The impact of modulational instability on coastal wave forecasting using quadratic models

Gal Akrish a,*, Ad Reniers a, Marcel Zijlema a, Pieter Smit b

a Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands
b Sofar Ocean, San Francisco, USA

Abstract

Coastal wave forecasting over large spatial scales is essential for many applications (e.g., coastal safety assessments, coastal management and developments, etc.). This demand explains the necessity for accurate yet effective models. A well-known efficient modelling approach is the quadratic approach (often referred to as frequency-domain models, nonlinear mild-slope models, amplitude models, etc.). The efficiency of this approach stems from a significant modelling reduction of the original governing equations (e.g., Euler equations). Most significantly, the description of wave nonlinearity essentially collapses into a single mode coupling term determined by the quadratic interaction coefficients. As a result, it is expected that the efficiency achieved by the quadratic approach is accompanied by a decrease in prediction accuracy. In order to gain further insight into the predictive capabilities of this modelling approach, this study examines six different quadratic formulations, three of which are of the Boussinesq type and the other three are referred to as fully dispersive. It is found that while the Boussinesq formulations reliably predict the evolution of coastal waves, the predictions by the fully dispersive formulations tend to be affected by false developments of modulational instability. Consequently, the predicted wave fields by the fully dispersive formulations are characterized by unexpectedly strong modulations of the sea-swell part and associated unexpected infragravity response. The impact of the modulational instability on wave prediction based on the quadratic approach is further demonstrated using existing laboratory results of bichromatic and irregular wave conditions.

1. Introduction

Over coastal waters, incoming ocean waves undergo a dramatic transformation due to the interaction with the bathymetry and due to nonlinear wave–wave interactions. Quasi-linear sinusoidal waves in deep water transform into skewed and asymmetric saw-tooth like shape over shallower water. Ultimately, the waves break and dissipate their energy close to shore.

These complex wave dynamics in coastal waters give rise to many important phenomena nearshore. Examples are, wave setup (e.g., Longuet-Higgins and Stewart, 1964), alongshore currents (e.g., Bowen (1969), Longuet-Higgins (1970), Reniers and Battjes (1997) and Ruessink et al. (2001)), return flow (e.g., Dyhr-Nielsen and Sørensen (1969) and Stive and De Vriend, 1994) and associated sediment transport processes (e.g., Fredsoe and Deigaard, 1992, Van Rijn, 1993) and the generation of infragravity waves which may significantly influence wave run-up and overtopping (e.g., Van Gent, 2001), dune erosion and sediment transport (e.g., Roelvink and Stive, 1989; Roelvink et al., 2009), and harbour oscillations (e.g., Bowers, 1977).

In practice, engineers and governmental agencies require accurate wave parameters over large spatial scales to correctly predict these processes. A well-known modelling approach that allows efficient large scale wave prediction is provided by the so-called quadratic formulation (other names are frequency-domain formulation, nonlinear mild-slope models, amplitude models, etc.). The name “quadratic approach” refers to the nonlinear representation through a single quadratic mode coupling term. Obviously, such a simplification constitutes a significant reduction of the original incompressible and inviscid Euler equations. Therefore, a central question that this study aims to answer is, how reliable is the forecasting of coastal waves as provided by the quadratic formulation?

The earliest quadratic formulation was suggested by Freilich and Guza (1984) based on the classical Boussinesq model of Peregrine (1967). Over relatively shallow waters, this formulation agrees well with observations (Freilich and Guza, 1984). However, due to its weak dispersion assumption, its prediction of wave shoaling from intermediate to shallower waters may deviate significantly (e.g., Agnon

* Corresponding author.
E-mail address: G.Akrish@tudelft.nl (G. Akrish).

https://doi.org/10.1016/j.coastaleng.2024.104502
Received 21 July 2023; Received in revised form 18 February 2024; Accepted 6 March 2024
Available online 11 March 2024
0378-3839/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
et al., 1993). Consequently, efforts were set forward to improve the dispersive behaviour within the quadratic approach, leading to the development of weakly nonlinear Boussinesq models with improved dispersion (e.g., Madsen and Sørensen, 1993). The success of this modelling improvement (e.g., Eldeberky and Battjes, 1996) motivated further developments of fully dispersive and weakly nonlinear models, aiming to improve wave predictions without additional computational costs. Quadratic models with such properties were proposed by Agnon et al. (1993) and Kaihatu and Kirby (1995) and later generalized and further developed by numerous following studies (e.g., Eldeberky and Madsen, 1999, Bredmose et al., 2005, Janssen, 2006, Sheremet et al., 2016, Ardani and Kaihatu, 2019, Kim and Kaihatu, 2021).

There is no doubt that these developments have improved the linear characteristics of the quadratic modelling (i.e., dispersion and shoaling), however doubt arises concerning the improvement in the prediction of the nonlinear evolution. This doubt stems from the fact that the improvement of the linear properties of the quadratic model is accompanied by a change in the quadratic coefficients, and therefore, also by a change in the truncation error obtained due to the modelling reduction associated with the formulation of the quadratic model. An indication for that is presented by the study of Bredmose et al. (2005), which proposes a fully dispersive quadratic model with exact second-order transfer. In other words, bound wave solutions obtained by this model exactly match the second-order bound wave predictions of Stokes theory (e.g., Hasselmann, 1962). Although expected to be promising, Bredmose et al. (2005) observed phase errors of the model predictions by comparing to laboratory experiments. These phase errors were explained by the significant reduction of the amplitude dispersion embedded in this model. However, apart from cumulative phase errors, errors in amplitude dispersion may lead to much more dramatic consequences. In specific, it is well known that the amplitude dispersion has a decisive impact on the evolution of narrow-banded fields (e.g., Lighthill, 2001, Whitham, 1974), controlling energy exchanges through the modulational instability mechanism (Benjamin and Feir, 1967). A well-known example that clearly highlights the effect of improving the linear dispersion relation on the stability characteristics in the context of weak nonlinear modelling is given by the Whitham equation. The Whitham equation was proposed by Whitham (1967) as a generalized Korteweg–de Vries (KdV) equation that incorporates the full linear dispersion relation. Such a generalization is expected to provide a more faithful description of wave field evolution which may also be composed of shorter wave components. However, it is now known that this generalization is accompanied by an unfavourable modification of the modulational instability mechanism triggered over shallower water than expected (the threshold is \( \mu > 1.146 \), which is lower than the usual threshold of \( \mu > 1.363 \), see, e.g., Van Groesen, 1998, Hur and Johnson, 2015, where \( \mu = k_\rho h, k_\rho \) is the characteristic wavenumber of a considered wave field and \( h \) represents the water depth). Thus, Whitham’s generalization turns the modulationaly stable KdV equation into a modulationaly unstable Whitham equation which will predict faulty focusing/defocusing recurrence of narrow-banded fields over regions of relatively small \( \mu \). This erroneous effect may not only lead to false energy exchanges and thus incorrect evolution of the peak frequency components, but may also contaminate the associated development of the infragravity components as a result of incorrect modulation of the wave field.

This section aims to reveal in further detail the nonlinear properties of the quadratic modelling approach and to gain insight into the prediction capabilities of different quadratic formulations to spectrally describe the nonlinear evolution of coastal wave fields, including the development of the sea-swell components (i.e., the primary harmonics and the secondary super-harmonics) and the generation and evolution of the infragravity components (i.e., the secondary sub-harmonics). In total six different formulations are examined and compared. These formulations consist of three Boussinesq formulations and three fully dispersive formulations. The leading order nonlinear properties of these formulations, including their second-order bound wave solutions, amplitude dispersion and stability characteristics are explored and compared in Section 2. Subsequently, Section 3 discusses the impact of these nonlinear properties on the predictive capabilities of coastal wave dynamics using comparisons with existing laboratory results of bichromatic and irregular wave conditions. Finally, conclusions are drawn in Section 4.

2. Model analysis over finite depth

This section aims to gain further insight into the leading order contributions of the nonlinear evolution of a considered wave field as obtained by different quadratic formulations. As in the rest of this study, the considered wave field is assumed to be composed of long-crested waves (i.e., this study is confined to one-dimensional wave propagation). For the purpose of the following analysis, the one-dimensional formulation (ignoring for now the effect of bathymetry changes) of the quadratic model is given as follows:

\[
\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} \left( h \frac{\partial \eta}{\partial x} \right) = -k_\rho \frac{\partial}{\partial x} \sum_{i} V_{ij} \exp(-i\omega_i t),
\]

where \( V_{ij} \) are the quadratic interaction coefficients and \( a_i \) and \( k_i \) are the \( n \)-th complex-amplitude and wavenumber of a time-periodic wave field, represented through the surface elevation function, \( \eta \), as

\[
\eta = \sum_{i} a_i \exp(-i\omega_i t),
\]

where \( \omega_i \) is the \( n \)-th wave angular-frequency and \( x, t \) represent the spatial and temporal coordinates. Further details on the derivation of the quadratic model (1) are provided in Appendix A.

