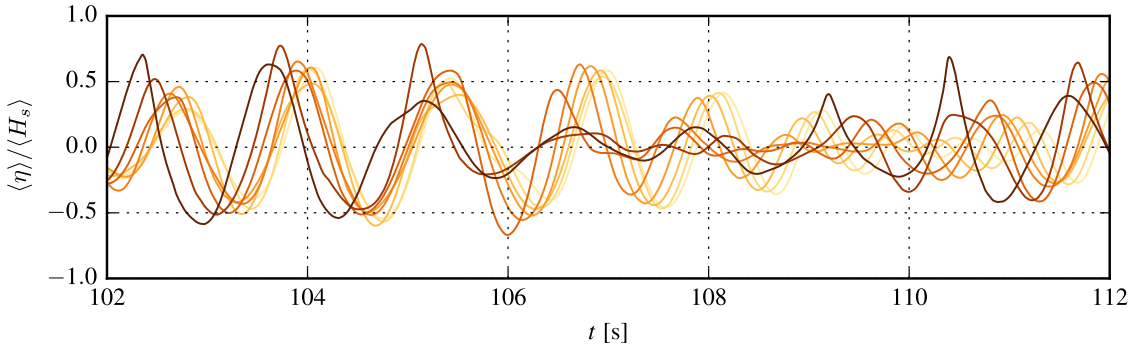

Figure V.10: Normalized mean free surface elevations for different H_s values recorded at the same location. The directional spreading is $s = 25$. The darker the color the larger the wave steepness. Only the amplitude of the wavemaker motion changes between each case (i.e., the initial set of wave phases is kept constant).

a larger velocity) and wave geometry gets sharper for higher steepness. Due to the modification of the relative phase shifts between wave components as well as to other time- and space-dependent nonlinear resonant phenomena, increasing the wave steepness also has nontrivial impacts on the surface elevation, such as for $107 < t < 111$ in figure V.10. In the following, we will investigate the improvements pertaining to ICWM compared to LWT in predicting such nonlinear wave fields.

V.2 Further Analysis of Long-Crested Waves

In this section, we further investigate the generated unidirectional sea states (set #4 in table V.1). We first detail a numerical setup that is used to compute a digital twin of the experiments, which later serves to improve the analysis pertaining to the experimental results. Then, we investigate the adequacy between the specified sea states and the actually generated ones in the physical and numerical wave tanks, and identify the sources of the differences in each case. Finally, we analyze and quantify the influence of the perturbations pertaining to the limitations of our experimental setup (referred to as "noise" throughout this work) on the recorded wave signal.

V.2.1 Digital Twin of Set #4

In order to provide surface elevation data that are free of noisy perturbations, we use a numerical model to compute similar cases to that described previously for unidirectional wave fields (i.e., $s = +\infty$). Numerical simulations are performed using the open-source code HOS-NWT developed at ECN. It makes use of the HOS method to simulate a numerical wave tank, and has been extensively used and validated against real wave tank experiments (Bonnefoy *et al.*, 2010; Ducrozet *et al.*, 2012). Based on a pseudo-spectral approach, the HOS method solves, to an arbitrary order M in wave steepness, the nonlinear free surface boundary conditions for a velocity potential (see section I.3.4.1 for details). A converged estimate of the potential (for which a value of $M = 7$ and a calculation grid made of 512 points are used hereafter) gives access to the fully nonlinear solution. As waves of characteristic wave steepness larger than $H_s/\lambda_p \sim 0.035$ will start breaking, a wave breaking model allowing to both detect impending breaking and absorb wave energy is used in HOS-NWT (Seiffert & Ducrozet, 2018; Seiffert *et al.*, 2017). The same wavemaker motion is specified in both laboratory experiments and numerical simulations, which ensures a consistent comparison between experimental and numerical results. The numerical wave tank geometry as well as the beach performance are similar to those of the physical wave tank in used for our experimental campaign.

V.2.2 Limitations of the Experimental/Modeling Setups

In both experiments and numerical simulations, we consider a full-scale one-directional wave field extracted from a JONSWAP spectrum with a $T_p = 10$ s peak period (≈ 1.41 s at tank scale) and a $\gamma = 3.3$ peakedness parameter. As detailed in section V.1.1.1, eight sea-states are generated using the same set of wave phases (table V.2), with their significant wave height $H_s = H_{m_0} = 4\sqrt{m_0}$ (where $m_0 = \int_0^{+\infty} S_\eta(f)\,\mathrm{d}f$) selected such that the characteristic steepness H_s/λ_p varies between $\sim 0.6\%$ and $\sim 5\%$, with $\lambda_p \approx 156$ m (3.12 m in tank scale). Note that

case	target		experiments		simulations	
	H_s [m]	H_s/λ_p [%]	H_s [m]	H_s/λ_p [%]	H_s [m]	H_s/λ_p [%]
A	1.00	0.64	0.89	0.57	1.01	0.65
B	2.00	1.28	1.86	1.19	2.02	1.29
C	3.00	1.92	2.81	1.80	3.02	1.93
D	4.00	2.56	3.79	2.43	4.01	2.57
E	5.00	3.20	4.64	2.97	4.98	3.18
F	6.00	3.84	5.60	3.59	5.87	3.76
G	7.00	4.48	6.46	4.14	6.69	4.28
H	9.00	5.76	8.02	5.13	8.06	5.16

Table V.2: Summary of the targeted and generated full-scale sea states in both experiments and numerical simulations. Each case is labeled in alphabetical order from the smallest to the largest characteristic wave steepness H_s/λ_p.

the H_s values reported in table V.2 are calculated as explained in section V.2.3, according to equation (V.4a). As already mentioned, we notice differences between the target and generated significant wave heights. The reduction of our estimates of H_s is easily explained by wave energy dissipation through the apparition of wave breaking events for high steepness, i.e., $H_s/\lambda_p \gtrsim 3.5\%$ (cases E to H). Wave-breaking dissipation is encountered in both experiments and simulations due to the wave breaking modeling in the numerical model. However, for sea states of small to moderate steepness (cases A to D), these differences are mainly explained by the method used to calculate the wavemaker transfer function, while the influence of wave reflection is very limited, as shown in the following sections.

V.2.2.1 Wavemaker Transfer Function

In order to calculate a theoretical wavemaker motion, we used the classical method employed in ocean-engineering wave-tank experiments based on the wavemaker geometry (which is identical in both numerical and physical configurations) and on the finite-depth wavemaker linear theory (e.g., Bonnefoy, 2005). The obtained theoretical motion is used to set both the numerical and physical wavemaker motions, meaning that the theoretical amount of energy provided by the wavemaker is the same in both cases. However, in addition to the use of the linear wavemaker theory – which impacts the wave generation accuracy for large wave steepness and may play a role the generation the nontrivial effects reported in the description of figure V.10 – two sources of experimental uncertainties have been identified:

- While the numerical wavemaker is perfectly one-directional, i.e., the entire kinetic energy of wavemaker is transmitted to the fluid in the form of a one-directional gravity wave field, the physical one is composed of 48 individual flaps spread along the width of the tank. Although each flap is perfectly aligned with the others and run with the exact same motion, small gaps between flaps (see figure V.11) introduce directional perturbations whose energy is not transmitted to the one-directional wave components of interest. Additionally, we experienced the formation of horse-shoe pattern resulting from nonlinear wave interactions

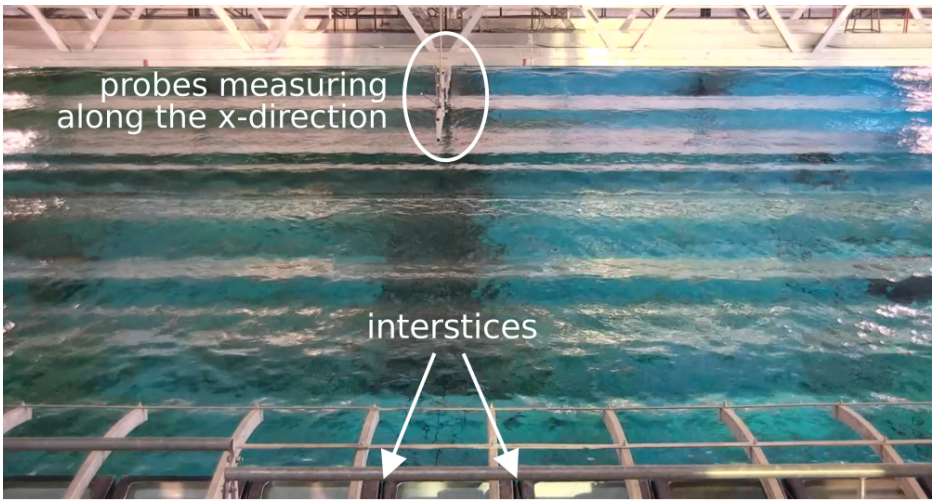

Figure V.11: Photo of the wave tank from above the wavemaker and looking along the x-direction. All flaps composing the wavemaker are aligned but there are small interstices between neighboring flaps.

during the propagation of steep waves, which contributes to the energy transfer between longitudinal waves to transverse waves.

- While the numerical wavemaker reproduces exactly the theoretical motion, we did not measure the actual motion of the physical one, which may slightly differ from the theory. Dissimilarities would directly affect the generated waves.

The important point here is that we did not correct the wavemaker transfer function for the calculation of the experimental wavemaker motion: the aforementioned sources of experimental uncertainties affect the generated wave fields.

Hence, in experiments, the wavemaker transfer function leads the physical wavemaker to generate waves of lower amplitude than according to the input. In contrast, simulations yield H_s values that are slightly higher than the targets by an amount that, as explained in the next section, is on the same order of magnitude as the expected effect pertaining to the wave reflection on the beach.

V.2.2.2 Wave Reflection

We know that the measured reflection coefficient C_r of the ECN wave tank beach is small (below 10 % of wave amplitudes) for the range of the wave periods at aim. The incident and reflected wave fields being incoherent, the total energy m_0 is the sum of incident energy m_0^{inc} and the reflected energy that follows $m_0^{\text{ref}} = C_r^2 m_0^{\text{inc}}$. As a consequence, we expect the contribution of the reflected waves on the estimated significant wave height $H_s = 4\sqrt{m_0}$ to be

$$H_s/H_s^{\text{inc}} = \sqrt{1 + C_r^2}. \tag{V.3}$$

For a reflection coefficient of 10%, this gives a contribution of ~ 0.5 % of the reflected waves on the estimated H_s. The ratio H_s/H_s^{inc} has been quantified through Matlab simulations based on the superposition of a linear wave field and its reflected wave field (in similar configurations to those of the experiments) as a function of the reflection coefficient. Both theoretical (equation (V.3)) and numerical results are shown in figure V.12 to agree very well. The described impact of wave reflection on the estimate of H_s is in agreement with our HOS-NWT simulations (in which the numerical beach is set such that its reflection rate is on the same order of magnitude as the physical one (Bonnefoy, 2005)): in cases that do not exhibit wave breaking (i.e., A to D),

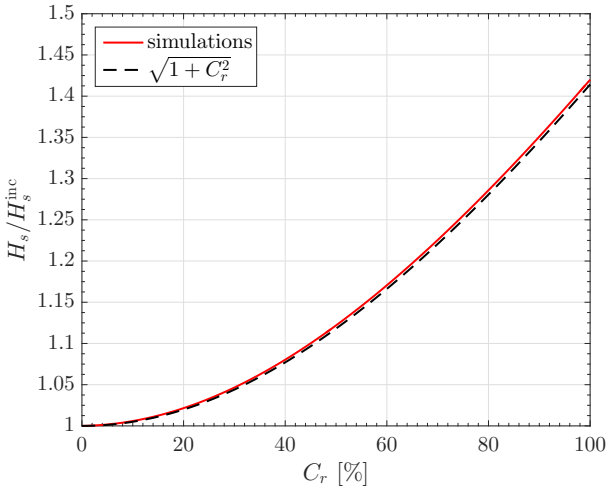

Figure V.12: Contribution of reflected waves on the estimated significant wave height as a function of the reflection coefficient. Results from Matlab simulations based on a linear JONSWAP spectrum and from the theoretical formula (V.3) are plotted.

we observe a very low overestimate of H_s values compared to the targets (lower than 1 % in each case). The wave reflection is expected to have a comparable effect on the experimental results, i.e., a very limited overestimation of H_s values.

Ideally, it would have been more relevant to either use a corrected experimental wavemaker transfer function to take into account the physical effects that differ from the theory in the wave generation, or calibrate our simulations to ensure consistency with the wave fields generated in the physical wave tank. Still, experiments and simulations yield wave fields of comparable energy content.

V.2.3 Characterization of the Experimental Noise

In order to fully characterize the experimental signals and understand what physical processes are recorded by the wave probes, we further study the influence of the noise (i.e., experimental perturbations) on the wave field of interest. Figure V.13 shows a typical time series of surface elevation measured at a resistive probe in laboratory experiments. Time $t = 0$ corresponds to the

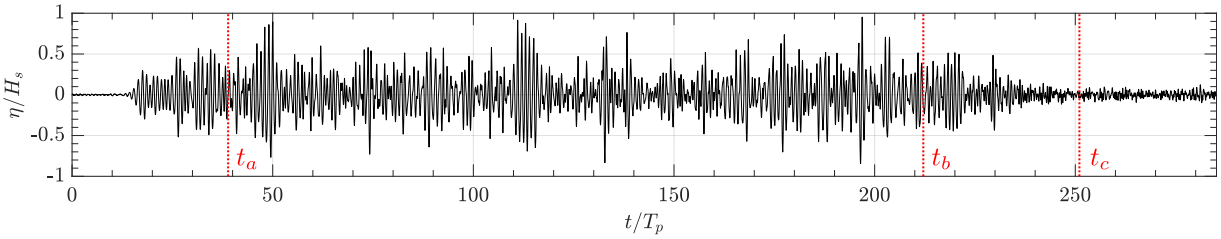

Figure V.13: Time series of surface elevation measured in experiments at probe 22 for case E (table V.2). Three characteristic times are marked on the record, t_a: all the generated wave components have been measured at all probes; t_b: shutdown of the wavemaker; t_c: all the generated wave components have propagated past all probes.

start of the wavemaker motion, i.e., the beginning of wave generation. At time t_a, it is estimated based on a truncation of the prescribed wave energy spectrum (similar to section III.3.4) that all the energetic wave components generated at the wavemaker have been seen by all probes. At time t_b, the wavemaker is shutdown and waves generation is interrupted. Finally, similar to the determination of t_a, at time t_c, it is estimated that all the generated energetic wave components

have propagated past all probes. Based on this, the data used for the wave field prediction study is restricted to the time interval $[t_a, t_b]$, with $t_b - t_a \approx 173 T_p$.

V.2.3.1 Amplitude Quantification

We consider that the rest of the data acquired at wave probes, for a few dozen peak wave periods beyond $t = t_c$, provides a representation of noise during the entire data acquisition duration. Based on this data, a noise to signal ratio NSR $= H_n/H_s$ is computed as a function of characteristic heights for both the primary wave ($t \in [t_a, t_b]$) and noise ($t \geq t_c$) signals following

$$\begin{cases} H_s = \dfrac{1}{N_p} \displaystyle\sum_{j=1}^{N_p} 4\,\sigma_\eta\left(x_j, t_a \leq t \leq t_b\right), & \text{(V.4a)} \\ H_n = \dfrac{1}{N_p} \displaystyle\sum_{j=1}^{N_p} 4\,\sigma_\eta\left(x_j, t \geq t_c\right), & \text{(V.4b)} \end{cases}$$

respectively, where $\sigma_\eta(x,t)$ denotes the standard deviation of the free surface elevation $\eta(x,t)$ and $N_p = 22$ wave probes. The NSR is computed for each case A to H in table V.2 and plotted in figure V.14 as a function of the corresponding characteristic wave steepness. We see that the

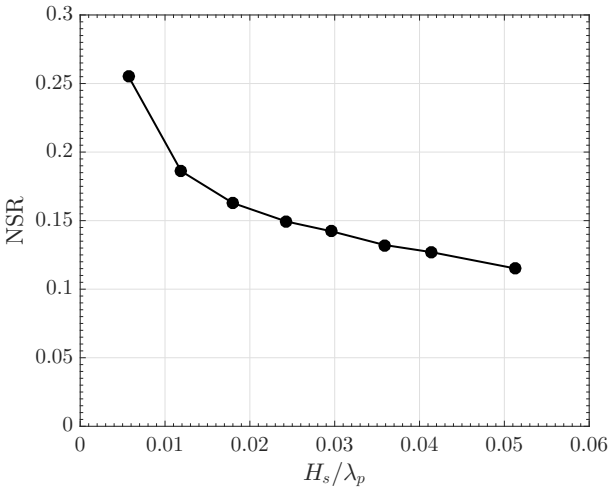

Figure V.14: Noise amplitude to wave amplitude ratio as a function of characteristic wave steepness.

NSR decreases as a function of wave steepness, with the largest value being about 25% for the smallest steepness and the smallest value being about 11.5% for the largest steepness. It thus appears that the geometry of our experimental set-up, in a 3D wave tank allowing the generation of directional wave fields, may have significantly affected the targeted one-directional wave fields. As will be detailed in section V.4.3, this potentially large NSR may affect the performance of the wave prediction algorithm.

V.2.3.2 Physical Qualification

To better quantify noise effects on the desired experimental data and relate the generated noise to a physical process, we computed the power spectral density S_n of the noisy part of the signal ($t \geq t_c$) (using Welch's method (Welch, 1967), with a Hanning window of length $\sim 10T_p$, overlap of $\sim 5T_p$, and a zero padded FFT of length 4096; the total length of the window is $\sim 34T_p$ sampled at $\sim 141 f_p$, where $f_p = 1/T_p$ is the peak frequency). For each steepness, the spectrum was averaged over results obtained at the 22 wave probes and normalized as $S_n^* = S_n f_p / m_0$, where $m_0 = H_n^2/16$ is the zeroth-order moment of the spectrum. These normalized noise spectra

were found to be nearly identical for each steepness. Figure V.15 shows their average, which is composed of a narrow-banded peak, centered on $\sim 1.1 f_p$, and a broad-banded high frequency part of much lower amplitude. The wave spectrum, calculated on $[t_a, t_b]$, is given on the same

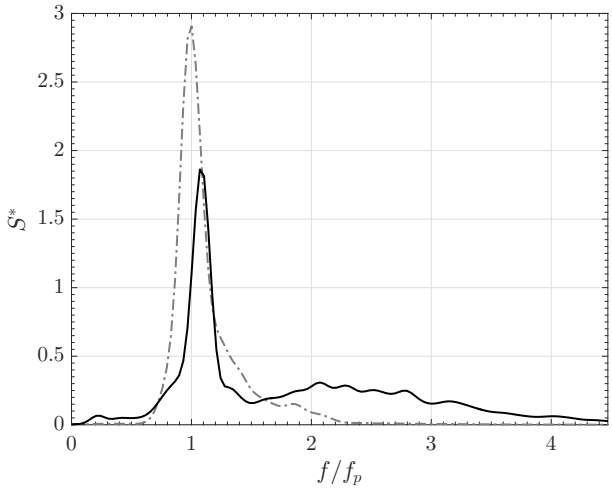

Figure V.15: Normalized noise (——) and wave (-·-·-) spectra averaged over all the characteristic wave steepnesses.

figure as a visual help for interpretation.

Figure V.16 shows normalized surface elevations of the noise signal η/H_n measured at various probes, in which frequencies $f > 1.4 f_p$ have been removed by filtering, i.e., these correspond to the dominant part of the noise signal. In figure V.16a, these are plotted at wave probe 22

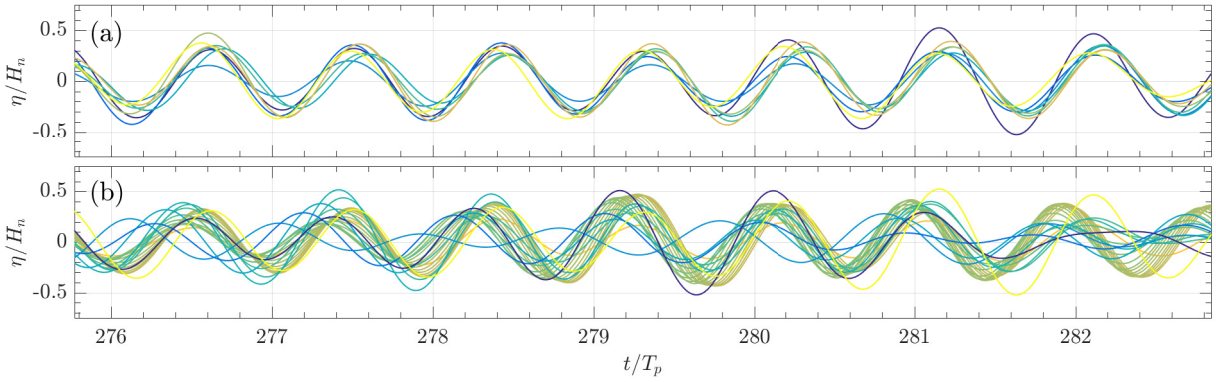

Figure V.16: Normalized surface elevations (a) at the location of gauge 22 for cases A to H, and (b) for case A at the locations of probes 1 to 22. Components of frequency higher that $1.4 f_p$ has been removed using a low-pass filter.

for each case A to H of different steepness, and we see that the dominant noise signals are in phase. This indicates that the noise generation process is deterministic: the chosen set of wave phases, which is the same in each case, leads to the generation of perturbations that are also in phase. Figure V.16b shows the dominant noise signal elevations for case A at wave probes 1 to 22 (which are aligned along the x-direction), and we see that noise signals are again mostly in phase, suggesting that the dominant experimental noise may be caused by resonant excitations of transverse modes in the 3D wave tank. This hypothesis is confirmed by our visual observations of these waves during calm-down times between measurements, and is explained by the presence of small interstices between the wavemaker flaps, locally generating transverse disturbances (see section V.2.2.1). Even during the waves generation and propagation, clear periodic transverse disturbances were observable, as shown in figure V.17. The much less energetic noisy components of the signal, with frequencies $f > 1.4 f_p$, were not found to be in phase, indicating that they

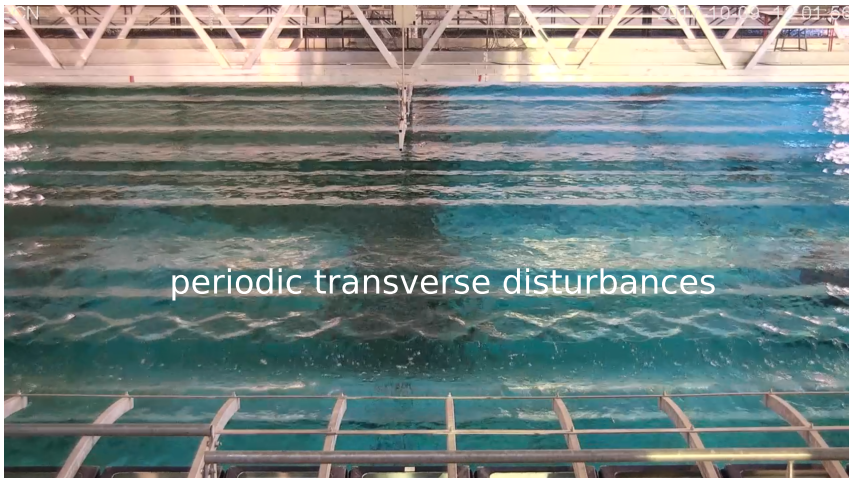

Figure V.17: Visual observation of periodic transverse disturbances during the generation of long-crested waves in the tank.

result from aleatory processes.

V.3 Sensitivity to Prediction Algorithm Parameters

Still on the basis the generated long-crested waves (set #4), this section focuses on the sensibility of the prediction accuracy to some algorithm parameters. After defining relevant indicators of the quality of free surface prediction and the procedure implemented to reliably evaluate them, we present the evolution of a global prediction error for different assimilation parameters, wave models and wave steepness.

Before proceeding to predictions, we briefly recall the important quantities that govern the assimilation procedure and the calculation of the accessible prediction zone. Wave probes provide observations that are used as input to the surface reconstruction and prediction algorithms. Here, the number of spatial observations is constant at $J = 20$ (probes 1 to 20), and the number of observation times K depends on the assimilation time duration T_a and data acquisition time step τ as $K = T_a/\tau$. The calculation of the one-directional prediction zone is depicted in figure V.18 for the setup used in this chapter (20 fixed observations). Further details regarding wave

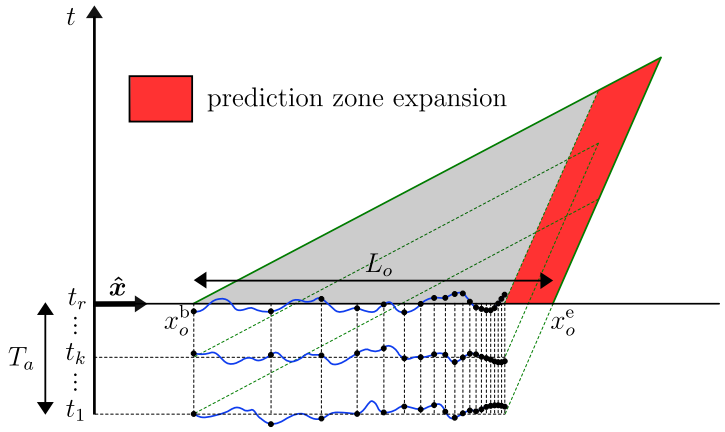

Figure V.18: Evolution of the wave prediction zone in time and space for the assimilation of spatio-temporal data. Green dash lines are prediction zones boundaries at times $t_1..t_k..t_r$. The increase in the prediction zone relative to that of spatial only observations (gray area) is highlighted in red.

model inversion and prediction zone calculation are presented in chapter III.

Here, contrary to the previous chapter in which we look for the wave model parameters $\boldsymbol{p} = \{a_n, b_n\}$ on a wavenumber grid that is linearly distributed in the interval $\left[k^{\min}, k^{\max}\right]$, we now use a logarithmic distribution of the wavenumbers . This allows a higher concentration of wave components in the energetic part of the spectrum, thus the use of a higher value for the high cutoff wavenumber without requiring the number of wave components N to be too large. The low cutoff frequency is, as before, calculated according the theoretical lowest measurable wavelength, i.e., $k^{\min} = 2\pi/L_o$ (see details in section III.3.3.1), but we now use a high cutoff wavenumber $k^{\max} = 20k_p$ (instead of $10k_p$ in the previous chapter).

V.3.1 Misfit Indicators Definitions

The misfit indicator used to quantify the accuracy of the predicted wave field is defined by equation (IV.2), in which η_{r} refers here to the measured or calculated reference surface depending on whether experimental or numerical data is used.

Ideally, we would like to analyze data generated before wave reflection starts impacting the measurements. Using the same method as in section V.2.3 to determine the propagation speeds of the longest and shortest waves of interest, the interval comprised between the time after the slowest wave reaches probe 22 and the time before the same probe is reached by the fastest reflected wave is about 10 seconds (i.e., approximately 7 peak periods). From a practical point of view, taking only a time interval on the order of 10 seconds necessitates a high number of wave field realizations with different sets of wave phases to generate enough data to properly estimate the prediction error. Indeed, as already mentioned in section IV.2.2, an unbiased estimate can only be obtained for a large number of samples from independent wave field realizations (i.e., of different sets of random wave phases) with, to the limit, $N_s \to \infty$. Even if this method (that consists in a Monte-Carlo approach for the estimate of the prediction error) can be used in a numerical framework, its application in experimental conditions is extremely time-consuming, thus not usable in practice.

Instead, as described in the next section, we chose to run only one single realization that we split into many time series that were used to perform statistical analyses. This method is largely employed in experimental hydrodynamics to study the statistical properties of wave fields, and, most importantly, wave reflections are expected to have a minor impact on the results, as quantified in section V.2.2.2 for H_s calculations. Moreover, having only time series limited to 10 seconds in length would have limited the extent of the sensitivity analysis described in section V.3.2, and the presence of residual reflections can be seen as a sensibility assessment of the developed prediction method to "environmental perturbations" that always exist in open-ocean conditions.

V.3.1.1 Generation of Surface Samples

We elect for this study to generate one single surface realization per sea state, but to record or compute wave probe data over a long time so that the signals can be split into a sufficiently large number of samples of meaningful duration T_a. Additionally, the number of samples is increased by selecting them as partially overlapping , i.e., shifting them in time by $\Delta t < T_a$, as illustrated in figure V.19. Therefore, the information used to estimate the misfit is the surface elevation data in the total time window covered by the samples, which has a duration

$$T_c = T_a + (N_s - 1)\,\Delta t. \tag{V.5}$$

A similar approach was employed by Naaijen *et al.* (2014) to investigate the spatio-temporal evolution of the prediction zone based on experimental and numerical data.

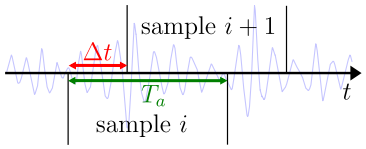

Figure V.19: Each sample consists of a time trace from the same generated surface realization. They can be partially overlapping, separated by a time span Δt from their neighbors.

V.3.1.2 Time-Averaged Prediction Error

The wave field prediction error at a specific location x is computed by averaging the corresponding misfit over the theoretical time-prediction zone $t \in \left[t^{\min}, t^{\max}\right]$ as

$$\mathcal{E}^{\mathcal{P}}(x) = \frac{1}{t^{\max} - t^{\min}} \int_{t^{\min}}^{t^{\max}} \mathcal{E}(x,t)\,\mathrm{d}t. \tag{V.6}$$

Figure V.20 shows, for case E and using simulated (i.e., HOS-NWT) reference data, the evolution of the wave field prediction error $\mathcal{E}^{\mathcal{P}}$ computed as a function of the amount of data used to calculate it, quantified by the relative duration T_c/T_p (equation (V.5)) of the time window used to evaluate the misfit $\mathcal{E}$. We see that the prediction error converges for $T_c/T_p \sim 60$ (similar results

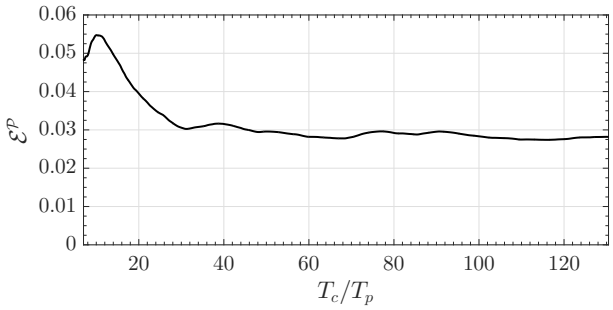

Figure V.20: Nonlinear (ICWM) prediction error estimate at the location of probe 22 as a function of the length of the time window used to evaluate the misfit $\mathcal{E}$, T_c (equation (V.5)), normalized by the peak spectral period T_p, and in which $\Delta t/T_p \approx 0.07$. Here, $T_a/T_p \approx 7$ and $\tau/T_p \approx 0.07$, and simulated reference data from case E are used.

are obtained using experimental data). Note that the wave probe network used to generate the observations covers a zone only slightly larger than $2\lambda_p$. If this zone was larger, the optimal number of peak wave periods for the sampling time window would likely be less than 60.

V.3.2 Prediction Error Sensitivity Assessment

We first assess the sensitivity of the prediction error to both the method used (linear or nonlinear) and parameters of the assimilation procedure, namely the assimilation time T_a and the time shift of the assimilated data τ.

For case A (which corresponds to a mild characteristic wave steepness of 0.65%) and simulated reference data, figure V.21a shows that the linear prediction error converges well as T_a increases, for the three considered τ values, although convergence is slower for larger τ. Hence, the converged error is independent of the time resolution of observations. This is a consequence of the characteristics of the physical description emerging from observations. As the assimilation time increases, the diversity of wave processes included in the assimilated information is enhanced, with respect to the relevant physics simulated in the model, causing the prediction error to converge. Additionally, the accuracy of the description of physical phenomena, which is directly function of the time resolution of observations, affects the prediction error convergence rate: for a given assimilation time T_a, a smaller time step τ will yield a prediction error closer

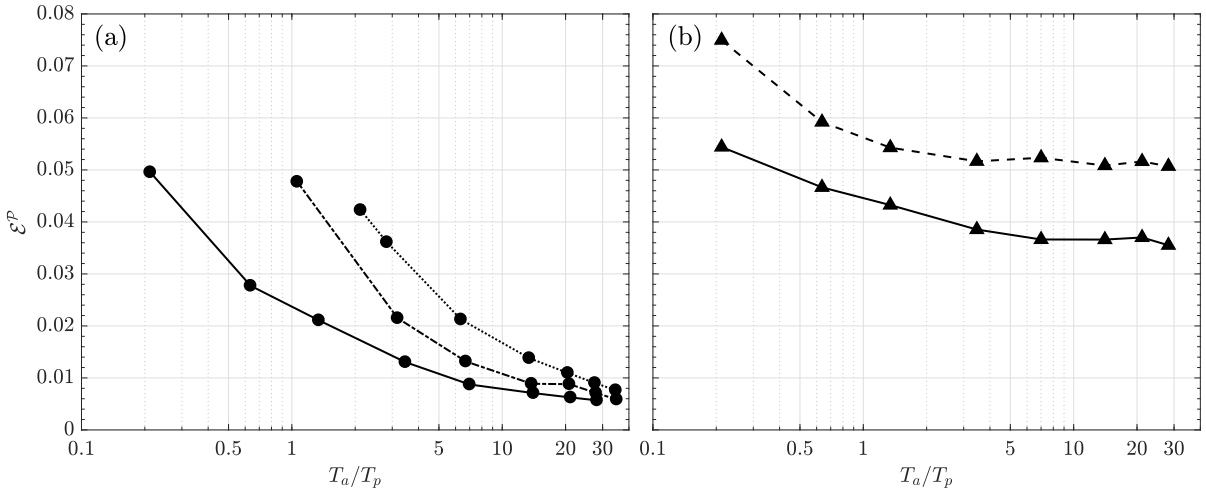

Figure V.21: Prediction error at the location of probe 22 as a function of the normalized assimilation time T_a/T_p in case (a) A (●) and (b) E (▲). In case A, the linear error is plotted for $\tau/T_p \approx 0.07$ (——), 0.35 (-·-·-·-) and 0.71 (·····). In case E, a time step $\tau/T_p \approx 0.07$ is used and both the linear (LWT, - - - -) and nonlinear (ICWM, ——) errors are plotted. Simulated reference data is used in both figures.

to the converged value. For the predictions presented later, we keep $\tau/T_p \approx 0.07$.