The analysis here focuses on the bound wave solutions and the parameters which control the evolution of narrow-banded seas. To this end, it is assumed here that nonlinear effects are relatively weak, such that Stokes theory can be applied to predict the physical parameters of the wave field (e.g., surface elevation, fluid velocities, etc.). A well-known parameter which provides an indication regarding the validity of Stokes expansion is the so-called Ursell parameter (or Stokes parameter), defined as

\[
U_r = \frac{H^2 L^2}{h^3},
\]

where \( H \) and \( L \) are the characteristic wave height and length of the considered wave field. Based on this definition, weak nonlinearity corresponds to \( U_r < 26 \), for which Stokes theory applies (see Le Méhauté, 1976). This parameter provides a convenient non-dimensional limit for the present analysis.

In total, six models are examined. The first three are the Boussinesq models proposed by Freilich and Guza (1984) (the “consistent shoaling model”), Madsen and Sørensen (1993) and Nwogu (1993). The other three are the fully-dispersive models proposed by Whitham (1967), Kaihatu and Kirby (1995) and Bredmose et al. (2005). Note that Nwogu (1993) and Whitham (1967) actually present time-domain model formulations. However, here those references are used to refer to the corresponding quadratic formulations which are derived based on these original time-domain models. The nonlinear interaction coefficients and also the linear parameters (for the linear dispersion and wave shoaling) of these six quadratic models are summarized in the Supplementary material document.

2.1. Second-order bound waves

The leading nonlinear contributions to the evolution of wave fields, as provided by the quadratic model (1), are extracted here under the conditions of unidirectional propagation over constant and finite depth. Based on the assumption of weak nonlinearity (i.e., \( U_r < 26 \)), the following multiple-scale expansion is being employed (e.g., Holmes, 1995). It is assumed that the waves evolve over two spatial scales. The
fast scale is denoted by \( x_f = x \) and the slow scale is represented by \( x_s = \varepsilon^2 x \), where the small parameter \( \varepsilon \) represents the ratio between the typical wave amplitude and wave length in deep/intermediate water or the ratio between the amplitude and water depth in shallow water. Note that \( x_s \) is defined as \( O(\varepsilon^2) \) variable since resonance condition can only be satisfied between four-waves as dictated by the dispersion relation (this statement contradicts the so-called near-resonance assumption that underlies the development of the quadratic model, see discussion in Appendix A).

In addition to the definition of the two new spatial variables, it is also assumed that the \( n \)th complex-amplitude, \( a_n \), can be written as follows:

\[
a_n = c(a_n^{(1)} + c^2 a_n^{(2)} + c^3 a_n^{(3)} + \ldots)
\]  

(4)

By substituting these assumptions into the quadratic model, (1), one obtains a set of equations, each balancing terms of mutual order. The first three equations are given as follows:

\[
\partial_x a_n^{(1)} - i k_s a_n^{(2)} = 0
\]  

(5)

\[
\partial_x a_n^{(2)} - i k_s a_n^{(3)} = -i \sum Y_{p,r-s} a_p^{(1)} a_r^{(1)}
\]  

(6)

\[
\partial_x a_n^{(3)} - i k_s a_n^{(1)} = -\partial_x a_n^{(1)} - i \sum Y_{p,r-s} (a_p^{(1)} a_r^{(2)} + a_r^{(3)} a_p^{(1)})
\]  

(7)

The interest here focuses on the bound wave solutions which can be obtained through the solution of (6). To this end, knowledge of \( a_n^{(3)} \) is required. This knowledge is achieved through the solution of (5), given by

\[
a_n^{(1)} = A_n \exp(i k_s x_1),
\]  

(8)

where \( A_n \) is a complex amplitude that depends on the boundary condition, say at \( x = 0 \), and the variable \( x_1 \). In addition, the wavenumber, \( k_s \), is obtained through the dispersion relation, \( D(a_n, k_s) = 0 \), as defined by the different model formulations (see details in Supplementary material).

The bound wave solutions are considered for each possible bichromatic pair. To this end, a pair of two incoming primary wave components which are represented by \( a_j \) and \( a_m \), are assumed. Using the linear solution, (8), and assuming that the homogeneous solution of \( a_n^{(2)} \) equals zero, the following expression is obtained:

\[
a_n^{(2)} = \begin{cases} 
2 G_{1,m} a_j a_m, & \text{if } m \neq l \\
G_{1,m} a_n, & \text{if } l = m \\
0, & \text{if } m = -l
\end{cases}
\]  

(9)

where \( n = l + m \), and the so-called quadratic transfer function, \( G_{1,m} \), is given by

\[
G_{1,m} = V_{1,m}/(k_s - k_{jm})
\]  

(10)

where \( k_{jm} = k_j + k_m \). Note that for the case where \( m = -l \), \( a_n^{(2)} \) is zero since \( V_{1,-m} \) is zero (as required to ensure real solutions). However, in the limit for which \( m \) approaches \( l \), the value of \( G_{1,m} \) does not converge to zero, but it converges to the coefficient that corresponds to the set-down associated with a monochromatic wave.

A computed demonstration of the bound solutions is presented in Fig. 1. The figure compares the solutions according to the quadratic model, the values where the waves are normalized by the values of the solutions according to the second-order Stokes theory (expressions of which are given by, e.g., Hasselmann, 1962, Sharma and Dean, 1981, Dalzell, 1999).

Following these results, the model of Bredmose et al. (2005) is clearly preferable over the others with regard to the prediction of the bound waves. In fact, the solutions following Bredmose et al. (2005) match exactly to the solutions according to Stokes theory. The deviation of the other models arises as a result of two factors. The first relates to the nonlinear terms of the underlying time-domain models which construct the quadratic coefficients, \( V_{1,m} \). These terms are subjected to some a priori assumed relation between the depth parameter \( \mu \) and the parameter for nonlinearity \( \varepsilon \). For example, a well-known relation is \( \varepsilon = O(\mu^5) \), which leads to the classical Boussinesq formulation (e.g., Peregrine, 1967). Therefore, under the classical Boussinesq regime, nonlinear terms involving corrections of the dispersion relation are neglected (also see Madsen and Schäffer, 1998). The second factor that leads to deviations in the bound wave predictions is the near-resonance assumption, that is usually taken to derive the quadratic formulation (see Appendix A). The exception in this regard is the model by Bredmose et al. (2005), being formulated through an operator splitting idea suggested by Agnon (1999). The operator splitting idea bypasses the necessity to rely on the near-resonance assumption, and therefore, avoids an additional error in the bound wave solutions (see further discussion in Bredmose et al., 2005 and in Supplementary material).

The exceptional performance of the model by Bredmose et al. (2005) in predicting the bound waves, raises the anticipation of its preferable wave prediction capabilities in general. However, account should also be paid to the fact that the different definitions of the quadratic coefficients and dispersion relation are also accompanied by different definitions for the truncation error (arising as a result of the modelling reduction associated with the formulation of the quadratic model), which may significantly influence the evolution of the waves, and thus, the model forecast.

2.2. Amplitude dispersion

An important nonlinear property that is deteriorated as a result of the modelling reduction introduced by the quadratic model is the amplitude dispersion. The amplitude dispersion is considered here in a complete sense based on the bichromatic case and using (7). Generally speaking, the solutions for \( a_n^{(1)} \), as also found for \( a_n^{(2)} \), consist of bound components which provide high-order corrections to the primary solutions, \( a_n^{(1)} \). However, the products of \( a_n^{(1)} \) and \( a_n^{(2)} \) can also result in resonance interactions as would be clearly understood for the product of the corresponding evolution equations for \( a_n^{(1)} \) (the multiple-scale method allows to keep the solution bounded by balancing such forcing terms with the derivatives of the primary solutions with respect to \( x_1 \) (as given by the right-hand-side of (7)). Namely, by requiring the following solvability condition:

\[
\partial_x a_n^{(1)} = -i \sum Y_{r,n-s} (a_r^{(1)} a_s^{(2)} + a_s^{(1)} a_r^{(2)})
\]  

(11)

which for the considered bichromatic case leads to

\[
\begin{cases} 
\partial_x a_1^{(1)} = C_{1,1} |a_1^{(1)}|^2 a_1^{(1)} + 2C_{1,m} |a_m^{(1)}|^2 a_1^{(1)} \\
\partial_x a_m^{(1)} = C_{m,1} |a_m^{(1)}|^2 a_1^{(1)} + 2C_{m,m} |a_m^{(1)}|^2 a_m^{(1)}
\end{cases}
\]  

(12)

where \( C_{1,1} \) and \( C_{m,m} \) are the cubic interaction coefficients which arise due to the trivial resonant quartets of self and mutual interactions respectively (see further details on the definition of the cubic interaction coefficients for trivial and non-trivial resonant quartets in Appendix B).

The solution for (12) can be obtained explicitly, since the magnitudes of square of \( a_1^{(1)} \) and \( a_m^{(1)} \) are constant in \( x_1 \) (as can be found from the corresponding evolution equations for \( |a_1^{(1)}|^2 \) and \( |a_m^{(1)}|^2 \)). This solution can be written as

\[
\begin{cases} 
|a_1^{(1)}|^2 = A_1^0 \exp(i k_1 x_1 - i (C_{1,1} |a_1^{(1)}|^2 + 2C_{1,m} |a_m^{(1)}|^2) x_1) \\
|a_m^{(1)}|^2 = A_m^0 \exp(i k_m x_1 - i (C_{m,1} |a_m^{(1)}|^2 + 2C_{m,m} |a_m^{(1)}|^2) x_1)
\end{cases}
\]  

(13)

where \( A_1^0 \) and \( A_m^0 \) are constants given at \( x_1 = 0 \). The solution expressed in (13) clearly shows the effect of the amplitudes on the dispersion relation through modifications of the wavenumbers. As an example, if the coefficients \( C_{1,1} \) are positive, then a particular wave component of a given wave field would not only travel faster due to self interaction but also due to the presence of other waves.
Fig. 1. Second-order bound wave solutions normalized by the solutions of Stokes theory (or simply $G_{l,m}/G_{s_{l,m}}$ where $G_{s_{l,m}}$ is the quadratic transfer function according to Stokes theory). Solutions due to sum interactions are given for $k_m \geq k_l$ (upper triangular of each panel), while solutions of sub interactions are provided for $k_m < k_l$ (lower triangular of each panel).

Fig. 2. Model prediction of the amplitude dispersion contributions due to self interactions, $\omega^{(2)}_{l,l}$. The results are normalized by third-order Stokes theory as given by Madsen and Fuhrman (2006).

The values of the amplitude dispersion due to self interactions as given by the different model formulations are compared with those of the third-order Stokes theory over finite depth (e.g., Madsen and Fuhrman, 2006) in Fig. 2. The computed values are not the corrections for the wavenumbers, but the corrections for the angular frequencies. The latter are obtained based on the following transformation (see, e.g., Bredmose et al., 2005):

$$\omega^{(2)}_{l} = c_{g,l} k^{(2)}_{l}$$

where $c_{g,l}$ is the group velocity and $k^{(2)}_{l}$ is the wavenumber correction, defined as

$$k^{(2)}_{l} = C_{l,l} |a^{(1)}_{l}|^2 + 2 \sum_{m} C_{l,m} |a^{(1)}_{m}|^2$$

including contribution due to self interaction, $k^{(2)}_{l,l}$, and sum of contributions due to mutual interactions, $k^{(2)}_{l,m}$.

The results demonstrated in Fig. 2 show a clear over prediction of the contribution due to the self interaction, $\omega^{(2)}_{l,l}$, by most of the models. This over prediction is reduced for increasing values of $\mu$. The one exception here is the formulation by Freilich and Guza (1984), which under predicts the self interaction contribution. Consequently, most of the models will typically predict faster travelling waves in coastal waters, leading to cumulative phase errors as demonstrated by Bredmose et al. (2004) and Bredmose et al. (2005). Ultimately, these results provide further evidence for the implications of the modelling reduction associated with the formulation of the quadratic model and how different assumptions (e.g., different relation between $\mu$ and $\epsilon$) affect the corresponding modelling errors. Similarly, the mutual interactions’ contributions to the amplitude dispersion can be illustrated as well. However, since this illustration is not required for the purposes of the present analysis, it is omitted here.

To summarize, the bichromatic case considered here provides a complete quantitative determination of the amplitude dispersion and emphasizes its direct physical consequences. Particularly, these refer to the change in wave velocity due to self and mutual interactions. Inaccuracy in the prediction of the amplitude dispersion will obviously result in phase errors with respect to, e.g., field/laboratory observations. However, beyond phase errors, inaccurate prediction of the amplitude dispersion may lead to much more dramatic deviations.
These deviations are related to the formation of instability mechanism known as modulational instability, discussed next.

2.3. Modulational instability

Narrowbandened wave fields which propagate over relatively deep water tend to develop modulational instability. Such an instability leads to a relatively rapid growth of a field's modulation at the expense of the carrier wave energy. Coastal waters though are typically shallow, and therefore, coastal wave fields are commonly not affected by modulational instability. Yet, model forecasting of coastal waves may be affected by such a mechanism if it wrongly predicts the amplitude dispersion (see qualitative description by Lighthill, 2001, page 462, on the role of the amplitude dispersion in the development of the modulational instability).

Modulational instability can be analysed by considering the interaction of three waves, namely, by considering the trichromatic case. To this end, the \( \text{O}(\epsilon) \) solution is now considered to be composed of three components, indicated by the carrier component, \( a_0(t) \), and the two side-bands, \( a_1(t) \) and \( a_2(t) \). The corresponding angular-frequencies are defined as \( \omega_0 = \omega_0(1 - \delta_\omega) \) and \( \omega_m = \omega_0(1 + \delta_\omega) \), where \( \delta_\omega \) is \( \Delta \omega/\omega_m \).

Accordingly, the solvability condition (11) for this case results in the following system of equations:

\[
\begin{align*}
\partial_t a_0 &= \left( C_{1,1}|a_0|^2 + 2C_{1,0}|a_0|^2 \right) a_0 + C_{1,m,p} a_m^2 a_{-m} \\
\partial_t a_1 &= \left( C_{p,p}|a_1|^2 + 2C_{m,m}|a_1|^2 \right) a_1 + \left( 2C_{p,m,a_0} a_m a_{-p} \right) \\
\partial_t a_2 &= \left( C_{m,n}|a_2|^2 + 2C_{m,p}|a_2|^2 \right) a_2 + \left( 2C_{p,m,a_0} a_m a_{-p} \right)
\end{align*}
\]

(16)

where the magnitude notation \(| \cdot |^2\) that accompanies the amplitudes is removed here to ease the presentation of the equations. Additionally, the formulation of each of the cubic coefficients, e.g., \( C_{i,j,p,m} \), is defined in Appendix B. The terms in the parenthesis on the right-hand-side of these equations are the amplitude dispersion components due to self and mutual interactions. These terms are real, and consequently, only result in phase corrections. In contrast, the last terms on the right allow exchange of energy among the components (e.g., Phillips, 1967). These last terms arise due to the interactions that satisfies the equality \( 2\omega_m - \omega_m - \omega_0 = 0 \). It could be argued that these last terms should not be included to satisfy the solvability condition, since they correspond to a wavenumber mismatch, i.e., \( 2k_m - k_m - k_m \neq 0 \) However, if the wavenumber mismatch is close enough to zero, that is to say, if \( 2k_m - k_m - k_m = O(\epsilon^2) \), its contribution is absorbed as part of the slow spatial variation. Therefore, for such conditions, these terms should be included as well.

2.3.1. Modulational instability of Stokes waves

Modulational instability concerns with the evolution of weakly modulated wave fields. Accordingly, it is assumed that the side-band amplitudes are small compared to the amplitude of the carrier component. This allows to reduce the above coupled system (16) to the following linear system (the so-called 'pump-wave' approximation, e.g., Craik, 1985):

\[
\begin{align*}
\partial_t a_0 &= \left( 2C_{1,0}|a_0|^2 + C_{1,m,p} a_m^2 a_{-m} \right) a_0 + \left( 2C_{p,p}|a_1|^2 \right) a_1 + \left( 2C_{p,m,a_0} a_m a_{-p} \right) \\
\partial_t a_1 &= \left( 2C_{m,p}|a_2|^2 \right) a_2 + \left( 2C_{p,m,a_0} a_m a_{-p} \right)
\end{align*}
\]

(17)

The solution for \( a_0 \) corresponds to a monochromatic Stokes wave, while the solution of either \( a_1 \) or \( a_2 \) is obtained through

\[
\partial_t a_1 = i(a_1^*(1 - 2C_{1,0}|a_0|^2)) a_1 - \left( 2C_{p,m,a_0} a_m a_{-p} \right)
\]

(18)

which is derived by combining the first and last equations of (17), through substitution for \( a_0 \) or for \( a_m \), where

\[
\begin{align*}
A_1 &= \tilde{A}_1 \exp(-i2C_{1,0}|A_0|^2) \\
A_2 &= \tilde{A}_2 \exp(-i2C_{p,m,a_0} a_m a_{-p} a_m a_{-p}) \\
\end{align*}
\]

(19)

and recall that the definition of the \( j \)th complex-amplitude, \( A_j \), is given by (8). The solution of (15) obeys to the following eigenvalues:

\[
\sigma_{1,2} = -i \left( -\frac{1}{2} (a_1^* - 2A_{1,0}^2|A_0|^2) \pm \sqrt{R} \right)
\]

(20)

where \( R \) is defined as

\[
R = \frac{1}{4} (a_1^* - 2A_{1,0}^2|A_0|^2 - C_{1,m,p} C_{m,0,p} |A_0|^4)
\]

(21)

Therefore, the side-bands are expected to grow when \( R < 0 \), where the growth rate value is provided by \( \text{Im} (\sqrt{R}) \) (\( \text{Im} (\cdot) \) is the imaginary part). In order to gain some insight into the conditions for which modulational instability is expected to emerge, the assumption of small modulation frequency is being employed. More specifically, it is assumed that \( \delta_\omega \ll \omega_0 \) (i.e., \( \Delta \omega \ll \omega_0 \)). Accordingly, the following asymptotic relations are assumed as well:

\[
C_{1,m,p} \sim C_{m,0,p} \sim C_{2}^2
\]