For case E, which corresponds to a larger wave steepness of 3.18% and hence a fairly nonlinear case, figure V.21b shows that, overall, prediction errors are larger than for case A, increasing from $[0.005, 0.05]$ to $[0.05, 0.075]$. Figure V.21b also shows that, as could be expected for this nonlinear case, the prediction errors are larger with the linear method than with the nonlinear method. Finally, the convergence of the nonlinear method to achieve an approximately constant value of $\mathcal{E}^{\mathcal{P}}$ requires a slightly larger T_a than for the linear method. This can be explained by the higher level of physics represented in the ICWM model than in LWT, which requires larger time scales to achieve convergence.

Note that, unlike section IV.2.3.3 in which the number of wave components for the wave model inversion is fixed to $N = 50$, we adapt here the value of N to prevent the prediction error from artificially increasing due to the insufficient number of wave components to accurately fit the data. As a simple illustration of the need to increase the number of wave components when the size of the assimilated data increases, figure V.22 shows the time propagation of cosine waves s_1 and s_2 of very close angular frequencies ω and $\omega + \Delta\omega$, respectively, with $\Delta\omega = \omega/50$. The sum $s_1 + s_2$ of the two waves can be formally seen as a third wave s_3 of angular frequency $\omega + \Delta\omega/2$ modulated by a long wave s_4 of angular frequency $\Delta\omega/2$. Then, if the propagation time is not too large (in the depicted case, approximately lower than four periods $2\pi/\omega$), the modulation effect remains negligible and $s_1 + s_2$ can be accurately approximated by s_3. Since a real ocean surface is formed by the superposition of an infinite number of elementary waves with $\Delta\omega \to 0$, a higher frequency resolution (thus a higher number of wave components) in the search for an accurate solution to the inverse problem (III.2) will always be required when the spatio-temporal extent of the assimilated data increases.

Attempts to automatize the determination of the optimal value of N (i.e., the lowest value to allow a consistent inversion) have not been successful. Here is how we proceed. We tried to find an optimal characteristic wavenumber discretization parameter $\Delta\tilde{k}$ for the prescribed frequency

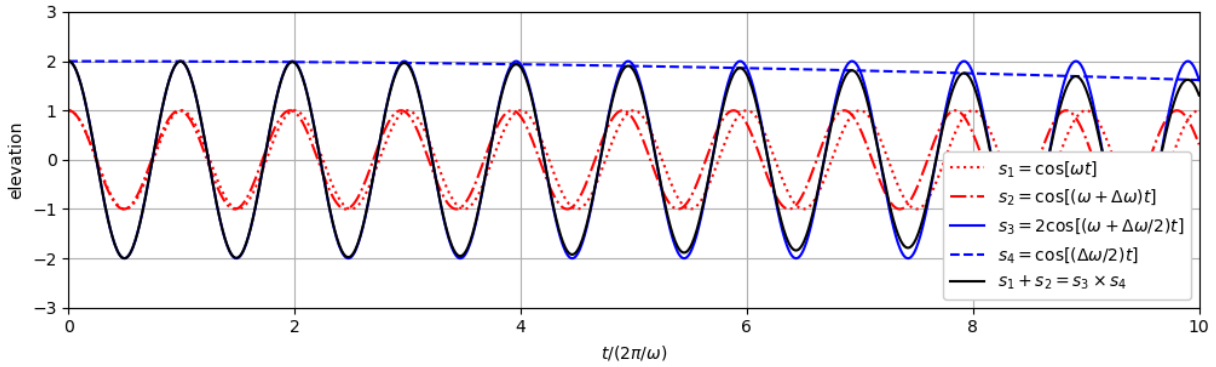

Figure V.22: Illustration of the modulation effect between two cosine waves s_1 and s_2 of very close angular frequencies ω and $\omega + \Delta\omega$, with $\Delta\omega = \omega/50$, respectively, and of the validity of the approximation of the sum of these waves by another wave of angular frequency $\omega + \Delta\omega/2$.

grid of our solution vector $\boldsymbol{p}$. The characteristic discretization parameter have been defined by

$$\Delta\tilde{k} = \frac{1}{N}\sum_{n=1}^{N} \Delta k_n \sqrt{2S_{\text{norm}}\left(k_n\right)\Delta k_n},$$

where $S_{\text{norm}} = S/m_0$ is the normalized wave spectrum, which was assumed to be known before the reconstruction. Then, the value of N have been selected such that

$$\Delta\tilde{k} \approx \min\left(\Delta k_{T_a}, \Delta k_{L_o}\right),$$

in which $\Delta k_{T_a} = 4\pi^2/\left(gT_a^2\right)$ and $\Delta k_{L_o} = 1/L_o$ are characteristic wavenumber steps related to the time and space scales of the inverse problem, respectively. Δk_{T_a} (Δk_{L_o}) corresponds to the wavenumber step of a time (space) DFT of a signal of length T_a (L_o). This method tends to yield values of N that are too high compared to the what appears to be necessary for an accurate inversion. For the results reported this section, we thus adapted N to avoid an unnecessary increase of the calculation time. For the value of $T_a \approx 7T_p$ and $\tau \approx 0.07T_p$ that are used hereafter, we selected $N = 50$.

V.4 Prediction Results and Discussion

Applications of the reconstruction and forecasting algorithms to cases of table V.2 (i.e., long-crested waves) are presented in the following and the accuracy of the wave field forecast is discussed, in particular, in terms of its sensitivity to the linear or nonlinear methods used.

V.4.1 Wave Group Analysis

All cases in table V.2 correspond to sea states generated using a JONSWAP spectrum with identical peak period $T_p = 10$ s (at full scale) and peakedness $\gamma = 3.3$, but a different significant wave height H_s and, hence, characteristic wave steepness H_s/λ_p. In both the physical and numerical wave tanks, these sea states are generated using the same set of random phases, so time series of surface elevations should be similar, except for small changes in amplitude due to nonlinear effects, proportional to H_s.

V.4.1.1 Surface Elevation Series

In the following, we analyze the prediction error for a group of 8 waves of elevation on the order of $\eta \sim 0.5H_s$ approximately centered at $t = 113T_p$, recorded/simulated at wave probe 22 for cases A to H (see figure V.13). The data used in the prediction algorithms was selected

for the prediction zone to span $t \approx 108T_p$ to $118T_p$ at the location of probe 22. Figure V.23 shows time series of surface elevations for these wave groups in cases A, E and H, compared to predictions of the linear (LWT), linear corrected (LWT-CDR), and nonlinear (ICWM) models, using experimental (a, c, e) and numerical (b, d, f) data. For each case, only small differences

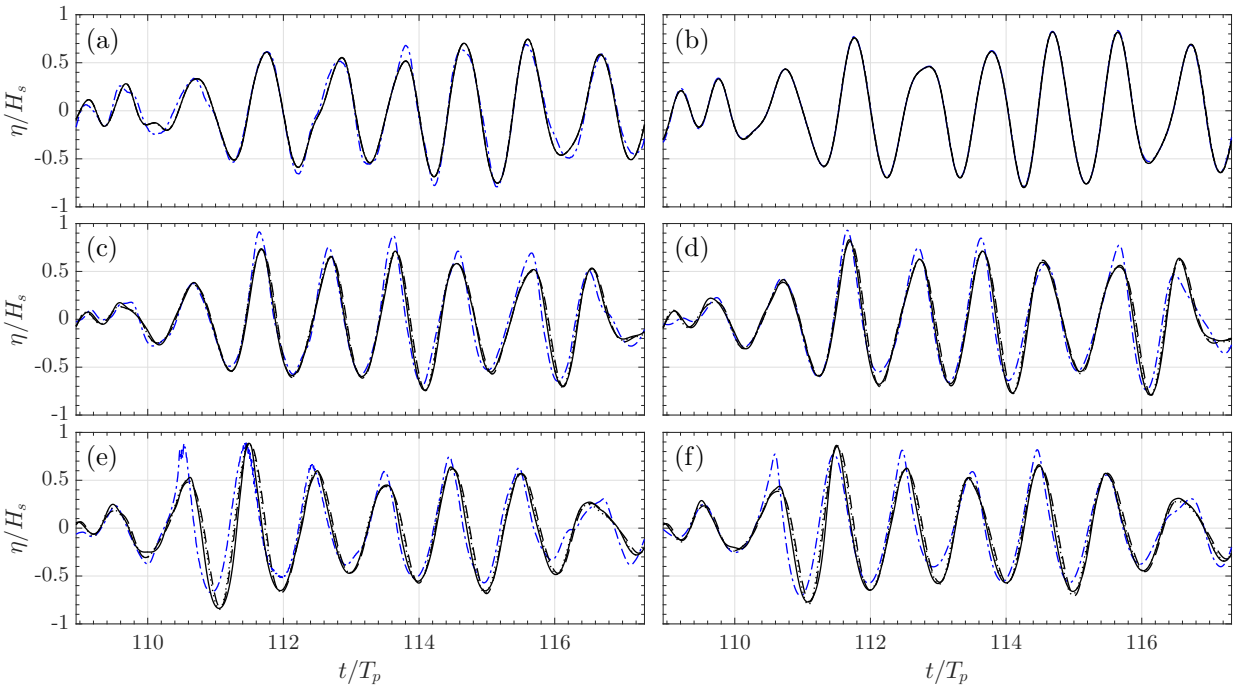

Figure V.23: Time series of surface elevations measured/simulated at probe 22 for a group of 8 waves (-·-·-) for cases: (a, b) A, (c, d) E and (e, f) H, of increasing nonlinearity. Predicted surface elevations are shown for the: linear (LWT, - - - -), corrected linear (LWT-CDR, ·····), and nonlinear (ICWM, ——) models. Left (a, c, e): experiments; right (b, d, f): simulations.

due to experimental noise can be seen between the experimentally and numerically generated reference surfaces. While there is an overall agreement between the reference and predicted surfaces, differences in wave phase and elevation increase with wave steepness, due to cumulative effects of nonlinearity during wave propagation.

Accordingly, for the smallest wave steepness (case A), all three models predict the same surface elevation, in good agreement with references, particularly for numerical data (b) for which predictions almost perfectly overlap the HOS solution, but predictions become increasingly different between the three algorithms, the larger the characteristic wave steepness. Although differences do not appear visually large, this is more pronounced for the algorithm based on ICWM, which, as will be shown next using various prediction error metrics, provides the most accurate prediction.

V.4.1.2 Indicators Based on Cross-Correlation

For the deterministic comparison of the surface elevations associated with the analyzed wave group, we make use of the cross-correlation between time series corresponding to the predicted and the measured surface elevations, which provides a correlation factor C as a function of a time-lag $\mathcal{T}$. The maximal value of the correlation factor and its corresponding time-lag can be interpreted as the correspondence in terms of shape and amplitude of the two elevations, and as an estimate of the time shift between the two elevations, respectively. The cross-correlation is defined by

$$C(\mathcal{T}) = \frac{1}{t^{\max} - t^{\min}} \int_{t^{\min}}^{t^{\max}} \eta^*(t) \times \eta_{\mathrm{r}}^*(t+\mathcal{T})\,\mathrm{d}t, \tag{V.7}$$

where $t^{\min,\max}$ are the prediction zone boundaries and $\eta^*(t) = \eta(t)/\sigma_\eta \left(t^{\min} \leq t \leq t^{\max}\right)$ is the normalized free surface elevation (similarly, $\eta_r^* = \eta_r/\sigma_{\eta_r}$ for the reference surface).

V.4.1.3 Influence of Wave Modeling

Differences between the reference (measured or simulated with HOS) surface elevations and those predicted by the three algorithms are quantified by their maximum cross-correlation $\max(C)$ (i.e., normalized convolution, equation (V.7)) and corresponding time-lag $\mathcal{T}_s = \arg\max(C)$. Both parameters are shown in figure V.24 for all cases in table V.2, based on time series measured or simulated at the location of wave probe 22 (e.g., figure V.23). The former quantifies the accuracy

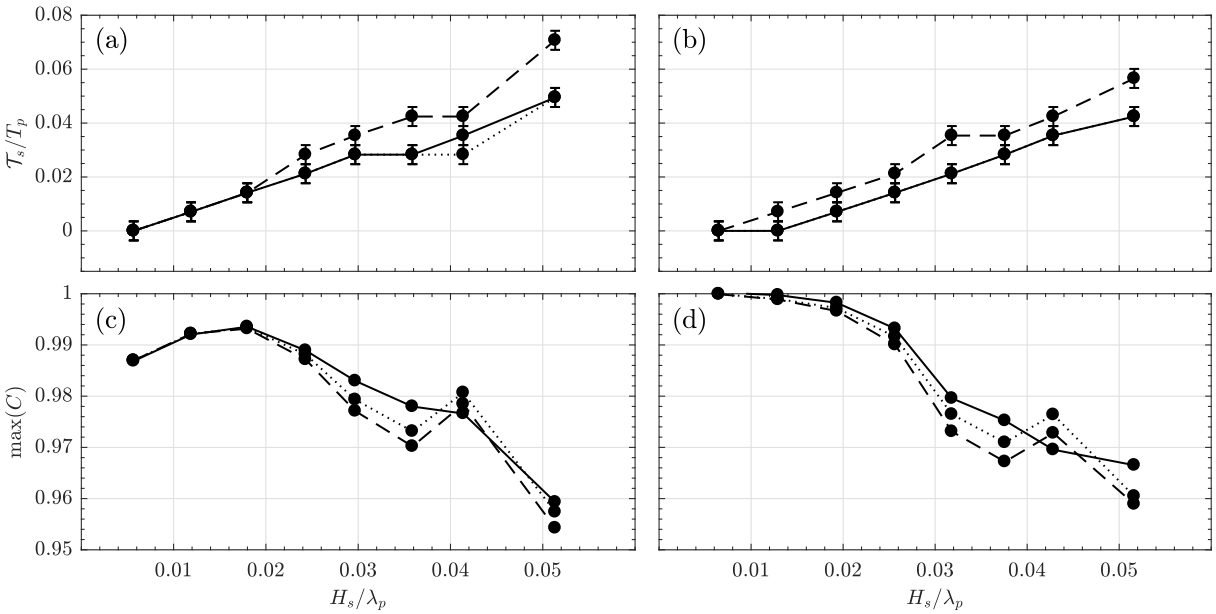

Figure V.24: Normalized time-lag (a, b) and maximum cross-correlation (c, d) computed for cases in table V.2, based on reference time series measured (left column)/simulated with HOS (right column) at the location of wave probe 22 for wave groups similar to those in figure V.23, based on: linear (LWT, - - - -), corrected linear (LWT-CDR, ·····), and nonlinear (ICWM, ——) prediction algorithms. Error bars in (a, b) result from the resolution of the time series.

of the prediction in terms of wave shape and amplitude, while the latter quantifies the time-shift of the predicted signal compared to the reference signal. Figures V.24a,b show that, for all prediction algorithms, time-lag increases with wave steepness (i.e., nonlinearity), from 0 for the smallest steepness to a few percent of T_p for the largest one, consistent with the expected effects of nonlinearity. As seen for instance in the time series of figure V.23, LWT yields the largest time-lags compared to the nonlinear models. LWT-CDR, which includes a phase shift correction, provides a time-lag very close to that of ICWM, particularly for the simulated data, indicating that the nonlinear phase shift prevails over the nonlinear wave geometry represented in the latter model for these cases. Also, the rate of increase of time-lag with wave steepness is similar whether or not the nonlinear phase shift is included in the model. This result is unexpected since this phase shift its due to nonlinear amplitude dispersion, which is function of wave steepness.

Figures V.24c,d show the maximum cross-correlations for the same cases. Consistent with the larger time-lag, $\max(C)$ mostly decreases, the larger the wave steepness, to reach a minimum of 96% for the largest wave steepness. Except for case G, which is discussed below, the maximum cross-correlation is larger using ICWM, which is expected since only this model is able to represent nonlinear wave geometry. The abnormal behavior of case G, which is seen in both the experimental and numerical data, likely results from a significant increase in wave breaking events within the considered wave group for this case. Note, for case H, which has an

even larger steepness, wave phases were such that breaking was not as widespread as for case G. Wave breaking affects wave geometry in a non-trivial manner and is not represented in ICWM. In the case of the wave group considered here, for some unknown reason, it appears that broken waves are better represented in the linear model than using ICWM.

V.4.1.4 Influence of Propagation Distance

Finally, figure V.25 shows results similar to figure V.24, using ICWM for numerical or experimental data, at wave probes 20, 21 and 22. Observations are acquired (and reconstructed) at

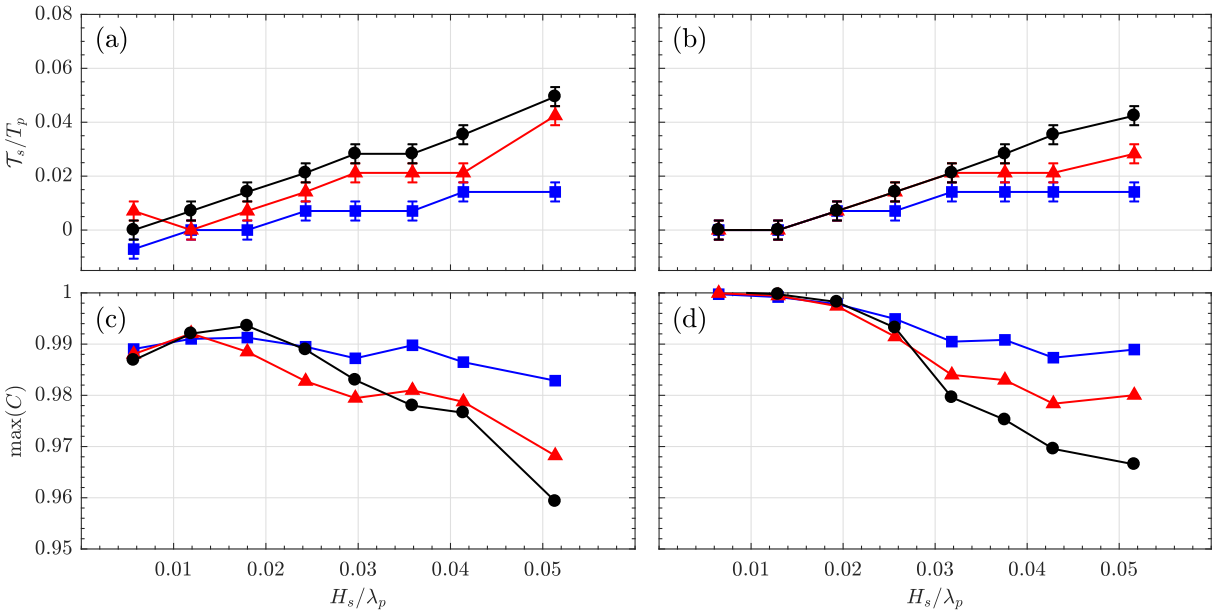

Figure V.25: Normalized time-lag (a, b) and maximum of cross-correlation (c, d) for the nonlinear (ICWM) predictions, at the location of probe 20 (■), 21 (▲) and 22 (●). Left (a, c): experiments; right (b, d): simulations. Error bars on (a, b) correspond to the time resolution of the time series.

wave probe 20, which is the last probe used in observations, and predictions are made at the other 2 probes, which are increasingly distant from it (figure V.2). For both the experimental and numerical data, the time-lags and their rates of increase with wave steepness are lower for prediction locations closer to the observation probe (figures V.25a,b). This results from inaccuracies in nonlinear wave propagation modeled by ICWM, which yield increasing differences in predicted surface elevations with time or space traveled, compared to the reference data. Consistent with this observation, figures V.25c,d show that, for steepnesses larger than $\sim 2.5\%$, the maximum cross-correlation decreases as the distance of the prediction probe to the observation location increases. Due to the disturbances related to the experimental conditions reported in section V.2.3, the values of $\max(C)$ in case of experimental data are generally lower than their numerical counterparts, especially for cases of small steepness for which the relative amplitude of the disturbances is high compared to the amplitude of the waves of interest. Also, while the numerical values of $\max(C)$ seem to reach a limit for the largest steepnesses, the experimental results do not follow that trend and continue to decrease significantly.

V.4.2 Instantaneous Misfit of Wave Prediction

We investigate next the evolution of the instantaneous misfit $\mathcal{E}(x,t)$ of the wave prediction for case E, which corresponds to a moderate steepness, although nonlinear effects already have a marked influence on the wave field dynamics. For both experimental and numerical cases, we compare the misfit obtained using the LWT, LWT-CDR and ICWM prediction algorithms.

Figure V.26 shows the temporal evolution of the wave prediction misfit computed using different algorithms, with respect to data simulated for case E, at both probe 22 (a) and for all probes (b). The misfit at probe 22 is significantly lower within the accessible prediction zone,

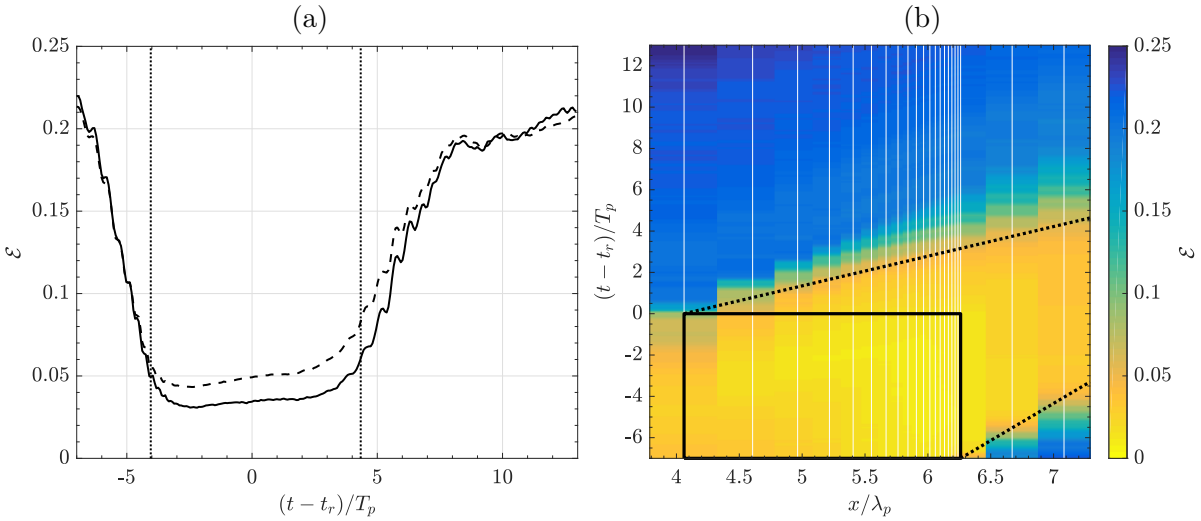

Figure V.26: Time evolution of wave prediction misfit $\mathcal{E}$, computed with respect to data simulated for case E: (a) at probe 22 using the linear (LWT, ----) and nonlinear (ICWM, ——) prediction algorithms; (b) for all probes using the nonlinear (ICWM) prediction algorithm (black rectangle encompasses assimilated observations; data was collected at probes' x-locations (vertical lines) and interpolated using the nearest neighbor method). In each subfigure, dotted lines (·····) mark the prediction zone boundaries.

$\left[t^{\min}, t^{\max}\right]$ (figure V.26a), reaching a minimum value of about 3.5% for ICWM, compared to about 4.5% for LWT, whose misfit is consistently about 30% larger than that of ICWM. Within the time prediction zone, the error gradually slightly increases due to the limited physics represented in both wave models. In the spatio-temporal domain (figure V.26b), ICWM's misfit is lowest within the theoretical prediction zone, reaching a maximum of about 7% along its boundary. As the x-location of the wave probe increases, the misfit gradient decreases across the prediction zone upper boundary $t^{\max}$, or in other words the transition of the misfit values from within to outside the prediction zone becomes more diffused, which is due to the dispersion of the assimilated information. More specifically, as detailed in section III.3, the energy associated with the reconstructed wave components disperses as x increases, gradually limiting the physical description of the wave field. Even within the spatio-temporal region corresponding to the observations (black rectangle in figure V.26b), the misfit is non-zero since observations are discrete rather than continuous samples of elevations. Hence, the reconstructed elevation (nowcast) is always an estimate of the reference solution. Note that the accessible prediction horizon in the depicted configuration is $t^{\max} - t_r \approx 3.7T_p$ and $4.3T_p$ at probes 21 and 22, respectively, and is expected to further increase at larger distances (at the expense of the prediction accuracy). Then, from the location where the beginning of the prediction zone $t^{\min}$ matches the reconstruction time t_r, the accessible horizon starts decreasing.

For the same case E, figure V.27 shows for both experimental or simulated reference data, the spatio-temporal evolution of the ratio of the nonlinear (ICWM) to linear (LWT; a, b) or linear with corrected dispersion (LWT-CDR; c, d) wave prediction misfit. The purpose of computing these error ratios is to characterize the improvements pertaining to ICWM relative to the other wave models. For the simulated data (figure V.27b), the misfit is reduced by up to $\sim 35\%$ within the prediction zone when using ICWM instead of LWT. Compared to LWT-CDR (figure V.27d), the reduction is smaller and mostly limited to the spatio-temporal region of assimilated data (within the solid box), particularly where the wave probe density is larger. Outside of this region ($t > T_a \approx 7T_p$; time propagation) or at the location of probes 21 and 22 (space propagation),

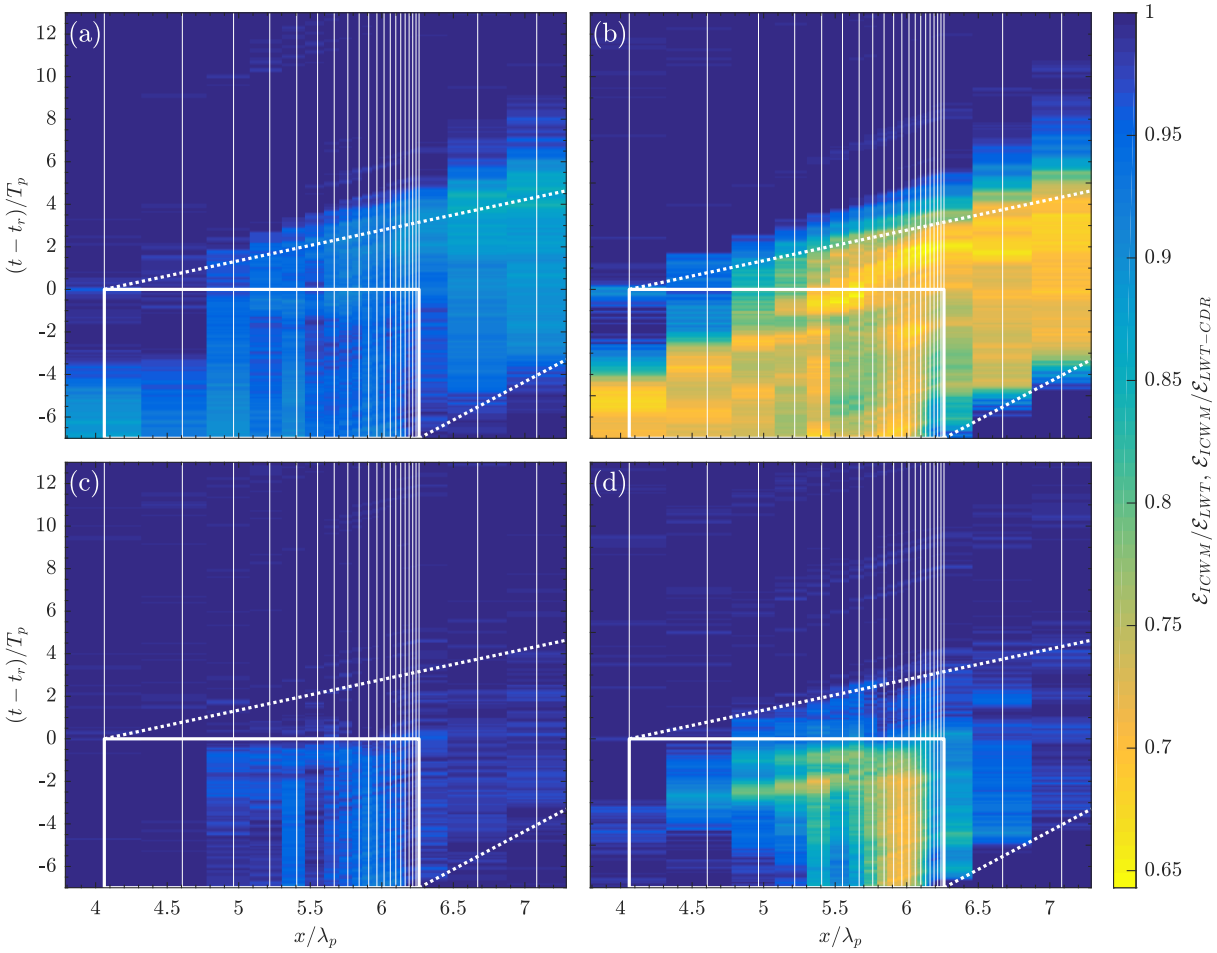

Figure V.27: Spatio-temporal evolution of the ratio of the nonlinear (ICWM) to (a, b) linear (LWT) and (c, d) linear with corrected dispersion relation (LWT-CDR) wave prediction misfit $\mathcal{E}$, with respect to: (a, c) experimental; and (b, d) simulated reference data for case E. Lines denote prediction zone boundaries (·····) and domain of assimilated data (——). Data was collected at probes' x-locations (vertical lines) and interpolated using the nearest neighbor method.

the misfit ratio rapidly approaches one, indicating that the improvement achieved using ICWM rather than LWT-CDR becomes negligible. For the experimental data (figures V.27a,c), similar patterns are observed, but the improvement achieved using ICWM is not as pronounced as for simulated data.

These results indicate that the improved representation of nonlinear wave geometry using ICWM mostly affects the accuracy of the reconstructed part of the wave field. Once reconstructed waves are propagated to the prediction zone, the nonlinear phase shift, which is corrected in LWT-CDR to the same level as in ICWM, becomes the main source of error and effects of nonlinear wave geometry become negligible compared to it. The wave models are parameterized to provide a relevant and consistent approximation of the wave field over the entire region covered by the observations. Hence, while the reconstructed wave field is constrained to fit the measurements, when waves are propagated to issue a prediction, only their propagation properties featured in the models come into play.