(22)

which can be understood by letting \( C_{1,m,p} \) and \( C_{m,0,p} \) to be defined through the continuous definition (instead of the discontinuous definition applied so far) of \( C_{i,j,p} \) (see details in Appendix B). Furthermore, the assumption that \( \delta_\omega \ll 1 \) also allows to obtain the relation

\[
\sigma_{1,2} \sim -\frac{1}{2} k_p \Delta \omega^2
\]

(23)

Using these relations, \( R \) can be approximated to fourth order (in \( \epsilon \) or \( \delta_\omega \)) as follows:

\[
R \sim \delta_\omega^2 k_p \Delta \omega^2 \left( \frac{1}{4} \delta_\omega^2 k_p \Delta \omega^2 - \sigma_{1,0}^2 \right)
\]

(24)

Consequently, it is found that modulational instability is determined by \( \sigma_{1,2} \), as the other terms are defined positive. Thus, modulational instability can only emerge if \( \sigma_{1,2} \) is positive as well. To get a better understanding of the meaning of \( \sigma_{1,2} \) and its role in the development of modulational instability, its definition in (19) is written more explicitly as follows:

\[
\begin{align*}
\sigma_{1,2} &= (2G_{p,m} V_{m,p} - 2G_{p,p} V_{2,p,p} + 2G_{p,p} V_{2,p,p} - 2G_{p,p} V_{2,p,p}) \\
&\quad + (2G_{p,p} V_{m,p} - 2G_{p,p} V_{2,p,p})
\end{align*}
\]

(25)

Through the assumption of small modulation frequency and that each of the terms above are smooth enough around \( \Delta \omega = 0 \) (see details in Appendix B), the first group of terms on the right-hand side of (25) converges to \( G_{p,p} = 2G_{p,p} V_{2,p,p} \), which is the normalized wavenumber correction due to self interaction. The second converges to a normalized correction due to interaction between the primary component and the component that represents the wave-induced current (e.g., \( 4G_{p,p} V_{0,0} \)). This explicit definition also highlights the opposite roles of these two contributions. The first group of terms leads to a positive contribution (since both the \( V \) terms and the super-harmonic forcing related \( G \) terms are positive), and thus, triggers energy focusing (as qualitatively explained in Lighthill, 2001, page 462), whereas the second group leads to a negative contribution (since the \( V \) terms are positive, but the sub-harmonic forcing related \( G \) terms are negative), and therefore, provides effect of stabilization (as discussed by Whitham, 1974 and Janssen and Onorato, 2007). As a result, the definition of \( \sigma_{1,2} \) under the assumption of small modulation frequency essentially converges to the normalized wavenumber correction as defined by third-order theory (e.g. Whitham, 1974).

A convenient third-order formulation, used here as a reference to evaluate the impact of modulational instability, is the so-called nonlinear Schrödinger equation (NLSE) (e.g., Mei et al., 2005). The NLSE
accurately captures third-order nonlinear effects, and thus, serves an adequate reference for comparison. This allows to quantitatively compare between $\Delta^2$, and the correct normalized wavenumber correction as given by the NLSE (see Eq. (27) in Supplementary material). This comparison is demonstrated for the different quadratic formulations in Fig. 3.

The comparison in Fig. 3 confirms the well-known modulational instability threshold of $\mu > 1.363$ as provided by the NLSE. As expected, the Boussinesq models are not exposed to this instability mechanism. Surprisingly though, this determination does not apply to the formulation by Nwogu (1993), demonstrating relatively weak positive values of $\Delta^2$. On the other hand, the fully dispersive models are strongly affected by modulational instability. The important result revealed here is that these models obey to a much lower instability threshold. Consequently, using these formulation, predictions of narrow-banded fields over coastal waters may be affected significantly by false unstable evolution. This finding raises questions concerning the growth rate of the modulation and the ranges of $\delta_{\omega}$ for which this mechanism is expected to emerge. Answers to these questions are discussed through Fig. 4.

Fig. 4 presents the growth rate of modulational instability as a function of normalized modulation frequency as obtained by the different model formulations for three different values of $\mu$ and using the same steepness value of $\epsilon = 0.13$, where the wave steepness is given by $\epsilon = 2 |A_p| k_p$. The maximum modulation frequency considered is $\Delta \omega_p$. Accordingly, the vertical dashed line in the middle of each panel indicates the value equal to a quarter of the peak frequency.

The results presented by the right panel of Fig. 4 (corresponding to $k_p h = 2.10$) show that the fully dispersive formulations of Bredmose et al. (2005) and Whitham (1967) are subjected to much stronger growth rates than the expected growth rate based on NLSE. The maximum growth rates and modulation frequencies presented by these models are even greater than the ones which are expected for infinitely deep water (see Eqs. (34) and (35) in Supplementary material).

It is remarkable to see that even for the other cases (corresponding to $k_p h = 1.07$ and $k_p h = 1.33$), for which modulational instability is not expected to emerge at all (as emphasized by the zero values of the NLSE, see the left and the middle panels of Fig. 4), the growth rates and modulation frequencies demonstrated using the fully dispersive formulations are significant.

2.3.2. The impact of modulational instability on the evolution of irregular waves

The stability analysis presented so far allows to explain unstable evolution for the three wave interaction case (i.e., monochromatic cases
subjected to small side-bands). However, the significance of modulational instability is not obviously determined for the more general irregular cases. Two different perspectives were proposed to give an indication for the expected impact of modulational instability on the evolution of irregular wave fields. The first, due to Alber (1978), relies on statistical arguments which result in the ratio between the expected modulation scale (a representative scale over which coherent structures due to modulational instability are formed) and the correlation scale (a representative scale over which the field is still correlated). The second, due to Onorato et al. (2001), relies on physical arguments which concern the ratio between nonlinear and dispersion effects (equivalent arguments underlying the Ursell parameter). Ultimately, these two perspectives share the same parameter, commonly referred to as the Benjamin–Feir Index (BFI) (Janssen, 2003). The BFI can be written as

\[ \text{BFI} = \frac{\lambda}{\delta \omega_p} \] (26)

where \( \delta \omega_p \) should be interpreted now as the bandwidth parameter, which defines the ratio between a representative of the spectral bandwidth and the peak frequency, \( \omega_p \) (see specific definition in Supplementary material). Additionally, the wave steepness, \( \epsilon \), should now be interpreted as a characteristic steepness value, which commonly taken as \( \epsilon = k_s H_s/2 \), where \( H_s \) is the significant wave height. Finally, \( \lambda \) is defined such that, using the NLSE parameters, the BFI value equals 1 for the typical deep water wave conditions characterized by a JONSWAP spectrum with peak-enhancement factor of \( \gamma = 3.3 \) and a characteristic steepness of \( \epsilon = 0.1 \) (see details in Supplementary material).

The BFI values as obtained by the NLSE and the different quadratic formulations are examined and compared through the following example. The example considers a shoaling wave field that is characterized by a JONSWAP spectrum with \( T_p = 2.25 \) s and \( H_s = 0.2 \) m (similar to the wave conditions of experiment A2 by Ruessink et al., 2013, introduced later in the text). The water depth is assumed to increase linearly from a value of \( h = 0.2 \) m and up to \( h = 4 \) m. It is assumed that over this shoaling region \( H_s \) stays constant (see, e.g., Ruessink et al., 2013). However, \( k_s \) does change and is determined by the linear dispersion relation according to the mild slope assumption. These wave conditions result in a steepness value of \( \epsilon = 0.08 \) at the shallowest point and a value of \( \epsilon = 0.08 \) at the deepest point. The results of BFI as a function of \( k_s h \) are presented in Fig. 5 for three different values of the peak-enhancement factor: \( \gamma = 1.33, 10 \).

The results shown in Fig. 5 provide further evidence for the effect of modulational instability on the predictions according to the fully dispersive quadratic formulations. The discrepancies are particularly significant over \( 0.5 < k_s h < 2 \), a region where the effect of modulational instability is expected to be weak or absent (as confirmed by the NLSE, and see also additional support provided by Akrish et al., 2016, Fig. 5). The results also demonstrate the effect of increasing \( \gamma \). Higher \( \gamma \) values correspond to narrower spectra, and thus also to weaker dispersion effects (or larger correlation scale). Consequently, as \( \gamma \) increases, modulational instability becomes more dominant. Finally, the inclusion of the \( U_r \) number highlights the significance of the results shown by the fully dispersive quadratic formulations. The \( U_r \) number suggests that over \( 0.5 < k_s h < 2 \) the expected wave evolution can be characterized as quasi-linear. However, the BFI due to the fully dispersive formulations suggests that over this region of \( k_s h \), the evolution may be significantly affected by the nonlinear modulational instability mechanism. This point is further demonstrated and discussed through the following numerical example.