V.4.3 Influence of Experimental Noise on Wave Prediction

The prediction misfits based on numerical and experimental reference data are compared in figure V.28, at probes 20, 21, and 22, for all cases listed in table V.2, using the linear (LWT) or nonlinear (ICWM) algorithms. To better assess the effect of experimental noise on the prediction misfit,

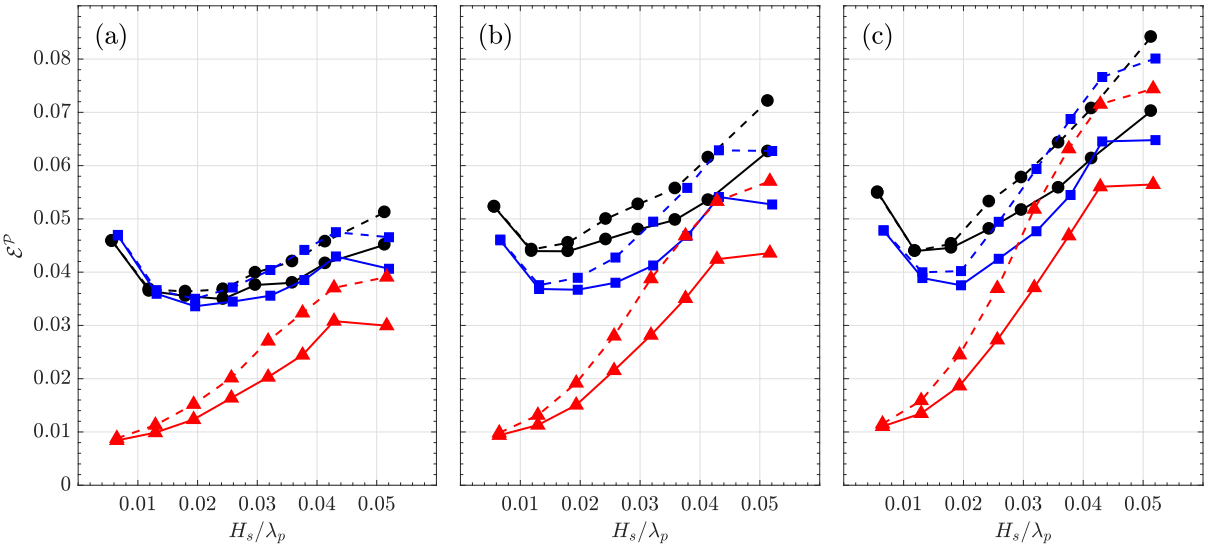

Figure V.28: Prediction misfits for all cases in table V.2, using the linear (LWT, ----) and nonlinear (ICWM, ——) algorithms, based on: experimental (●), numerical (▲), and noisy numerical (■) reference data, at probes (a) 20, (b) 21 and (c) 22.

a "noisy numerical dataset" was created by adding to the numerical data a noise signal having the same spectral shape (or NSR) as that analyzed for the experiments (figure V.15), scaled by the measured noise amplitude H_n (figure V.14), with independent random phases for each wave probe. As would be expected, for both algorithms, the prediction misfits are larger at all wave probes using experimental data, as compared to noise-free numerical data, particularly for cases with a lower steepness, which have relatively larger noise levels. Using the noisy numerical data, however, prediction misfits increase to nearly match those of the experimental data. This indicates that the noisy numerical data is consistent with the experimental data and provides a *digital twin* of experiments that explains, for the most part, differences observed between predictions issued for experimental and noise-free numerical data.

V.4.4 Application to Remote Sensing: Free Surface Slope Prediction

In the free surface elevation predictions described above, the nonlinear phase correction was responsible for the main relative improvement in prediction misfit, rather than the nonlinear wave geometry represented in ICWM. While for many ocean engineering applications predicting instantaneous free surface elevations is most important, such as when computing wave forces or runup on structures, or controlling a wave energy converter, in some applications such as remote sensing the main parameter of interest is free surface slope, which governs the backscattered signal to the radar or optical sensor used (e.g., Nouguier *et al.*, 2014, 2010). Hence, in the following, we quantify the improvement in free surface slopes representation of achieved using ICWM rather than LWT-CDR. More specifically, at the location of wave probe 20, we analyze the evolution as a function of wave steepness of the maximum prediction misfit ratio $\mathcal{E}^{\mathcal{P}}_{ICWM}/\mathcal{E}^{\mathcal{P}}_{LWT-CDR}$, for both surface elevation and slope. As the distance between probes 19 and 20 is small ($\sim 0.02\lambda_p$), the wave surface slope can be approximated by

$$s\left(t\right) = \frac{\eta\left(x_{20},t\right) - \eta\left(x_{19},t\right)}{x_{20} - x_{19}},$$

where η denotes the reference or predicted surface elevation and $x_{19,20}$ the location of probe 19 and 20. We calculate the slope prediction misfit by replacing η with s in equation (IV.2) for the determination of $\mathcal{E}^{\mathcal{P}}$ (equation (V.6)).

Results in figure V.29 show a consistently lower misfit ratio for the slope prediction, whether experimental or numerical data is used (although in the latter case the ratios are lower in abso-

lute value), indicating that, unlike with surface elevation predictions, geometrical nonlinearities included in ICWM provide a significant improvement for predicting surface slopes.

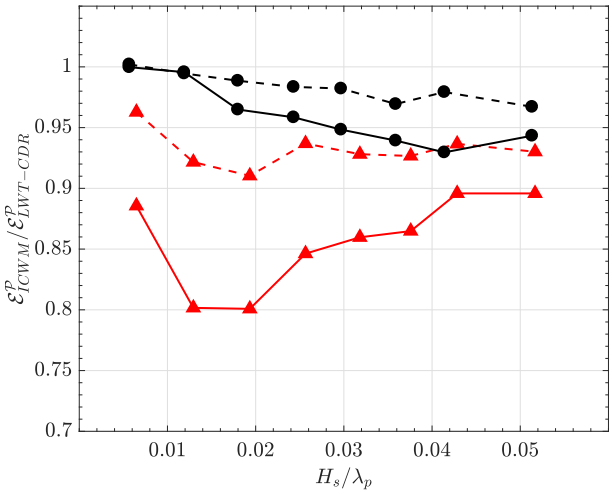

Figure V.29: Ratio of the nonlinear (ICWM) to the linear with corrected dispersion relation (LWT-CDR) prediction misfit for the surface elevation (----) and the surface slope (——) at the location of probe 20 using experimental (●) and numerical (▲) reference data.

Further, in ICWM, geometrical nonlinearities are second-order in wave steepness whereas nonlinear phase shifts are third-order (Guérin *et al.*, 2019; Nouguier *et al.*, 2009, 2010). Accordingly, in figure V.29, the improved performance of ICWM for predicting surface slopes is much more significant at small wave steepness, when third-order effects are not prominent yet. Additionally, when using numerical data, the misfit ratios for surface elevation and slope predictions become increasingly close, the larger the wave steepness. With experimental data, the residual noise causes higher-frequency surface oscillations that significantly affect the slope calculation and prevent a proper evaluation of the corresponding prediction misfit. Hence, in figure V.29 both elevation and slope misfit ratios are close to one at low wave steepness.

When initializing a nonlinear model based on a superposition of harmonics linearly extracted from a standard wave spectrum (such as here a JONSWAP), nonlinearity in the model equations will cause the generation of higher-frequency wave components that will translate into additional spectral energy at those frequencies from the onset. Likewise, waves generated in a wave tank based on a similar standard spectrum will evolve nonlinearly, which results in energy transfer towards higher frequencies. This energy redistribution in wave spectra is referred to as the *dressing process* (Nouguier *et al.*, 2009). In our applications, it is thus the dressed spectrum that is calculated based on the reference surface, both in experiments and simulations. This is illustrated in figure V.30, which shows the normalized energy density spectra of the free surface, $S^*_\eta = S_\eta f_p / \left(H_s^2/16\right)$, computed for case E over the time interval corresponding to the prediction zone $\left[t^{\min}, t^{\max}\right]$ at the location of probe 20. Similar method and parameters to that of the calculation of the noise spectrum (section V.2.3.2) are used to perform the spectral analysis. Waves are only generated by both experimental and numerical wavemakers in the frequency range $f \in [f_l, f_h]$, with $f_l \approx 0.68 f_p$ and $f_h \approx 2.18 f_p$, the low and high cut-off frequencies, respectively. Within this frequency range (figures V.30a,c), spectral amplitudes agree well with each other for the reference data and all wave models. However, as a result of nonlinearity in the reference data, for $f > f_h$ (figures V.30b,d), energy increases in all cases as compared to the targeted spectrum. More specifically, while for $f > f_h$ spectral amplitudes are lower in all the models than in the reference data, ICWM provides results in much better agreement with the latter, particularly for the numerical reference data, and LWT and LWT-CDR provide similar results, both lower than those of ICWM. These results confirm that high-frequency wave components that predominantly affect wave slope predictions, are more accurately predicted with

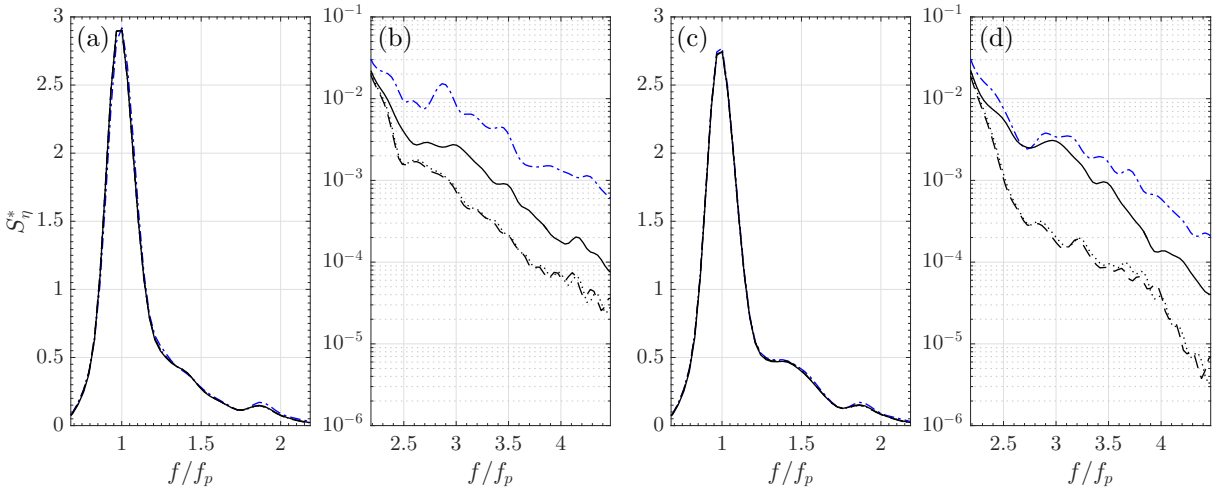

Figure V.30: Normalized energy density spectra of free surface elevations computed for case E over $t \in [t^{\min}, t^{\max}]$, at probe 20, for the: reference data (-·-·-), ICWM model (——), LWT model (- - - -) and LWT-CDR model (·····), in: (a, b) experiments; (c, d) simulations.

the latter model, due to its representation of nonlinear wave geometry (e.g., sharper crests).

Conclusion

We assessed the accuracy of the developed ocean wave prediction algorithm based on three wave models having different nonlinear properties. The experimental campaign led to the generation of data (time series of surface elevation) acquired for both long- and short-crested wave fields using wave probes, in wave tank experiments as well as in simulations of similar experiments with a high-order numerical wave tank. A number of test cases were defined with different directional spreading and characteristic wave steepness, from mildly to more strongly nonlinear. The set of wave probes (physical or numerical) was non-uniformly distributed in space, to mimic the typical uneven sampling made by a lidar camera at grazing incidence.

The prediction method was tested against data generated for long-crested waves. For these cases, residual oscillations observed in the physical wave tank were investigated and found to result from perturbations caused by the wavemaker. These oscillations were identified as the principal cause for observed differences between the experimental and numerical data. Adding the experimental noise to numerical reference data in fact led to similar levels of wave prediction misfit (or accuracy) as when using experimental data.

Our study showed that the prediction accuracy converged with the duration of the assimilated surface elevation time series (i.e., amount of data used). In addition, smaller data acquisition time steps yielded higher convergence rates, and the larger the characteristic wave steepness, the larger the prediction error, in the form of nonlinear phase shifts (related to wave phase velocity) and wave shape discrepancies. ICWM yielded improved predictions within the accessible prediction zone, as compared to LWT, with a maximum prediction error reduction of $\sim 35\%$ for an intermediate wave steepness, $H_s/\lambda_p \approx 3.18\%$ (based on numerical data). In the spatio-temporal region corresponding to the observations, the wave shape asymmetry represented in ICWM improved the surface elevation representation. For waves that had propagated (in space and/or time), the main factor for reducing the prediction misfit was the accurate modeling of nonlinear wave phase velocity. For surface slope predictions, however, the improved representation of wave shape asymmetry in ICWM allowed to achieve an average prediction misfit reduction of 20% over the prediction zone, for a low wave steepness of $\sim$ 1–2% and based on numerical data, as compared to a linear model with a phase speed correction (LWT-CDR). Through a procedure that has been fully validated experimentally, these relative improvements demonstrate the benefits of

using ICWM for the prediction of deterministic fields of the ocean surface.

General Conclusion

During this work, we addressed the problem of deterministic prediction of ocean waves on the basis of datasets acquired by an optical sensor remotely observing the ocean surface. Efforts were focused on the inclusion of relevant nonlinear wave properties in the prediction algorithm, while preserving the computational efficiency of the wave model in order to prevent calculation times from unreasonably growing.

To this end, the Lagrangian approach was selected for the free surface representation at a low computational cost. Developments to the second order in wave steepness were used, which includes both shape nonlinear effects from the modeling of bound waves and a partial modeling of wave resonant effects, affecting the dispersion relation, thus the wave propagation velocity. An extended qualification of the solution properties was proposed and a numerical assessment was performed using a fully nonlinear solution in several configurations. Due to a more suitable choice of the particle location around which the Lagrangian expansion is performed, the proposed wave model, referred to as ICWM, was shown to provide, after a propagation time of several wave periods, a more accurate description of the free surface elevation than the classical second-order Lagrangian solution, even if the explicit terms modeling bound waves and resulting from the interaction between waves of different frequencies are omitted in ICWM.

On the basis of the expected properties of remote optical measurements of surface elevation – which are to yield measurement points with a highly nonuniform spatial distribution and at, a priori, unknown locations –, a methodology for the extraction of the wave information (i.e., amplitudes and phases for a set of frequencies and directions of propagation) from the measurements was developed. Through the minimization of a quadratic cost function, this data assimilation procedure consists in inverting the wave model to provide a description of the observed waves, referred to as the reconstructed wave field. No preprocessing (e.g., interpolation, windowing) of the surface elevation measurements is required. This initial solution is then propagated in time and space with the wave model to provide a prediction. The theoretical validity of the computed solution as well as the definition of the target measurement zone for a desired forecast horizon are addressed based on the available wave information included in the assimilated free surface elevation measurements.

This methodology is then applied to synthetically generated optical measurements (computed from a realistic, fully-nonlinear wave model), providing a numerical demonstration of the validity of the approach. It was proven to issue accurate predictions of both long- and short-crested waves within the theoretically accessible prediction zone. The impact of the spatial non-uniformity of the measurements – which reduces the accuracy of the wave field reconstruction in the region of poor measurement density – decreases as the temporal extent of the assimilated measurements increases. Comparisons between linear and nonlinear predictions showed that the nonlinear features included in ICWM allows improving the prediction accuracy, especially for large wave steepness. Experimentally generated data helped to characterize the improvements pertaining to ICWM relative to a linear model, and to a linear model with a similar wave velocity correction to that of ICWM. Based on data recorded for long-crested waves and their numerical counterpart generated with a digital twin, we highlighted the benefits of the nonlinear wave velocity correction for the accurate representation of the free surface elevation of propagated wave fields,

while the nonlinear wave shape correction is important for the wave field reconstruction and the free surface slope prediction. This provides a demonstration, fully validated experimentally, of the usefulness of using ICWM for the prediction of deterministic fields.

The core of the proposed prediction methodology is the inversion of a physical wave model, consisting, in our case, in the minimization of a quadratic cost function. The resolution of the problem is based on a singular value decomposition plus a regularization, and an iterative process is used for the evaluation of the nonlinear terms. Despite the low computational cost for the spatio-temporal propagation of a known wave field using ICWM, the considered inversion method (required to initialize the wave model), chosen for its robustness and adaptation to non-uniform observations, is not numerically efficient and represents the main limitation of the proposed approach for its use in real-time applications. The calculation time of this approach could possibly be managed for predictions using the linear wave theory through the definition of a better conditioned inverse system or the use of a high-performance numerical framework, but the estimation of wave nonlinearities remains out of reach in real time. Given the efficient numerical properties of ICWM, a reformulation of the inverse problem that effectively uses them could significantly improve the current computational performances, probably at the expense of a preprocessing procedure of the observations or further approximations regarding the wave properties. More generally, the development of inversion strategies that make compatible the non-uniformity and non-periodicity of the wave data with fast numerical methods are worth investigating as this incompatibility represents the major constraint for the real-time initialization of nonlinear wave models, even the most computationally efficient ones. Nevertheless, the proposed nonlinear prediction algorithm can already be used in off-line systems for analysis purposes, providing a more accurate surface representation than the one pertaining to the linear approach. Plus, the limitation of the availability of optical measurements such as provided by a lidar camera does not affect the applicability of the developed methodology, since it can easily be adapted to any kind of wave measurements, either from remote sensors (e.g., radar) or in situ sensors (e.g., wave buoy).

Although the prediction method has been developed within a framework that is fully compatible with directional wave fields, the work presented in this thesis mainly deals with unidirectional wave fields. Considering that the nonlinear properties of ICWM are not affected by directional considerations (i.e., no emergence of nonlinear phenomena in case of a directional distribution), results extended to directional wave fields are expected to be in line with those of unidirectional ones. Nonetheless, in order to thoroughly apprehend the capabilities of ICWM in realistic conditions, a proper assessment of its performances in predicting directional wave fields is required. To this end, the investigation procedure and knowledge provided in this thesis are of great interest and should be used as a basis. In many aspects, the analysis developed herein can be extended to directional wave fields straightforwardly. Questions regarding the quantification of the indicators for the prediction accuracy have to be addressed carefully in order to avoid excessive calculation times due to the necessary large number of surface samples. The experimental datasets described in the last chapter, generated for directional sea states and already partly characterized, provide an appropriate framework for further experimental validation.

Finally, despite the validation of the proposed approach in a wave tank environment, which already provides a solid proof of concept, the use of field data acquired in the open ocean is required to get the full picture of its potentiality.

Appendix A

Assimilation using CWM1 and LWT-CDR

We explicitly formulate here the system for the inversion of the nonlinear models CWM1 (first-order Lagrangian solution) and LWT-CDR (LWT with corrected dispersion relation). We recall that we aim at inverting the system described by equation (III.2), which, after straightforward developments, leads to

$$\begin{cases} \displaystyle\sum_{n=1}^{2N} (\eta_\ell)_{a_m} (\eta_\ell - \overline{\eta}_\ell) = 0, \\ \displaystyle\sum_{n=1}^{2N} (\eta_\ell)_{b_m} (\eta_\ell - \overline{\eta}_\ell) = 0. \end{cases} \tag{A.1}$$

We thus have to determine the derivative terms $(\eta_\ell)_{a_m,b_m}$ to explicitly formulate the system that is solved for the wave parameters $\boldsymbol{p} = \{a_n, b_n\}$.

A.1 System Formulation for CWM1

Following a similar procedure to the one detailed in section II.2.2 for the approximation of ICWM in an Eulerian framework, the Eulerian version of CWM1 surface elevation – including the (second-order) mean water surface level correction – is given by

$$\eta(\boldsymbol{r}, t) = \sum_{n=1}^{N} a_n \cos\Phi_n + b_n \sin\Phi_n + \sum_{n=1}^{N} \frac{1}{2}\left(a_n^2 + b_n^2\right) k_n,$$
$$\text{in which} \quad \Phi_n = \boldsymbol{k}_n \cdot \left[\boldsymbol{r} - \sum_{i=1}^{N} \hat{\boldsymbol{k}}_i \left(-a_i \sin\psi_i + b_i \cos\psi_i\right)\right] - \omega_n t.$$

The above formulation gives

$$\begin{aligned} (\eta_\ell)_{a_m} &= (a_m \cos\Phi_{m\ell})_{a_m} + (b_m \sin\Phi_{m\ell})_{a_m} + a_m k_m, \\ (\eta_\ell)_{b_m} &= (a_m \cos\Phi_{m\ell})_{b_m} + (b_m \sin\Phi_{m\ell})_{b_m} + b_m k_m. \end{aligned}$$

We then calculate

$$\begin{aligned} (a_m \cos\Phi_{m\ell})_{a_m} &= \cos\Phi_{m\ell} - a_m \sin\Phi_{m\ell} (\Phi_{m\ell})_{a_m}, \\ (b_m \sin\Phi_{m\ell})_{a_m} &= b_m \cos\Phi_{m\ell} (\Phi_{m\ell})_{a_m}, \\ (a_m \cos\Phi_{m\ell})_{b_m} &= -a_m \sin\Phi_{m\ell} (\Phi_{m\ell})_{b_m}, \\ (b_m \sin\Phi_{m\ell})_{b_m} &= \sin\Phi_{m\ell} + b_m \cos\Phi_{m\ell} (\Phi_{m\ell})_{b_m}. \end{aligned}$$

The derivatives of the phase term are

$$\begin{aligned}(\Phi_{m\ell})_{a_m} &= k_m \sin\psi_{m\ell},\\ (\Phi_{m\ell})_{b_m} &= -k_m \cos\psi_{m\ell}.\end{aligned}$$

It is now possible to retrieve the modal derivatives of the CWM1 free surface elevation

$$\begin{aligned}(\eta_\ell)_{a_m} &= \cos\Phi_{m\ell} - k_m\left(a_m \sin\Phi_{m\ell} - b_m\cos\Phi_{m\ell}\right)\sin\psi_{m\ell} + a_m k_m,\\ (\eta_\ell)_{b_m} &= \sin\Phi_{m\ell} + k_m\left(a_m \sin\Phi_{m\ell} - b_m\cos\Phi_{m\ell}\right)\cos\psi_{m\ell} + b_m k_m.\end{aligned}$$

Denoting $(\eta_\ell)_{a_m} = P_{m\ell}$ and $(\eta_\ell)_{b_m} = Q_{m\ell}$, system (A.1) turns into

$$\begin{cases}\displaystyle\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n\left(\cos\Phi_{n\ell} + \frac{1}{2}a_n k_n\right)P_{m\ell} + b_n\left(\sin\Phi_{n\ell} + \frac{1}{2}b_n k_n\right)P_{m\ell} = \sum_{\ell=1}^{L}\overline{\eta}_\ell P_{m\ell},\\ \displaystyle\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n\left(\cos\Phi_{n\ell} + \frac{1}{2}a_n k_n\right)Q_{m\ell} + b_n\left(\sin\Phi_{n\ell} + \frac{1}{2}b_n k_n\right)Q_{m\ell} = \sum_{\ell=1}^{L}\overline{\eta}_\ell Q_{m\ell}.\end{cases} \tag{A.2}$$

This system is then recast in a matrix form $\mathsf{A}\boldsymbol{p} = \mathbf{B}$, where, $\boldsymbol{p}$ is our wave model parameters vector containing $2N$ unknown elements

$$p_n = a_n, \quad p_{N+n} = b_n,$$

$\mathbf{B}$ is a vector containing the observation information

$$\mathsf{B}_m = \sum_{\ell=1}^{L}\overline{\eta}_\ell P_{m\ell}, \qquad \mathsf{B}_{N+m} = \sum_{\ell=1}^{L}\overline{\eta}_\ell Q_{m\ell},$$

and A is a $2N \times 2N$ matrix that follows

$$\mathsf{A}_{mn} = \sum_{\ell=1}^{L}\left(\cos\Phi_{n\ell} + \frac{1}{2}a_n k_n\right)P_{m\ell}, \qquad \mathsf{A}_{m,N+n} = \sum_{\ell=1}^{L}\left(\sin\Phi_{n\ell} + \frac{1}{2}b_n k_n\right)P_{m\ell},$$

$$\mathsf{A}_{N+m,n} = \sum_{\ell=1}^{L}\left(\cos\Phi_{n\ell} + \frac{1}{2}a_n k_n\right)Q_{m\ell}, \qquad \mathsf{A}_{N+m,N+n} = \sum_{\ell=1}^{L}\left(\sin\Phi_{n\ell} + \frac{1}{2}b_n k_n\right)Q_{m\ell},$$

System (A.2) is finally solved for optimal wave parameters (a_n, b_n). Since both A and $\mathbf{B}$ depend on wave parameters we make use of the iterative procedure detailed in section III.2.3.

A.2 System Formulation for LWT-CDR

The LWT-CDR surface elevation is given by

$$\begin{aligned}\eta(\boldsymbol{r}, t) &= \sum_{n=1}^{N} a_n\cos\tilde{\phi}_n + b_n\sin\tilde{\phi}_n,\\ \text{in which} \quad \tilde{\phi}_n &= \boldsymbol{k}_n\cdot\boldsymbol{r} - \tilde{\omega}_{\mathrm{Ln}}t.\end{aligned}$$

The above formulation gives

$$\begin{aligned}(\eta_\ell)_{a_m} &= \left(a_m\cos\tilde{\phi}_{m\ell}\right)_{a_m} + \left(b_m\sin\tilde{\phi}_{m\ell}\right)_{a_m},\\ (\eta_\ell)_{b_m} &= \left(a_m\cos\tilde{\phi}_{m\ell}\right)_{b_m} + \left(b_m\sin\tilde{\phi}_{m\ell}\right)_{b_m}.\end{aligned}$$

We then calculate

$$\begin{aligned}
\left(a_m \cos\tilde{\phi}_{m\ell}\right)_{a_m} &= \cos\tilde{\phi}_{m\ell} - a_m \sin\tilde{\phi}_{m\ell}\left(\tilde{\phi}_{m\ell}\right)_{a_m},\\
\left(b_m \sin\tilde{\phi}_{m\ell}\right)_{a_m} &= b_m \cos\tilde{\phi}_{m\ell}\left(\tilde{\phi}_{m\ell}\right)_{a_m},\\
\left(a_m \cos\tilde{\phi}_{m\ell}\right)_{b_m} &= -a_m \sin\tilde{\phi}_{m\ell}\left(\tilde{\phi}_{m\ell}\right)_{b_m},\\
\left(b_m \sin\tilde{\phi}_{m\ell}\right)_{b_m} &= \sin\tilde{\phi}_{m\ell} + b_m \cos\tilde{\phi}_{m\ell}\left(\tilde{\phi}_{m\ell}\right)_{b_m}.
\end{aligned}$$

We now have to calculate the modal derivatives of the phase term $\tilde{\phi}_{m\ell}$, for which we recall here the expression

$$\begin{cases}
\tilde{\phi}_{m\ell} = \boldsymbol{k}_m \cdot \boldsymbol{r}_\ell - \tilde{\omega}_{\mathrm{L}m} t_\ell,\\
\tilde{\omega}_{\mathrm{L}m} = \omega_m + 1/2 \boldsymbol{k}_m \cdot \boldsymbol{\mathcal{U}}_s,\\
\boldsymbol{\mathcal{U}}_s = \sum_{n=1}^{N} \left(a_n^2 + b_n^2\right) \omega_n \boldsymbol{k}_n.
\end{cases}$$

Its derivative yields

$$\left(\tilde{\phi}_{m\ell}\right)_{a_m,b_m} = \left(-\tilde{\omega}_{\mathrm{L}m} t_\ell\right)_{a_m,b_m} \implies \begin{cases}
\left(\tilde{\phi}_{m\ell}\right)_{a_m} = -a_m \omega_m k_m^2 t_\ell,\\
\left(\tilde{\phi}_{m\ell}\right)_{b_m} = -b_m \omega_m k_m^2 t_\ell.
\end{cases}$$

It is now possible to retrieve the expression of the modal derivatives of the LWT-CDR free surface elevation

$$\begin{aligned}
\left(\eta_\ell\right)_{a_m} &= \cos\tilde{\phi}_{m\ell} + k_m \left(a_m \sin\tilde{\phi}_{m\ell} - b_m \cos\tilde{\phi}_{m\ell}\right) a_m \omega_m k_m t_\ell,\\
\left(\eta_\ell\right)_{b_m} &= \sin\tilde{\phi}_{m\ell} + k_m \left(a_m \sin\tilde{\phi}_{m\ell} - b_m \cos\tilde{\phi}_{m\ell}\right) b_m \omega_m k_m t_\ell.
\end{aligned}$$

Denoting $\left(\eta_\ell\right)_{a_m} = P_{m\ell}$ and $\left(\eta_\ell\right)_{b_m} = Q_{m\ell}$, system (A.1) turns into

$$\begin{cases}
\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n \cos\tilde{\phi}_{n\ell} P_{m\ell} + b_n \sin\tilde{\phi}_{n\ell} P_{m\ell} = \sum_{\ell=1}^{L} \overline{\eta}_\ell P_{m\ell},\\
\sum_{\ell=1}^{L}\sum_{n=1}^{N} a_n \cos\tilde{\phi}_{n\ell} Q_{m\ell} + b_n \sin\tilde{\phi}_{n\ell} Q_{m\ell} = \sum_{\ell=1}^{L} \overline{\eta}_\ell Q_{m\ell}.
\end{cases} \tag{A.3}$$

This system is then recast in a matrix form $\mathsf{A}\boldsymbol{p} = \mathbf{B}$, where, $\boldsymbol{p}$ is our wave model parameters vector containing $2N$ unknown elements

$$p_n = a_n, \quad p_{N+n} = b_n,$$

$\mathbf{B}$ is a vector containing the observation information

$$\mathsf{B}_m = \sum_{\ell=1}^{L} \overline{\eta}_\ell P_{m\ell}, \qquad \mathsf{B}_{N+m} = \sum_{\ell=1}^{L} \overline{\eta}_\ell Q_{m\ell},$$

and A is a $2N \times 2N$ matrix that follows

$$\begin{aligned}
\mathsf{A}_{mn} &= \sum_{\ell=1}^{L} \cos\tilde{\phi}_{n\ell} P_{m\ell}, & \mathsf{A}_{m,N+n} &= \sum_{\ell=1}^{L} \sin\tilde{\phi}_{n\ell} P_{m\ell},\\
\mathsf{A}_{N+m,n} &= \sum_{\ell=1}^{L} \cos\tilde{\phi}_{n\ell} Q_{m\ell}, & \mathsf{A}_{N+m,N+n} &= \sum_{\ell=1}^{L} \sin\tilde{\phi}_{n\ell} Q_{m\ell},
\end{aligned}$$

System (A.3) is finally solved for optimal wave parameters (a_n, b_n). Since both A and $\mathbf{B}$ depend on wave parameters we make use of the iterative procedure detailed in section III.2.3.

Bibliography

Abusedra, L. & Belmont, M. 2011 Prediction diagrams for deterministic sea wave prediction and the introduction of the data extension prediction method. *International Shipbuilding Progress* **58**, 59–81.

Allis, M., Peirson, W. & Banner, M. 2011 Application of lidar as a measurement tool for waves. In *Proceedings of the International Offshore and Polar Engineering Conference*, pp. 373–379.

Alpers, W. R., Ross, D. B. & Rufenach, C. L. 1981 On the detectability of ocean surface waves by real and synthetic aperture radar. *Journal of Geophysical Research: Oceans* **86** (C7), 6481–6498.

Aragh, S. & Nwogu, O. 2008 Variation assimilating of synthetic radar data into a pseudo-spectral wave model. *Journal of Coastal Research* pp. 235–244.

Aragh, S., Nwogu, O. & Lyzenga, D. 2008 Improved estimation of ocean wave fields from marine radars using data assimilation techniques. In *The Eighteenth International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers.

Ardhuin, F. 2019 *Ocean waves in geosciences*, Chap. Wave spectra and signficant wave height: theory and measurement, pp. 29–49. open-access online.

Atanassov, V., Rosenthal, W. & Ziemer, F. 1985 Removal of ambiguity of two-dimensional power spectra obtained by processing ship radar images of ocean waves. *Journal of Geophysical Research: Oceans* **90** (C1), 1061–1067.

Barstow, S. F., Bidlot, J.-R., Caires, S., Donelan, M. A., Drennan, W. M., Dupuis, H., Graber, H. C., Green, J. J., Gronlie, O., Guérin, C., Gurgel, K.-W., Günther, H., Hauser, D., Hayes, K., Hessner, K., Hoja, D., Icard, D., Kahma, K. K., Keller, W. C., Krogstad, H. E., Lefevre, J.-M., Lehner, S., Magnusson, A. K., Monbaliu, J., Nieto Borge, J. C., Pettersson, H., Plant, W. J., Quentin, C. G., Reichert, K., Reistad, M., Rosenthal, W., Saetra, O., Schulz-Stellenfleth, J., Walsh, E. J., Weill, A., Wolf, J., Wright, C. W. & Wyatt, L. R. 2005 *Measuring and Analysing the directional spectrum of ocean waves. COST 714; EUR 21367* . COST Office.

Belmont, M. 1995 An extension of nyquist's theorem to non-uniformly sampled finite length data. *International journal of adaptive control and signal processing* **9** (2), 163–181.

Belmont, M. & Horwood, J. 1999 The effect of frequency distribution in sea model spectra on simulations of deterministic sea wave prediction. *International shipbuilding progress* **46** (447), 265–276.

Belmont, M., Horwood, J., Thurley, R. & Baker, J. 2006 Filters for linear sea-wave prediction. *Ocean Engineering* **33** (17), 2332 – 2351.

Belmont, M. R., Baker, J. & Horwood, J. M. K. 2003 Avoidance of phase shift errors in short term deterministic sea wave prediction. *Journal of Marine Engineering & Technology* **2** (2), 21–26.

Belmont, M. R., Christmas, J., Dannenberg, J., Hilmer, T., Duncan, J., Duncan, J. M. & Ferrier, B. 2014 An examination of the feasibility of linear deterministic sea wave prediction in multidirectional seas using wave profiling radar: Theory, simulation, and sea trials. *Journal of Atmospheric and Oceanic Technology* **31** (7), 1601–1614.

Belmont, M. R., Horwood, J. M. K., Thurley, R. W. F. & Baker, J. 2007 Shallow angle wave profiling lidar. *Journal of Atmospheric and Oceanic Technology* **24** (6), 1150–1156.

Belmont, M. R. & Morris, E. L. 1994 Adaptive measurement and signal processing strategies associated with deterministic sea wave prediction. In *Sixth International Conference on Electronic Engineering in Oceanography, 1994.*, pp. 181–188.