2.3.3. Bichromatic group evolution over constant depth

The example considered here concerns the evolution of a bichromatic wave group in a 60 m long flume with constant depth of \( h = 0.8 \) m. The group is assumed to be composed of a primary frequency, \( f_1 = 0.60 \) Hz, that is subjected to a side-band frequency of \( f_2 = 0.45 \) Hz, and a forced sub-harmonic frequency of \( f_3 = 0.15 \) Hz (\( f_1 = f_3 - f_2 \)). The corresponding incoming amplitudes are assumed to be \( a_1 = 0.08 \) m and \( a_2 = 0.008 \) m, where the amplitude of the forced sub-harmonic \( a_3 \) is obtained based on (9). These specifications result in the following primary wave parameters:

\[
\begin{aligned}
\delta \omega_p &= 0.25 \\
\gamma &= 1.33 \\
\epsilon &= 0.13 \\
\delta a_1 &= 0.08
\end{aligned}
\] (27)

Therefore, as suggested by Fig. 4 (middle panel), the modulation of the group is expected to grow along the flume following the predictions by Whitham (1967) and Bredmose et al. (2005). Modulational growth indeed occurs in both of these model formulations as implied by the sub-harmonic growth shown in Fig. 6.

The results shown in Fig. 6 are obtained numerically using a spatial step of \( \Delta x = 0.05 \) m and a spectral step of \( \Delta f = 0.05 \) Hz, where the highest frequency considered is \( f_{\text{max}} = 4 f_3 \). Based on these parameters, the spatial-dependent solution of the quadratic system, (1), is achieved through the classical fourth-order Runge–Kutta (RK4) method. Also note that the solutions shown here are phase-averaged over 10 different realizations (assuming random phases for \( a_1 \) and \( a_2 \) at the flume’s boundary for each realization). Finally, the results of the quadratic models are compared with the highly accurate SWASH model (Zijlema et al., 2011), which is implemented here using two vertical layers, spatial step of \( \Delta x = 0.02 \) m, time step of \( \Delta t = 0.005 \) s and simulation time of 10 min, where the results shown here are time-averaged over the last 5 min.

The significance of the results shown in Fig. 6 can be explained using the Ursell parameter, (3). The parameter value of the present example is estimated as \( U_r \sim 4.4 \), indicating that the amplitude spectrum, and accordingly also the significant wave heights, are expected to stay approximately constant along the flume. The results show that most of the quadratic formulations and SWASH indeed describe this permanent behaviour, demonstrating roughly constant \( H_s \) values of the shorter and longer waves. The weak deviations of SWASH from this...
Fig. 6. Bichromatic group evolution in a flume of constant depth (shown in terms of the significant wave heights, $H_s$), as predicted by the different model formulations and the SWASH model. The vertical dashed lines shown in the upper panel (located at $x = 4$ m, 24 m, 44 m) approximately indicate the spatial recurrence of the group’s focusing/defocusing as obtained by the prediction following Bredmose et al. (2005).

permanent behaviour (i.e., the short waves $H_s$ drop near the incoming boundary and the weak short and long waves $H_s$ oscillations) are due to mismatch between the imposed incoming wave field and the resulted downwave field (see Vasarmidis et al., 2024). In contrast, the predictions of Whitham (1967) and Bredmose et al. (2005) present significant deviations from the expected permanent behaviour outcome. These deviations are the result of modulational instability, giving rise to significant energy transfers between the carrier and the side-band components. Accordingly, the initial weak modulation, which forces relatively small sub-harmonic response, develops into a strong modulation and an abnormal sub-harmonic growth. As demonstrated by the lower panel of Fig. 6, the prediction of the sub-harmonic $H_s$ following Whitham (1967) and Bredmose et al. (2005) is an order of magnitude higher than expected. Eventually, the significant growth of the side-bands and the corresponding modulations is restricted by energy conservation. This leads to back and forth transfers of energy and to the well-known long distance behaviour of focusing/defocusing recurrence (e.g., Lake et al., 1977), which is demonstrated through the prediction of Bredmose et al. (2005) in Fig. 7. This model prediction shows that the initial weakly modulated monochromatic field evolves into an almost fully modulated bichromatic field due to intensive energy exchanges between the carrier and the side-band components. Close to the end of the flume, the initial energy spectrum is almost recovered.

To summarize, the analysis presented in this section shows that the positive effect of improving the dispersion on the linear behaviour of wave fields also involves an unfavourable effect on the nonlinear evolution. Nonlinear effects are governed by the quadratic interaction coefficients, $V_{lmn}$, and generally, also by higher-order factors which are neglected based on the quadratic modelling approach. As a consequence, a change of $V_{lmn}$ due to improvement of the linear dispersion relation also involves a change of the neglected higher-order residual. It turns out that this residual becomes significant for the fully dispersive models, leading to over prediction of the amplitude dispersion, and as a consequence, to changes in the modulational instability mechanism, including the instability threshold, the growth rate and the modulation

ranges over which this mechanism is expected. Hence, fully dispersive quadratic formulations not only tend to develop significant phase errors, but are also exposed to intensive energy exchanges triggered by false modulational instability, which may even take place over relatively shallow waters and under conditions for which quasi-linear behaviour is expected.

3. Model prediction of waves over coastal waters

The impact of modulational instability on coastal wave prediction and the prediction capabilities of the different quadratic formulations are studied here through comparisons with different laboratory and SWASH results. The considered cases concern incoming wave fields which are either bichromatic or described through a continuous spectrum. Since the examples presented here involve bathymetry changes, the quadratic formulation (1) discussed so far should be modified to include the effect of wave shoaling. This can be readily implemented by using the energy-flux related amplitude, \( c \), which is described by the thick green lines. The thin vertical lines, plotted along the still water level \((h = 0)\), indicating measurement locations. The vertical dashed lines indicate the locations where computed and measured wave spectra are compared.

### Table 1

Incoming wave parameters for the bichromatic examples of Van Noorloos (2003). The incoming forced amplitude \( a_{mpb} \) of the sub-harmonic indicated by \( f_j \) is calculated based on second-order Stokes theory. Additionally, \( U_{r,max} \) estimates the Ursell number at the breaking point and \( \xi_{max} \) estimates the maximum location for which \( U_r \leq 26 \).

<table>
<thead>
<tr>
<th>Exp</th>
<th>( f_1 ) (Hz)</th>
<th>( f_2 ) (Hz)</th>
<th>( f_1 ) (Hz)</th>
<th>( a_{mpb} ) (m)</th>
<th>( a_{mp} ) (m)</th>
<th>( \mu )</th>
<th>( U_{r} )</th>
<th>( U_{r,max} )</th>
<th>( \xi_{max} ) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.6714</td>
<td>0.4761</td>
<td>0.1953</td>
<td>0.06</td>
<td>0.012</td>
<td>1.43</td>
<td>4.7</td>
<td>41.8</td>
<td>21</td>
</tr>
<tr>
<td>B3</td>
<td>0.6470</td>
<td>0.5005</td>
<td>0.1465</td>
<td>0.06</td>
<td>0.024</td>
<td>1.35</td>
<td>5.6</td>
<td>39.2</td>
<td>20</td>
</tr>
<tr>
<td>B5</td>
<td>0.6470</td>
<td>0.5005</td>
<td>0.1465</td>
<td>0.06</td>
<td>0.036</td>
<td>1.35</td>
<td>6.0</td>
<td>38.6</td>
<td>19</td>
</tr>
</tbody>
</table>

3.1. Evolution of bichromatic groups and irregular waves over a slope

Two sets of laboratory experiments, conducted by Van Noorloos (2003) and Ruessink et al. (2013), are considered. Generally speaking, these experiments describe one-dimensional nonlinear shoaling of wave fields over a mild slope. The settings of these experiments are described schematically in Fig. 8 and the parameters of the incoming wave fields are detailed in Tables 1 and 2. Model capabilities are examined by comparisons to measured results and to the predictions of SWASH up to the breaking points beyond which the quadratic formulations become invalid. The comparison focuses on the evolution of the primary components and the generation and development of the secondary components (the super and sub-harmonics). Finally, since the quadratic formulations only account for the incoming wave components, the examined cases considered here are such that the effect of wave reflection on the evolution of the primary and secondary components is negligible (Rijnsdorp et al., 2014; De Bakker et al., 2015). Accordingly, the measured data is not separated into incoming and reflected wave components. To avoid unwanted reflections in SWASH, a radiation condition is applied on the downwave side of the domain at a depth of \( h \sim 0.057 \text{ m} \) accompanied by a sponge layer of 5 m in front of it. The combination of these measures allows an effective absorption of both the long and the short wave components, as verified in Fig. 9 and Fig. 12.

3.1.1. Bichromatic groups over a slope

Three bichromatic examples introduced by Van Noorloos (2003) (i.e., A1, B3 and B5) are considered. Wave predictions for these examples as obtained by the different quadratic formulations are compared to measured and SWASH results in Fig. 9. The comparison is presented in terms of the \( H_s \) of the primary and super-harmonics (defined by \( f > f_{s,max} \), and referred to as the sea-swell components) and the \( H_s \) of the sub-harmonics (defined by \( f \leq f_{s,max} \) and referred to as the infragravity (IG) components), where the separating frequency takes the following value: \( f_{s,max} = 0.3 \text{ Hz} \). Additionally, the results of the quadratic formulation are computed using spectral resolution of \( \Delta f = 0.025 \text{ Hz} \) and maximum frequency of \( f_{s,max} = 4 f_s \) (recall that \( f_s \) is the peak frequency). Furthermore, these results are phase-averaged over 10 realizations. Finally, the results according to SWASH are time-averaged over the last 6 min, where the total simulation time is chosen to be 10 min.