Benoit, M., Frigaard, P. & Schäffer, H. A. 1997 Analysing multidirectional wave spectra: a tentative classification of available methods. In *Proceedings of the 1997 IAHR conference*, pp. 131–158.

Birch, R., Fissel, D. B., Borg, K., Lee, V. & English, D. 2004 The capabilities of doppler current profilers for directional wave measurements in coastal and nearshore waters. In *Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No.04CH37600)*, , vol. 3, pp. 1418–1427.

Bishop, C. T. & Donelan, M. A. 1987 Measuring waves with pressure transducers. *Coastal Engineering* **11** (4), 309–328.

Blenkinsopp, C., Mole, M., Turner, I. & Peirson, W. 2010 Measurements of the time-varying free-surface profile across the swash zone obtained using an industrial lidar. *Coastal Engineering* **57** (11), 1059–1065.

Blondel, E. 2009 Reconstruction et prévision déterministe de houle à partir de données mesurées. PhD Thesis, École Centrale de Nantes.

Blondel, E., Bonnefoy, F. & Ferrant, P. 2010 Deterministic non-linear wave prediction using probe data. *Ocean Engineering* **37** (10), 913–926.

Blondel-Couprie, E., Bonnefoy, F. & Ferrant, P. 2013 Experimental validation of non-linear deterministic prediction schemes for long-crested waves. *Ocean Engineering* **58**, 284–292.

Blondel-Couprie, E. & Naaijen, P. 2012 Reconstruction and Prediction of Short-Crested Seas Based on the Application of a 3D-FFT on Synthetic Waves: Part 2 — Prediction. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 5: Ocean Engineering; CFD and VIV, pp. 55–70.

Bonnefoy, F. 2005 Modélisation expérimentale et numérique des états de mer complexes. PhD thesis, Université de Nantes.

Bonnefoy, F., Ducrozet, G., Touzé, D. L. & Ferrant, P. 2010 *Advances in Numerical Simulation of Nonlinear Water Waves*, chap. Time-Domain Simulation of Nonlinear Water Waves Using Spectral Methods, pp. 129–164. World Scientific.

Bonneton, P. & Lannes, D. 2017 Recovering water wave elevation from pressure measurements. *Journal of Fluid Mechanics* **833**, 399–429.

Bonneton, P., Lannes, D., Martins, K. & Michallet, H. 2018 A nonlinear weakly dispersive method for recovering the elevation of irrotational surface waves from pressure measurements. *Coastal Engineering* **138**, 1–8.

Bouferrouk, A., Saulnier, J.-B., Smith, G. H. & Johanning, L. 2016 Field measurements of surface waves using a 5-beam adcp. *Ocean Engineering* **112**, 173–184.

Calvetti, D., Reichel, L. & Shuibi, A. 2004 L-curve and curvature bounds for tikhonov regularization. *Numerical Algorithms* **35** (2-4), 301–314.

Chabchoub, A., Hoffmann, N. P. & Akhmediev, N. 2011 Rogue wave observation in a water wave tank. *Phys. Rev. Lett.* **106**, 204502.

Clamond, D. 2007 On the lagrangian description of steady surface gravity waves. *Journal of Fluid Mechanics* **589**, 433–454.

Clauss, G. F., Klein, M., Dudek, M. & Onorato, M. 2014 Application of Higher Order Spectral Method for Deterministic Wave Forecast. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 8B: Ocean Engineering. V08BT06A038.

Clauss, G. F., Klein, M., Dudek, M., Onorato, M. *et al.* 2015 Deterministic non-linear wave forecast and motion prediction for short-term offshore operations. In *The Twenty-fifth International Ocean and Polar Engineering Conference*. International Society of Offshore and Polar Engineers.

Connell, B. S. H., Rudzinsky, J. P., Brundick, C. S., Milewski, W. M., Kusters, J. G. & Farquharson, G. 2015 Development of an Environmental and Ship Motion Forecasting System. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 11: Prof. Robert F. Beck Honoring Symposium on Marine Hydrodynamics. V011T12A058.

Dalzell, J. 1999 A note on finite depth second-order wave–wave interactions. *Applied Ocean Research* **21** (3), 105–111.

Dankert, H., Horstmann, J. & Rosenthal, W. 2005 Wind- and wave-field measurements using marine x-band radar-image sequences. *IEEE Journal of Oceanic Engineering* **30** (3), 534–542.

Dankert, H. & Rosenthal, W. 2004 Ocean surface determination from x-band radar-image sequences. *Journal of Geophysical Research: Oceans* **109** (C4).

Dannenberg, J., Hessner, K., Naaijen, P., van den Boom, H. & Reichert, K. 2010 The on board wave and motion estimator owme. In *Proceedings of the International Offshore and Polar Engineering Conference*, , vol. 3, pp. 424–431.

Dean, R. G. & Dalrymple, R. A. 1991 *Water Wave Mechanics for Engineers and Scientists*. World Scientific.

Desmars, N., Bonnefoy, F., Grilli, S., Ducrozet, G., Perignon, Y., Guérin, C.-A. & Ferrant, P. 2020 Experimental and numerical assessment of deterministic nonlinear ocean waves prediction algorithms using non-uniformly sampled wave gauges. *Ocean Engineering* **212**, 107659.

Desmars, N., Perignon, Y., Ducrozet, G., Guérin, C.-A., Grilli, S. T. & Ferrant, P. 2018*a* Phase-Resolved Prediction of Nonlinear Ocean Wave Fields from Remote Optical Measurements. In *16èmes Journées de l'Hydrodynamique*. Marseille, France.

Desmars, N., Perignon, Y., Ducrozet, G., Guérin, C.-A., Grilli, S. T. & Ferrant, P. 2018*b* Phase-Resolved Reconstruction Algorithm and Deterministic Prediction of Nonlinear Ocean Waves From Spatio-Temporal Optical Measurements. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 7B: Ocean Engineering, pp. 1–12. V07BT06A054.

Desouky, M. A. & Abdelkhalik, O. 2019 Wave prediction using wave rider position measurements and narx network in wave energy conversion. *Applied Ocean Research* **82**, 10–21.

Dommermuth, D. 2000 The initialization of nonlinear waves using an adjustment scheme. *Wave Motion* **32** (4), 307–317.

Dommermuth, D. G. & Yue, D. K. P. 1987 A high-order spectral method for the study of nonlinear gravity waves. *Journal of Fluid Mechanics* **184**, 267–288.

Ducrozet, G., Bonnefoy, F., Le Touzé, D. & Ferrant, P. 2007 3-d hos simulations of extreme waves in open seas. *Natural Hazards and Earth System Sciences* **7**, 109–122.

Ducrozet, G., Bonnefoy, F., Le Touzé, D. & Ferrant, P. 2016 Hos-ocean: Open-source solver for nonlinear waves in open ocean based on high-order spectral method. *Computer Physics Communications* **203**, 245–254.

Ducrozet, G., Bonnefoy, F., Touzé, D. L. & Ferrant, P. 2012 A modified high-order spectral method for wavemaker modeling in a numerical wave tank. *European Journal of Mechanics - B/Fluids* **34**, 19–34.

Duncan, P. & Drake, K. 1995 A note on the simulation and analysis of irregular non-linear waves. *Applied Ocean Research* **17** (1), 1–8.

Dysthe, K. B. 1979 Note on a modification to the nonlinear schrödinger equation for application to deep water waves. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **369** (1736), 105–114.

Edgar, D., Horwood, J., Thurley, R. & Belmont, M. 2000 Effects of parameters on the maximum prediction time possible in short term forecasting of the sea surface shape. *International Shipbuilding Progress* **47**, 287–301.

Filipot, J.-F., Guimaraes, P., Leckler, F., Hortsmann, J., Carrasco, R., Leroy, E., Fady, N., Accensi, M., Prevosto, M., Duarte, R., Roeber, V., Benetazzo, A., Raoult, C., Franzetti, M., Varing, A. & Le Dantec, N. 2019 La jument lighthouse: a real-scale laboratory for the study of giant waves and their loading on marine structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **377** (2155), 20190008.

Fucile, F., Bulian, G. & Lugni, C. 2018 A probabilistic approach for the quantification of prediction error in deterministic phase-resolved wave forecasting. *Ocean Engineering* **163**, 718–736.

Fujimoto, W. & Waseda, T. 2020 Ensemble-based variational method for nonlinear inversion of surface gravity waves. *Journal of Atmospheric and Oceanic Technology* **37** (1), 17–31.

Gerstner, F. 1809 Theorie der wellen. *Annalen der Physik* **32** (8), 412–445.

Giron-Sierra, J. M. & Jimenez, J. F. 2010 State-of-the-art of wave measurement for ship motion prediction. *IFAC Proceedings Volumes* **43** (20), 295–300, 8th IFAC Conference on Control Applications in Marine Systems.

Gjosund, S. H. 2000 Kinematics in regular and irregular waves based on a lagrangian formulation. PhD Thesis, NTNU.

Gjosund, S. H. 2003 A lagrangian model for irregular waves and wave kinematics. *Journal of Offshore Mechanics and Arctic Engineering* **125**, 94–102.

Gouin, M., Ducrozet, G. & Ferrant, P. 2016 Development and validation of a non-linear spectral model for water waves over variable depth. *European Journal of Mechanics - B/Fluids* **57**, 115–128.

Grilli, S. T., Guérin, C.-A. & Goldstein, B. 2011 Ocean wave reconstruction algorithms based on spatio-temporal data acquired by a flash lidar camera. In *Proceedings of the Twenty-first International Offshore and Polar Engineering Conference (ISOPE2011)*, pp. 275–282. International Society of Offshore and Polar Engineers.

Guérin, C.-A., Desmars, N., Grilli, S. T., Ducrozet, G., Perignon, Y. & Ferrant, P. 2019 An improved lagrangian model for the time evolution of nonlinear surface waves. *Journal of Fluid Mechanics* **876**, 527–552.

Guimarães, P. V., Ardhuin, F., Bergamasco, F., Leckler, F., Filipot, J.-F., Shim, J.-S., Dulov, V. & Benetazzo, A. 2020 A data set of sea surface stereo images to resolve space-time wave fields. *Scientific Data* **7** (1), 1–12.

Hansen, P. C. 2000 The l-curve and its use in the numerical treatment of inverse problems. In *Computational Inverse Problems in Electrocardiology*, pp. 119–142. WIT Press.

Harry, M., Zhang, H., Lemckert, C. & Colleter, G. 2010 3d spatial definition of a water surface. In *Proceedings of the 9th (2010) ISOPE Pacific/Asia Offshore Mechanics Symposium, PACOMS-2010*, pp. 255–261.

Harry, M., Zhang, H., Lemckert, C., Colleter, G. & Blenkinsopp, C. 2018 Observation of surf zone wave transformation using lidar. *Applied Ocean Research* **78**, 88–98.

Hashimoto, N. & Kobune, K. 1988 Estimation of directional spectrum through a bayesian approach. *Coastal Engineering in Japan* **31** (2), 183–198.

Hasselmann, K. 1962 On the non-linear energy transfer in a gravity-wave spectrum part 1. general theory. *Journal of Fluid Mechanics* **12** (4), 481–500.

Hasselmann, K., Barnet, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Krusemann, P., Meerburg, A., Müller, P., Olbers, D., Richter, K., Seil, W. & Walden, H. 1973 Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap). *Deutsche Hydrographische Zeitschrift, Reihe A* **8** (12), 1–95.

Hilmer, T. & Thornhill, E. 2014 Deterministic wave predictions from the wamos ii. In *OCEANS 2014 - TAIPEI*, pp. 1–8.

Hilmer, T. & Thornhill, E. 2015 Observations of predictive skill for real-time deterministic sea waves from the wamos ii. In *OCEANS 2015 - MTS/IEEE Washington*, pp. 1–7.

Howell, G. L. 1998 Shallow water directional wave gages using short baseline pressure arrays. *Coastal Engineering* **35** (1), 85–102.

Huchet, M., Gilloteaux, J.-C., Babarit, A., Ducrozet, G., Perignon, Y. & Ferrant, P. 2018 Deterministic wave prediction using horizontal velocity measurements. In *16èmes Journées de l'Hydrodynamique*. Marseille, France.

Irish, J. L., Wozencraft, J. M., Cunningham, A. G. & Giroud, C. 2006 Nonintrusive measurement of ocean waves: Lidar wave gauge. *Journal of Atmospheric and Oceanic Technology* **23** (11), 1559–1572.

Isobe, M., Kondo, K. & Horikawa, K. 1984 Extension of mlm for estimating directional wave spectrum. In *Proc. Symp. on Description and modeling of directional seas*, pp. 1–15.

Jeans, G., Primrose, C., Descusse, N., Strong, B. & van Weert, P. 2003 A comparison between directional wave measurements from the rdi workhorse with waves and the datawell directional waverider. In *Proceedings of the IEEE/OES Seventh Working Conference on Current Measurement Technology, 2003.*, pp. 148–151.

Kabel, T., Georgakis, C. T. & Rod Zeeberg, A. 2019 Mapping ocean waves using new lidar equipment. In *Proceedings of the Twenty-ninth (2019) International Ocean and Polar Engineering Conference*, pp. 2558–2562. International Society of Offshore and Polar Engineers.

Kalman, R. E. 1960 A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* **82** (1), 35–45.

Klein, M., Dudek, M., Clauss, G., Ehlers, S., Behrendt, J., Hoffmann, N. & Onorato, M. 2020 On the deterministic prediction of water waves. *Fluids* **5** (9).

Köllisch, N., Behrendt, J., Klein, M. & Hoffmann, N. 2018 Nonlinear real time prediction of ocean surface waves. *Ocean Engineering* **157**, 387–400.

Krogstad, H. E. 1988 Maximum likelihood estimation of ocean wave spectra from general arrays of wave gauges. *Modeling, Identification and Control* **9** (2), 81–97.

Kusters, J. G., Cockrell, K. L., Connell, B. S. H., Rudzinsky, J. P. & Vinciullo, V. J. 2016 Futurewaves$^{\text{TM}}$: A real-time ship motion forecasting system employing advanced wave-sensing radar. In *OCEANS 2016 MTS/IEEE Monterey*, pp. 1–9.

Law, Y., Santo, H., Lim, K. & Chan, E. 2020 Deterministic wave prediction for unidirectional sea-states in real-time using artificial neural network. *Ocean Engineering* **195**, 106722.

Le Touzé, D. 2003 Spectral methods for modelling non-linear unsteady free-surface flows. Thesis, Ecole Centrale de Nantes (ECN) ; Université de Nantes.

Lindgren, G. 2009 Exact asymmetric slope distributions in stochastic gauss–lagrange ocean waves. *Applied Ocean Research* **31** (1), 65–73.

Lo, E. & Mei, C. C. 1985 A numerical study of water-wave modulation based on a higher-order nonlinear schrödinger equation. *Journal of Fluid Mechanics* **150**, 395–416.

Lo, E. Y. & Mei, C. C. 1987 Slow evolution of nonlinear deep water waves in two horizontal directions: A numerical study. *Wave Motion* **9** (3), 245–259.

Longuet-Higgins, M. S. 1962 Resonant interactions between two trains of gravity waves. *Journal of Fluid Mechanics* **12** (3), 321–332.

Longuet-Higgins, M. S. 1963 The effect of non-linearities on statistical distributions in the theory of sea waves. *Journal of fluid mechanics* **17** (3), 459–480.

Longuet-Higgins, M. S. & Phillips, O. M. 1962 Phase velocity effects in tertiary wave interactions. *Journal of Fluid Mechanics* **12** (3), 333–336.

Lyzenga, D. R. & Walker, D. T. 2015 A simple model for marine radar images of the ocean surface. *IEEE Geoscience and Remote Sensing Letters* **12** (12), 2389–2392.

Martins, K., Blenkinsopp, C. E., Power, H. E., Bruder, B., Puleo, J. A. & Bergsma, E. W. 2017 High-resolution monitoring of wave transformation in the surf zone using a lidar scanner array. *Coastal Engineering* **128**, 37–43.

Mérigaud, A., Herterich, J., Flanagan, J., Ringwood, J. & Dias, F. 2018 Incorporating wave spectrum information in real-time free-surface elevation forecasting: Real-sea experiments. *IFAC-PapersOnLine* **51** (29), 232–237, 11th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2018.

Mérigaud, A. & Ringwood, J. V. 2019 Incorporating ocean wave spectrum information in short-term free-surface elevation forecasting. *IEEE Journal of Oceanic Engineering* **44** (2), 401–414.

Morris, E., Zienkiewicz, H. & Belmont, M. 1998 Short term forecasting of the sea surface shape. *International shipbuilding progress* **45** (444), 383–400.