The depth parameter values, \( \mu_j \), given in Table 1 indicate that the considered examples describe incoming wave groups over intermediate water depth. Additionally, these groups are characterized by relatively small incoming \( U_r \) value (see \( U_{r,j} \) in Table 1). The Ursell number only becomes significant around the breaking area. Thus, the evolution is expected to be quasi-linear, namely, dominated by linear dispersion and shoaling along most of the domain (up to \( x \sim x_{s,max} \)) for all the considered examples.

Based on these expectations, the predictions presented by the fully dispersive models in Fig. 9 are surprising. These predictions describe...
significant energy transfers from the primary component (i.e., the component with frequency $f_1$) to secondary components, as implied by the decrease of the sea-swell $H_s$ and the relatively rapid $H_s$ growth of the IG components. The mechanism which triggers these energy exchanges is attributed to modulational instability.

The impact of modulational instability on the group evolution may be explained through the data detailed in Table 1. The depth parameter of the primary component (given by $\mu_s$) provides an indication whether modulation instability can emerge. Following Fig. 3, it is clear that all the three fully dispersive formulations (i.e., the formulations by Whitham, 1967, Kaihatu and Kirby, 1995 and Bredmose et al., 2005) are subjected to modulational instability for all the considered examples. However, the significance of the energy transformation due to modulational instability is determined by the steepness parameter (i.e., $\delta_m$) and the normalized modulation frequency (i.e., $\delta_m$). As an example, the wave conditions, based on which the growth rate shown in Fig. 4 (left panel) is calculated, approximately describe the wave conditions of example B3 and B5 at $x \sim 15$. Consequently, it is expected that the predictions for these examples following the formulations by Whitham (1967) and Bredmose et al. (2005) will be strongly affected by the modulational instability mechanism. These models indeed describe rapid $H_s$ growth of the IG components, indicating significant modulational growth. On the other hand, the Boussinesq formulations agree better with the measured and SWASH results and with the expectation of quasi-linear evolution. The largest deviations obtained based on these models observed by the sea-swell $H_s$ prediction due to Freilich and Guza (1984) and by the $H_s$ prediction of the infragravity response due to Madsen and Sørensen (1993). The former develops due to over prediction of the primary components by Freilich and Guza, 1984). Especially, the results of Freilich and Guza (1984) and Nwogu (1993) show accurate development of the complete spectrum (excluding the over prediction of the primary components by Freilich and Guza, 1984). The prediction of Madsen and Sørensen (1993) though, tend to under predict the development of the super-harmonics.

3.1.2. Irregular waves over a slope

This part is devoted to the evolution of irregular wave fields. The considered examples are the irregular cases which were experimentally investigated by Ruessink et al. (2013). The generated wave fields are defined based on the JONSWAP spectrum using the parameters detailed in Table 2. The computations through the quadratic formulations are based on a spectral resolution of $\Delta f = 0.015$ Hz, maximum frequency of $f_{\text{max}} = 4 f_p$ and averaging over 60 realizations. The computations through SWASH are based on a simulation time of 60 min, where the results represented are time-averaged over the last 54 min.

The computed and measured results are compared in Fig. 12 in terms of $H_s$. Here again, the values of $H_s$ are presented separately for the shorter waves (denoted as the sea-swell components and satisfy $f > f_{\text{ig,max}}$) and for the longer waves (denoted as the infragravity (IG) components and satisfy $f \leq f_{\text{ig,max}}$), where the separation frequency, $f_{\text{ig,max}}$, is provided in Table 2 for each of the considered examples. For the intermediate to shallow water depth conditions that characterize these examples, the values of $x_{\text{ig, max}}$ define the regions over which second-order Stokes theory is expected to be valid. Over these regions, wave evolution is expected to be dominated by linear dispersion and shoaling, while evidence of nonlinear exchanges of energy is expected to be weak. This highlights again the abnormal infragravity responses as presented by the fully dispersive formulations in Fig. 12. Here, the explanation for these results is based on the BFI parameter. As an example, refer to the middle panel of Fig. 5 which shows the BFI values due to the fully dispersive formulations under the same wave conditions as of example A2. These BFI values indeed suggest that for the considered $\mu$ values, wave prediction using the fully dispersive formulations (especially using the model by Whitham, 1967) would be strongly affected by the modulational instability mechanism. This may lead to unexpectedly strong prediction of wave field’s modulations, and consequently, to the prediction of unexpectedly strong responses of the infragravity components as indeed suggested by Fig. 12. Furthermore,

![Fig. 9. A comparison of computed and measured $H_s$ for the bichromatic examples A1 (upper row), B3 (middle row) and B5 (lower row) of Van Noorloos (2003). The vertical dashed lines provide estimation for the initial wave breaking locations.](image-url)
Fig. 10. A comparison of amplitude spectra as obtained by the measurements, SWASH and the fully dispersive models.

Fig. 11. A comparison of amplitude spectra as obtained by the measurements, SWASH and the Boussinesq models.

Table 2

<table>
<thead>
<tr>
<th>Exp.</th>
<th>( f_p ) (Hz)</th>
<th>( H_s ) (m)</th>
<th>( \gamma )</th>
<th>( f_{1max} )</th>
<th>( r_l )</th>
<th>( U_{r,1} )</th>
<th>( U_{r,max} )</th>
<th>( x_{s,max} ) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.6329</td>
<td>0.1</td>
<td>3.3</td>
<td>0.37</td>
<td>1.5</td>
<td>2.1</td>
<td>47.5</td>
<td>61</td>
</tr>
<tr>
<td>A2</td>
<td>0.4444</td>
<td>0.2</td>
<td>3.3</td>
<td>0.26</td>
<td>0.9</td>
<td>11.4</td>
<td>52.5</td>
<td>37</td>
</tr>
<tr>
<td>A3</td>
<td>0.4444</td>
<td>0.1</td>
<td>20</td>
<td>0.26</td>
<td>0.9</td>
<td>5.7</td>
<td>77.3</td>
<td>49</td>
</tr>
</tbody>
</table>

The results of Fig. 12 provide additional evidence to the reliability of wave prediction using the Boussinesq models. These models agree well with the measured and SWASH results up until the breaking points. However, also here, the inaccurate shoaling prediction of Freilich and Guza (1984) and the inadequate nonlinear balance due to the quadratic coefficients of Madsen and Sørensen (1993) result in over prediction of the sea-swell \( H_s \) and the infragravity \( H_s \), respectively.

Further details explaining the prediction capabilities of the different quadratic formulations are presented in Figs. 13–14. These results provide a limited view on the spectral evolution as obtained for example A2 close to the breaking point. In order to highlight the modelling capabilities of the infragravity components (which also provide possible indication to the significance of modulational instability), the spectra are plotted through logarithmic scales. The results provide another perspective on the effect of modulational instability, which induces much faster spectral broadening than predicted by the measurements (especially notable by the results of Whitham, 1967 and Bredmose et al., 2005, but also seen less obviously through the results of Kähler and Kirby, 1995 in Fig. 13). In addition, the tendency of the model by Madsen and Sørensen (1993) to over predict the sub-harmonic responses and to under predict the super-harmonic responses is revealed again through Fig. 14. Moreover, Fig. 14 provides additional demonstration of the tendency of the model by Freilich and Guza (1984) to over predict the shoaling primary components. In summary, the model by Nwogu (1993) seems to generate the most accurate prediction for this example.

To summarize, the verification conducted for both bichromatic and irregular wave conditions shows the preferable prediction capabilities of the Boussinesq models. However, all the examples considered showed the tendency of Freilich and Guza (1984) to over predict the sea-swell components due to inaccurate formulation of linear shoaling and the tendency of Madsen and Sørensen (1993) to under predict the
Fig. 12. A comparison of computed and measured $H_s$ for the irregular examples A1 (upper row), A2 (middle row) and A3 (lower row) of Ruessink et al. (2013). The vertical dashed lines provide estimation for the initial wave breaking locations.

Fig. 13. A comparison of variance spectra as obtained by the measurements, SWASH and the fully dispersive models.

Fig. 14. A comparison of variance spectra as obtained by the measurements, SWASH and the Boussinesq models.

sea-swell components and to over predict the IG components due to inadequate nonlinear balance provided by the quadratic coefficients. The model of Nwogu (1993) present the most satisfying general agreement with the measured and SWASH results, and together with the model
of Freilich and Guza (1984) showed the most accurate prediction of the infragravity response. Finally, the predictions of the fully dispersive formulations deviated considerably from the measurement results. The mechanism that leads to the observed deviations appears to be the modulational instability. This instability mechanism triggered rapid energy exchanges and accompanying growth of wave modulations over regions for which the waves are expected to develop (almost) linearly. As a result, an unexpectedly strong response of the infragravity field is presented by the predictions of these fully dispersive formulations.

4. Discussion and concluding remarks

The quadratic modelling approach proposes a significant modelling reduction of the original Euler equations. The description of wave nonlinearity essentially collapses into a single mode coupling term determined by the quadratic interaction coefficients. Therefore, doubt arises regarding the prediction capabilities of this modelling approach especially for coastal waves which are characterized by significant nonlinearity.

Since the first development of the quadratic formulation by Freilich and Guza (1984), efforts were mainly devoted to the improvement of the dispersion relation and the inclusion of highly dispersive and weakly nonlinear terms. Beyond the improvement of the linear wave properties, accounting for high-order dispersive terms also improves the representation of the second-order bound waves (see Madsen and Schäffer, 1998). As an example, the formulation by Bredmose et al. (2005) provides second-order bound wave solutions which exactly match those of Stokes theory. However, as demonstrated throughout this study, these added capabilities also involve unfavourable consequences on the nonlinear evolution. This unexpected impact can be explained by considering the neglected residual (or truncation error) arising due to the modelling reduction of the quadratic modelling approach. It turns out that this residual may lead to significant consequences for the prediction of coastal waves using quadratic formulations with improved dispersion. Specifically, the residuals corresponding to the fully dispersive models of Whitham (1967), Kaihatu and Kirby (1995) and Bredmose et al. (2005) lead to considerable overpredictions of the so-called amplitude dispersion, and consequently, also to unfavourable modifications of the modulational instability mechanism. Therefore, beyond phase errors, errors related to nonlinear energy exchanges are expected to evolve as well using these models. Stability analysis showed that these models become modulationally unstable over much shallower water than expected and are subjected to much stronger growth rates and much larger modulation ranges. As a result, predictions using these models may be significantly affected by modulational instability even under conditions (e.g., water depth, spectral bandwidth) for which this mechanism is expected to be weak or absent. Specifically, this study shows the consequences of false development of modulational instability on the prediction of waves in coastal waters. As an example, the evolution of relatively linear waves (characterized by small Ursell number) over relatively shallow waters ($\mu < 1.36$) was examined several times along this study. The expected evolution for such conditions should be well described by linear theory. However, the fully dispersive formulations demonstrated entirely different results, which are characterized by rapid growth of wave modulations and associated growth of the infragravity field.

In contrast to the fully dispersive formulations, the Boussinesq models (i.e., the models by Freilich and Guza, 1984, Madsen and Sørensen, 1993 and Nwogu (1993)) tend to underestimate the amplitude dispersion, and consequently, modulational instability has little or no effect on these models’ prediction. Therefore, based on the laboratory results presented here (designated to characterize coastal wave evolution), for which modulational instability is negligible, the Boussinesq models demonstrated more reliable predictive capabilities. The difference between the predictive capabilities of the two model types (i.e., the Boussinesq approach and the fully dispersive approach) stood out in particular in the shoaling regions, over which wave conditions were typically quasi-linear. In contrast to the fully dispersive formulations, the Boussinesq models which are not exposed to false development of modulational instability described more adequately the quasi-linear evolution of the shoaling waves. Nevertheless, these models also demonstrated some deviations in comparison to the measurements. Specifically, the model by Freilich and Guza (1984), while accurately predicted the infragravity field, showed shoaling deviations of the shorter waves as a result of its simplified description of linear dispersion. Furthermore, the model by Madsen and Sørensen (1993) demonstrated deviations of both the sub and super-harmonics. These latter deviations are explained by the inadequate nonlinear balance characterizing the quadratic coefficients of Madsen and Sørensen (1993).

Ultimately, based on the examples presented here, it appears that the model by Nwogu (1993) provides the most adequate coastal wave predictions considering the evolution of both the sea-swell field and the infragravity field. It should be emphasized that the investigation in this study is limited to coastal waters. Under deep water conditions, nonlinear wave evolution is controlled by effects related to the interaction between wave quartets (e.g., amplitude dispersion, modulational instability). Therefore, all the quadratic formulations presented here would not be suitable to describe wave nonlinearity in deep water, since none of them correctly captures these nonlinear effects for relatively high values of $\mu$ (see Figs. 2 and 3).

To conclude, this study raises the need to improve the capabilities of the quadratic approach for the prediction of coastal waves. The Boussinesq approach is relatively reliable, but shows discrepancies related to inaccurate description of linear dispersion or inadequate nonlinear balance between super and sub interactions. On the other hand, the improved linear properties of the fully dispersive approach are accompanied by the tendency to develop false modulational instability. Therefore, a clear modelling gap remains. This gap indicates the need for deriving of a fully dispersive quadratic model, which adequately describes nonlinear wave developments over water depths and bathymetrical structures that characterize the coastal environment.

CRediT authorship contribution statement

Gaia Akrish: Conceptualization, Methodology, Writing – original draft. Ad Reniers: Supervision, Writing – review & editing. Marcel Zijlema: Supervision, Writing – review & editing. Pieter Smit: Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work is part of the research programme Earth and Life Sciences (ALW) with project number ALWOP.167, which is (partly) financed by the Dutch Research Council (NWO).

Appendix A. Formulation of the quadratic model

Generally speaking, the derivation of the quadratic model starts with the underlying time-domain model. The latter is usually written
as a set of two equations for the surface elevation, \( \eta \), and for the fluid velocity variable (may be the depth-averaged horizontal velocity or the horizontal velocity at a certain elevation level or the surface potential velocity etc.). However, in order to simplify the discussion, it is assumed that the underlying time-domain model can be written through the following combined form:

\[
D(i\partial_x - i\partial_y)\eta = N(i(i\partial_y, i\partial_z, -i\partial_x, -i\partial_z)\eta) + \eta^1(\eta^2)
\]

(29)

where \( D \) is a linear differential operator and \( N \) defines a nonlinear operator. The superscripts used on the right-hand-side of (29) (indicated by the numbers 1, 2) specify the partial derivatives with superscript \( (j) \) operates on \( \eta^j \). Additionally, as implied by (29), the effect of slow bottom variations is ignored here (otherwise the differential operators were dependent on \( x \)). This is done based on the common assumption that this effect, which is manifested by the so-called linear shoaling term, is of the same order as of the quadratic nonlinear term. As such, the shoaling term can be simply added to the quadratic model separately at a later stage.

The usual procedure to derive the quadratic model, (1), is through the multiple-scale method. A detailed account for this derivation can be found in Dingemans (1997), Chapter 7. Here, this derivation is briefly summarized based on the combined model form, (29). To start with, two spatial scales are defined, \( x_1 = x \) and \( x_2 = cx \), where recall that \( c \) represents a small value of the field’s nonlinearity. In addition, \( \eta \) is expanded as \( \eta = \eta_1 + \epsilon^2 \eta_2 + \ldots \). By substituting these assumptions into (29), one obtains a set of equations, each balancing terms of mutual order. Here, only the first two equations are required. These are given by,

\[
D_1\eta_1 = 0
\]

(30)

\[
D_1\eta_2 = iD'_r\partial_{x_1}\eta_1 + N_1\eta_1^2 + \eta_2^2
\]

(31)

where the subscript in \( D_1 \) and \( N_1 \) indicates that these are operators functions of the spatial derivative \( \partial_{x_1} \). In addition, the definition of the operator \( D'_r \) stems from the following symbolic Taylor expansion:

\[
D(i\partial_x - i\partial_y) - i\epsilon\partial_{x_1} = D_1 - i\epsilon D'_r\partial_{x_1} - \frac{\epsilon^2}{2} D''_r\partial_{x_1}^2 + \ldots
\]

(32)

where the tag notation defines derivative with respect to the factor \( -i\partial_{x_1} \).

The non-trivial solution, \( \eta_1 \), is assumed to be periodic in time and slowly modulated in space, and therefore, it is assumed to take the following form:

\[
\eta_1 = \sum_n A_n(x_2)\exp(i k_n x_1 - i\omega_n t)
\]

(33)

By substituting this solution into (30), the following linear dispersion relation is obtained

\[
D(i\omega_n, k_n) = 0
\]

(34)

and the substitution of \( \eta_1 \) into (31) provides

\[
D_1\eta_2 = \sum_n \exp(-i\omega_n t) \left[ iD'_n\partial_{x_1} A_n \exp(i k_n x_1) + \sum_{m} N_{n,m} A_{n,m} \exp((ik_n + k_{m,-})x_1) \right]
\]

(35)

where now the tag \( D'_n \) indicates derivative with respect to \( k_n \) and \( D_n = D(i\omega_n, k_n) \). Additionally, \( N \) is associated with the operator \( N_1 \), such that \( N_{n,m} = N(i\omega_n, i\omega_m, k_n, k_{m,-}) \).

It is well-known that the dispersion relation does not allow resonance to occur between three waves, and therefore, at this stage, only bound wave solutions are expected to exist. However, as the water depth becomes shallower, the wavenumber mismatch, \( k_n + k_{m,-} - k_{l,0} \) becomes smaller, creating a weaker condition that is commonly referred to as “near resonance”, which practically allows energy exchange to take place. Therefore, in order to keep the solution bounded over relatively shallow waters, it is required that the right-hand-side of (35) is set to zero. This condition is expressed as follows:

\[
iD'_n\partial_{x_1} A_n \exp(i k_n x_1) + \sum_m N_{n,m} A_{n,m} \exp((ik_n + k_{m,-})x_1) = 0
\]

(36)

or, in terms of \( a_n = A_n \exp(i k_n x_1) \)

\[
\partial_{x_1} a_n - i k_n a_n = -i \sum_m V_{n,m} a_{n,m}
\]

(37)

As can be observed, this is exactly the quadratic model which is introduced by (1), where since the scale separation has stopped to play a role at this stage, it is simply reduced. Additionally, the transition from (36) to (37) suggests the following relation:

\[
V_{n,m} = -N_{n,m}/D'_n
\]

(38)

where \( n = l + m \). This relation can be conveniently used to derive the missing quadratic coefficients of Nwogu (1993) and Whitham (1967) (see derivation in Supplementary material).