Morris, E., Zienkiewicz, H., Pourzanjani, M., Flower, J. & Belmont, M. 1992 Techniques for sea state prediction. In *Proceedings of the 2nd International Conference on Maneuvering and Control of Marine Craft, Southampton. Elsevier, Amsterdam, The Netherlands*, pp. 547–571.

Naaijen, P. & Blondel-Couprie, E. 2012 Reconstruction and Prediction of Short-Crested Seas Based on the Application of a 3D-FFT on Synthetic Waves: Part 1 — Reconstruction. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 5: Ocean Engineering; CFD and VIV, pp. 43–53.

Naaijen, P., van Dijk, R. R. T., Huijsmans, R. H. M. & El-Mouhandiz, A. A. 2009 Real Time Estimation of Ship Motions in Short Crested Seas. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 4: Ocean Engineering; Ocean Renewable Energy; Ocean Space Utilization, Parts A and B, pp. 243–255.

Naaijen, P. & Huijsmans, R. 2008 Real Time Wave Forecasting for Real Time Ship Motion Predictions. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 4: Ocean Engineering; Offshore Renewable Energy, pp. 607–614.

Naaijen, P., Trulsen, K. & Blondel-Couprie, E. 2014 Limits to the extent of the spatio-temporal domain for deterministic wave prediction. *International Shipbuilding Progress* **61** (3-4), 203–223.

Naaijen, P., Van Oosten, K., Roozen, K. & Van't Veer, R. 2018 Validation of a Deterministic Wave and Ship Motion Prediction System. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 7B: Ocean Engineering, pp. 1–8. V07BT06A032.

Naaijen, P. & Wijaya, A. P. 2014 Phase Resolved Wave Prediction From Synthetic Radar Images. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 8A: Ocean Engineering. V08AT06A045.

Neill, S. P. & Hashemi, M. R., ed. 2018 *Chapter 7 - In Situ and Remote Methods for Resource Characterization*, pp. 157–191. Academic Press.

Newman, J. N. 1985 Algorithms for the free-surface green function. *Journal of Engineering Mathematics* **19** (1), 57–67.

Nieto Borge, J. C., Rodríguez, G. R., Hessner, K. & González, P. I. 2004 Inversion of marine radar images for surface wave analysis. *Journal of Atmospheric and Oceanic Technology* **21** (8), 1291–1300.

Nouguier, F., Chapron, B. & Guérin, C.-A. 2015 Second-order lagrangian description of tri-dimensional gravity wave interactions. *Journal of Fluid Mechanics* **772**, 165–195.

Nouguier, F., Grilli, S. T. & Guérin, C.-A. 2014 Nonlinear ocean wave reconstruction algorithms based on spatiotemporal data acquired by a flash lidar camera. *IEEE Transactions on Geoscience and Remote Sensing* **52** (3), 1761–1771.

Nouguier, F., Guérin, C.-A. & Chapron, B. 2009 "choppy wave" model for nonlinear gravity waves. *Journal of Geophysical Research* **114**, 1–16.

Nouguier, F., Guérin, C.-A. & Chapron, B. 2010 Scattering from nonlinear gravity waves: the "choppy wave" model. *IEEE Transactions on Geoscience and Remote Sensing* **48** (12), 4184–4192.

Nwogu, O. 1989 Maximum entropy estimation of directional wave spectra from an array of wave probes. *Applied Ocean Research* **11** (4), 176–182.

Paalvast, M. S. M., Naaijen, P. & Huijsmans, H. R. M. 2014 Propagation of a Radar Modulated Ocean Wave Field. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. Volume 8B: Ocean Engineering. V08BT06A055.

Pedersen, T., Nylund, S. & Dolle, A. 2002 Wave height measurements using acoustic surface tracking. In *OCEANS '02 MTS/IEEE*, , vol. 3, pp. 1747–1754.

Peregrine, D. H. 1983 Water waves, nonlinear schrödinger equations and their solutions. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics* **25** (1), 16–43.

Perignon, Y. 2011 Modélisation déterministe des états de mer - application à la rétrodiffusion d'ondes radar. PhD Thesis, École Centrale de Nantes.

Perignon, Y. L., Bonnefoy, F., Ferrant, P. & Ducrozet, G. 2010 Non-Linear Initialization in Three-Dimensional High Order Spectra Deterministic Sea State Modeling. In *International Conference on Offshore Mechanics and Arctic Engineering*, , vol. 29th International Conference on Ocean, Offshore and Arctic Engineering: Volume 4, pp. 525–532.

Perlin, M. & Bustamante, M. D. 2016 A robust quantitative comparison criterion of two signals based on the sobolev norm of their difference. *Journal of Engineering Mathematics* **101** (1), 115–124.

Phillips, O. M. 1960 On the dynamics of unsteady gravity waves of finite amplitude part 1. the elementary interactions. *Journal of Fluid Mechanics* **9** (2), 193–217.

Pierson, W. J. 1961 Models of random seas based on the lagrangian equations of motion. Technical report. New York University.

Pierson Jr., W. J. & Moskowitz, L. 1964 A proposed spectral form for fully developed wind seas based on the similarity theory of s. a. kitaigorodskii. *Journal of Geophysical Research (1896-1977)* **69** (24), 5181–5190.

Plant, W. J. & Farquharson, G. 2012 Wave shadowing and modulation of microwave backscatter from the ocean. *Journal of Geophysical Research: Oceans* **117** (C8).

Plassman, G. E. 2005 A survey of singular value decomposition methods and performance comparison of some available serial codes. Technical report. NASA, cr-2005-213500.

Prislin, I., Zhang, J. & Seymour, R. J. 1997 Deterministic decomposition of deep water short-crested irregular gravity waves. *Journal of Geophysical Research: Oceans* **102** (C6), 12677–12688.

Qi, Y. 2017 Phase-resolved reconstruction and forecast of nonlinear irregular wave field based on direct numerical simulations. PhD Thesis, Massachusetts Institute of Technology.

Qi, Y., Wu, G., Liu, Y., Kim, M.-H. & Yue, D. K. P. 2018*a* Nonlinear phase-resolved reconstruction of irregular water waves. *Journal of Fluid Mechanics* **838**, 544–572.

Qi, Y., Wu, G., Liu, Y. & Yue, D. K. 2018*b* Predictable zone for phase-resolved reconstruction and forecast of irregular waves. *Wave Motion* **77**, 195–213.

Qi, Y., Xiao, W. & Yue, D. K. P. 2016 Phase-resolved wave field simulation calibration of sea surface reconstruction using noncoherent marine radar. *Journal of Atmospheric and Ocean Technology* **33**, 1135–1149.

Rodenbusch, G. & Forristall, G. 1986 An empirical model for random directional wave kinematics near the free surface. *Proceedings of the 18th Offshore Technology Conference* pp. 137–143.

Rosenbrock, H. H. 1960 An Automatic Method for Finding the Greatest or Least Value of a Function. *The Computer Journal* **3** (3), 175–184.

Seiffert, B. R. & Ducrozet, G. 2018 Simulation of breaking waves using the high-order spectral method with laboratory experiments: wave-breaking energy dissipation. *Ocean Dynamics* **68** (1), 65–89.

Seiffert, B. R., Ducrozet, G. & Bonnefoy, F. 2017 Simulation of breaking waves using the high-order spectral method with laboratory experiments: Wave-breaking onset. *Ocean Modelling* **119**, 94–104.

Serafino, F., Lugni, C., Nieto Borge, J. C. & Soldovieri, F. 2011 A simple strategy to mitigate the aliasing effect in x-band marine radar data: Numerical results for a 2d case. *Sensors* **11** (1), 1009–1027.

Simanesew, A., Trulsen, K., Krogstad, H. E. & Nieto Borge, J. C. 2017 Surface wave predictions in weakly nonlinear directional seas. *Applied Ocean Research* **65**, 79–89.

Socquet-Juglard, H., Dysthe, K., Trulsen, K., Krogstad, H. E. & Liu, J. 2005 Probability distributions of surface gravity waves during spectral changes. *Journal of Fluid Mechanics* **542**, 195–216.

Spell, C., Zhang, J. & Randall, R. E. 1996 Hybrid wave model for unidirectional irregular waves—part ii. comparison with laboratory measurements. *Applied Ocean Research* **18** (2), 93–110.

Steinhagen, U. 2001 Synthesizing nonlinear transient gravity waves in random seas. PhD Thesis, Technical University of Berlin.

Stokes, G. G. 1847 On the theory of oscillatory waves. *Transactions of the Cambridge Philosophical Society* **8**, 197–225.

Sviridov, S. A. 1993 States of arts on laser remote sensing techniques of sea surface roughness in russia. In *Proceedings of OCEANS '93*, pp. I473–I481 vol.1.

Takagi, K., Hamamichi, S., Wada, R. & Sakurai, Y. 2017 Prediction of wave time-history using multipoint measurements. *Ocean Engineering* **140**, 412–418.

Terray, E., Gordon, R. L. & Brumley, B. 1997 Measuring wave height and direction using upward-looking adcps. In *Oceans '97. MTS/IEEE Conference Proceedings*, , vol. 1, pp. 287–290.

Terray, E. A., Brumley, B. H. & Strong, B. 1999 Measuring waves and currents with an upward-looking adcp. In *Proceedings of the IEEE Sixth Working Conference on Current Measurement (Cat. No.99CH36331)*, pp. 66–71.

Tikhonov, A. N. 1963 Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.* pp. 1035–1038.

Trefethen, L. N. & Bau, D. 1997 *Numerical Linear Algebra*. Society for Industrial & Applied Mathematics.

Trulsen, K. 2007 *Weakly Nonlinear Sea Surface Waves — Freak Waves and Deterministic Forecasting*, pp. 191–209. Berlin, Heidelberg: Springer Berlin Heidelberg.

Trulsen, K. & Dysthe, K. B. 1996 A modified nonlinear schrödinger equation for broader bandwidth gravity waves on deep water. *Wave Motion* **24** (3), 281–289.

Trulsen, K. & Dysthe, K. B. 1997 Frequency downshift in three-dimensional wave trains in a deep basin. *Journal of Fluid Mechanics* **352**, 359–373.

Trulsen, K., Kliakhandler, I., Dysthe, K. B. & Velarde, M. G. 2000 On weakly nonlinear modulation of waves on deep water. *Physics of Fluids* **12** (10), 2432–2437.

Trulsen, K. & Stansberg, C. T. 2001 Spatial evolution of water surface waves: Numerical simulation and experiment of bichromatic waves. In *Proceedings of the Eleventh International Offshore and Polar Engineering Conference*, , vol. 3, pp. 71–77. International Society of Offshore and Polar Engineers.

Tsai, C.-H., Huang, M.-C., Young, F.-J., Lin, Y.-C. & Li, H.-W. 2005 On the recovery of surface wave by pressure transfer function. *Ocean Engineering* **32** (10), 1247–1259.

Wada, R., Mondal, R., Sakurai, Y., Takagi, K. *et al.* 2016 Prediction of individual waves based on multi-point measurements. In *Offshore Technology Conference Asia*. Offshore Technology Conference.

Wanis, P. 2013 Design and applications of a vertical beam in acoustic doppler current profilers. In *2013 OCEANS - San Diego*, pp. 1–5.

Welch, P. 1967 The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **15** (2), 70–73.

West, B. J., Brueckner, K. A., Janda, R. S., Milder, D. M. & Milton, R. L. 1987 A new numerical method for surface hydrodynamics. *Journal of Geophysical Research: Oceans* **92** (C11), 11803–11824.

Wheeler, J. *et al.* 1970 Method for calculating forces produced by irregular waves. *Journal of petroleum technology* **22** (03), 359–367.

Wijaya, A. & van Groesen, E. 2016 Determination of the significant wave height from shadowing in synthetic radar images. *Ocean Engineering* **114**, 204–215.

Wijaya, A., Naaijen, P., Andonowati & van Groesen, E. 2015 Reconstruction and future prediction of the sea surface from radar observations. *Ocean Engineering* **106**, 261 – 270.

Work, P. A. 2008 Nearshore directional wave measurements by surface-following buoy and acoustic doppler current profiler. *Ocean Engineering* **35** (8), 727–737.

Wu, G. 2004 Direct simulation and deterministic prediction of large-scale nonlinear ocean wave-field. PhD Thesis, Massachusetts Institute of Technology.

Yoon, S., Kim, J. & Choi, W. 2016 An explicit data assimilation scheme for a nonlinear wave prediction model based on a pseudo-spectral method. *IEEE Journal of Oceanic Engineering* **41** (1), 112–122.

Young, I. R., Rosenthal, W. & Ziemer, F. 1985 A three-dimensional analysis of marine radar images for the determination of ocean wave directionality and surface currents. *Journal of Geophysical Research: Oceans* **90** (C1), 1049–1059.

Zakharov, V. E. 1968 Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Journal of Applied Mechanics and Technical Physics* **9** (2), 190–194.

Zhang, J., Chen, L., Ye, M. & Randall, R. E. 1996 Hybrid wave model for unidirectional irregular waves—part i. theory and numerical scheme. *Applied Ocean Research* **18** (2), 77–92.

Zhang, J., Hong, K. & Yue, D. K. P. 1993 Effects of wavelength ratio on wave modelling. *Journal of Fluid Mechanics* **248**, 107–127.

Zhang, J., Prislin, I., Yang, J. & Wen, J. 1999*a* Deterministic wave model for short-crested ocean waves: Part ii. comparison with laboratory and field measurements. *Applied Ocean Research* **21** (4), 189–206.

Zhang, J., Yang, J., Wen, J., Prislin, I. & Hong, K. 1999*b* Deterministic wave model for short-crested ocean waves: Part i. theory and numerical scheme. *Applied Ocean Research* **21** (4), 167–188.

CENTRALE NANTES

Titre : Reconstruction et prédiction en temps réel de champs de vagues par télédétection optique

Mots clés : Vagues, Houle, Ondes de gravité, Prédiction déterministe, Ondes non-linéaires, Hydrodynamique

Résumé : Les recherches présentées dans le cadre de cette thèse portent sur le traitement de mesures par télédétection optique de la surface océanique en vue de la prédiction de champs de vagues au voisinage d'une structure marine, information cruciale pour l'analyse et le contrôle d'une large gamme d'opérations en mer. Elles visent notamment à inclure, à moindre coût calcul, la modélisation de phénomènes non-linéaires, conservant une représentation réaliste de la surface en cas d'état de mer sévère.
Une approche Lagrangienne faiblement non-linéaire (ICWM), dont les propriétés hydrodynamiques sont étudiées par inter-comparaison avec des modèles de référence, est sélectionnée pour la description de la surface libre. Le problème de prédiction est formulé comme un problème inverse dont le but est de faire correspondre la solution décrite par le modèle de vagues à des observations, composées ici d'élévations de surface mesurées par un capteur lidar synthétique balayant la surface en incidence rasante. Les prédictions résultent de la propagation en temps/espace du modèle ainsi paramétré.
L'applicabilité de la méthodologie est validée à l'aide d'observations de champs de vagues unidirectionnels et directionnels, acquises à des instants différents pour pallier à leur forte non-uniformité spatiale. La comparaison relative des performances de ICWM et de modèles d'ordre inférieur met en évidence les améliorations dues à la modélisation des non-linéarités, notamment celles issues de la correction de la relation de dispersion. Une démonstration de la pertinence de l'utilisation de ICWM est ensuite proposée au moyen d'une procédure entièrement validée expérimentalement en bassin de houle.

Title: Real-time reconstruction and prediction of ocean wave fields from remote optical measurements

Keywords: Ocean waves, Gravity waves, Deterministic prediction, Nonlinear waves, Hydrodynamics, Inverse problem

Abstract: Researches conducted in this thesis address the problem of deterministic prediction of ocean wave fields around a marine structure, a key parameter for the analysis and control of a vast range of offshore operations, on the basis of datasets acquired remotely by an optical sensor. Efforts focus on the inclusion, at low computational cost, of the modeling of nonlinear hydrodynamic phenomena, preserving the reliability the surface representation in case of severe sea state.
A weakly nonlinear Lagrangian approach (ICWM), whose hydrodynamic properties are evaluated by inter-comparison with reference wave models, is selected for the description of the free surface. The prediction problem is then formulated as an inverse problem that aims at fitting the solution described by the wave model to observations, here composed of free surface elevation datasets generated by a synthetic, yet realistic, lidar sensor scanning the ocean surface at grazing angle. Predictions are then issued through the propagation in time and space of the parameterized wave model.
The applicability of the methodology is validated using observations of both unidirectional and directional wave fields, obtained at different instants to compensate for their strong spatial non-uniformity. The relative performance comparison between ICWM and lower-order wave models highlights the improvements due to the modeling of wave nonlinearities, especially those pertaining to the correction of the dispersion relation. A demonstration of the usefulness of ICWM is then provided by means of a procedure that is fully validated experimentally in a wave tank.