A.1. Shoaling term

As mentioned earlier, the shoaling term can be added into the quadratic formulation separately by assuming it to be of the same order as the quadratic nonlinear term. The derivation of the shoaling term is based on the usual WKB assumption allowing the wavenumbers to be weakly dependent on \( x \). Namely, the first order solution, \( \eta_1 \), takes now the following form:

\[
\eta_1 = \sum_n A_n \exp(iS_n(x) - i\omega_n t)
\]

(39)

where \( \partial_{x_1} S_n = k_n \). The linear operator, \( D_1 \), is now weakly dependent on \( x \) as well. For model formulations which are based on the so-called free-surface velocity potential (see definition in, e.g., Zakharov, 1968), the slowly varying operator, \( D_1 \), can be treated as a Weyl operator (Akrish et al., 2023). This is convenient since the Weyl operator provides the following general formula for the shoaling term (see, e.g., Akrish et al., 2020, Eq. (AB8)):

\[
\frac{\partial_{x_1} A_n}{A_n} = \frac{\partial_{x_1} k_n D''_n}{2D'_n} + \frac{\partial_{x_1} D'_n}{2D'_n}
\]

(40)

where \( D'_n \) indicates second derivative with respect to \( k_n \). This equation leads to the following well-known linear shoaling definition (which corresponds to the well-known energy flux conservation):

\[
\partial_{x_1} \left[ A_n \frac{\sqrt{g k_n}}{\sqrt{x}} \right] = 0
\]

(41)

where \( \sqrt{g k_n} \) is the fully dispersive group velocity if one of the fully dispersive models is considered, while for the model by Freilich and Guza (1984), the shallow water approximation, \( \sqrt{g k_n} \) is used. Note that the shoaling term formula, (40) (or (41)), is applied here also for the model by Whitham (1967) which was developed for constant depth. This applicability is argued heuristically based on the derivations of the Whitham systems in Moldabayev et al. (2015) and Akrish (2023). A more rigorous derivation is beyond the scope of this study. Finally, recall that the Weyl formula, (40), is valid for models which are formulated based on the free-surface velocity potential. Accordingly, this formula does not lead to the correct shoaling terms of Madsen and Sørensen (1993) and Nwogu (1993). Furthermore, the expression in (41), written in terms of \( \sqrt{g k_n} \), is also not found to agree with the shoaling terms of Madsen and Sørensen (1993) and Nwogu (1993) (see discussions by Beji and Nadaoka, 1996 and Schäffer and Madsen, 1998). However, Fig. 3 in Schäffer and Madsen (1995) shows that the deviation of the shoaling terms of Madsen and Sørensen (1993) and Nwogu (1993) from the fully dispersive shoaling is practically negligible over depths that characterize coastal waters. Therefore, for the computations performed here, the fully dispersive shoaling term is applied for the models by Madsen and Sørensen (1993) and Nwogu (1993) as well.
Appendix B. Cubic interaction coefficients

The cubic interaction coefficients, $C_{i,j,k,l}$, are formulated through the solvability condition, (11), that is obtained at third order. This equation describes the interaction of resonant quartets, which satisfy the conditions

$$\begin{align*}
\omega_i + \omega_j &= \omega_k + \omega_l \\
k_i + k_j &= k_l + k_l + O(\epsilon^2)
\end{align*}$$

(42)

for non-trivial interactions, while exact resonance is obtained for the trivial interactions (the interactions which result in amplitude dispersion contributions). As implied by the formulation of the solvability condition, the cubic coefficients of the cubic terms (appearing implicitly on the right-hand-side of (11)) are constructed as sums of multiplications of the quadratic coefficients, $V$, and the quadratic transfer function, $G$. In the following, this is demonstrated explicitly based on the assumption that $\omega_k \leq \omega_i \leq \omega_j \leq \omega_l$. Consider first the more common resonant case which satisfies the following:

$$\begin{align*}
\omega_i - \omega_l &= \omega_i - \omega_j \\
\omega_j - \omega_l &\neq \omega_i - \omega_l \\
\omega_i &= \omega_i + \omega_j - \omega_j \\
\omega_i &= \omega_i + \omega_j - \omega_k \\
\omega_j &= \omega_j + \omega_k - \omega_l \\
\omega_l &= \omega_l + \omega_j - \omega_j
\end{align*}$$

(43)

In this case, any wave component participates in four trivial resonant interactions and one non-trivial interaction, as is demonstrated for the component corresponding to $\omega_k$ as follows:

$$\begin{align*}
\omega_i &= \omega_i + \omega_j - \omega_j \\
\omega_j &= \omega_i + \omega_i - \omega_j \\
\omega_i &= \omega_i + \omega_j - \omega_k \\
\omega_i &= \omega_i + \omega_i - \omega_j \\
\omega_j &= \omega_i + \omega_j - \omega_j
\end{align*}$$

(44)

The cubic coefficient corresponding to each of these interactions consists of different products of $G$ and $V$ which are determined by the right-hand-side of the equations in (44). For example, consider the cubic coefficient of the non-trivial interaction which is determined according to the last equation of (44). Each pair of frequencies on the right-hand-side of this equation forces a second-order bound solution which is resonant with the third frequency and with $\omega_i$, as a result, the following expression is obtained:

$$C_{i,j,k,l} = 2G_{i,k}V_{l,k-j} + 2G_{i,j}V_{l,j-k} + 2G_{i,j}V_{k,j}$$

(45)

The formulation for the cubic coefficients of the trivial interactions can be derived in a similar fashion based on the other equations of (44). The resulted evolution equation for $a_l$ is given by

$$\partial_x a_l = \left( C_{i,k,j}a_i^2 + 2C_{i,k,j}a_i^2 + 2C_{i,l,j}a_i^2 + 2C_{i,l,k}a_i^2 \right)a_i + 2C_{i,k,j}a_ia_ia_{i,j}$$

(46)

which is coupled with the corresponding equations for $a_k$, $a_l$, and $a_j$. Note that the compact notations defined for the cubic coefficients of the trivial interactions $C_{i,k,j}$ and $C_{i,l,j}$, conveniently represent the notations $C_{i,k,j}$, and $C_{i,l,j}$, respectively.

The above described formulation, when applied for cubic coefficients that correspond to the trivial resonant interactions, raises some doubt due to apparent singularity demonstrated by terms like $G_{i,l,j}$. The quadratic model naturally bypasses this difficulty since it excludes the set-down terms due to self interactions as a results of the property which states that $V_{\omega_i}=0$. The formulation of the cubic coefficients based on the quadratic model is referred here as the discontinuous definition. On the other hand, the continuous definition requires some assumptions regarding the limit of $G_{i,j,k}$, which should be identical through the different convergence paths described by $G(\omega_i \pm \Delta \omega, -\omega_j \pm \Delta \omega)$ and by taking the limit $\Delta \omega \rightarrow 0$. Thus, consider for instance the cubic coefficient, $C_{i,j}$, defined by the right-hand-side of the first equation in (44). The discontinuous definition expresses this coefficient as

$$C_{i,j} = 2G_{i,j}V_{2i-j}$$

(47)

whereas the formulation following the continuous definition gives

$$C_{i,j} = 2G_{i,j}V_{2i-j} + 4G_{i,j}V_{i,j}$$

(48)

The continuous definition is used in this study to relate between the different cubic coefficients for the case of three-wave interaction. An example of such case is defined by the condition (43) and assuming that $\omega_i = \omega_j$. The three cubic coefficients that correspond to the non-trivial interaction are given as follows:

$$\begin{align*}
C_{k,l,i,j} &= 2G_{i,k}V_{2i-k} + 4G_{i,k}V_{i,k} \\
C_{k,l,i,j} &= 2G_{i,k}V_{i,k} + 2G_{i,k}V_{i-k} + 2G_{i,k}V_{k-i}
\end{align*}$$

(49)

Under the assumption of small modulation frequency, namely, $\omega_i - \omega_k = \omega_j - \omega_l = \Delta \omega$, where $\Delta \omega \ll \omega_i$, one obtains the approximation

$$C_{k,l,i,j} \sim C_{k,l,i,j} \sim C_{i,j}$$

(50)

if the following conditions hold:

$$\begin{align*}
G(\omega_i \pm \Delta \omega, -\omega_j \pm \Delta \omega) &= G_{i,j} + O(\Delta \omega) \\
G(\omega_i \pm \Delta \omega, \omega_j \pm \Delta \omega) &= G_{i,j} + O(\Delta \omega) \\
V(\omega_i \pm \Delta \omega, -\omega_l \pm \Delta \omega) &= V_{i,j} + O(\Delta \omega) \\
V(\omega_i \pm \Delta \omega, \omega_l \pm \Delta \omega) &= V_{i,j} + O(\Delta \omega)
\end{align*}$$

(51)

Appendix C. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.coastaleng.2024.104502.

References


