Testing and Calibrating Parametric Wave Transformation Models On Natural Beaches

Alex Apotsos[1], Britt Raubenheimer[2], Steve Elgar[3]
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

R. T. Guza[4]
Scripps Institution of Oceanography, La Jolla, California, USA

[1] Corresponding author, Department of Applied Ocean Physics and Engineering, MS #9, phone 508 289 3756, email: aapotsos@whoi.edu
[2] email: britt@whoi.edu
[3] email: elgar@whoi.edu
[4] email: rguzaucsd.edu

Abstract

To provide coastal engineers and scientists with a detailed inter-comparison of widely used parametric wave transformation models, several models are tested and calibrated with extensive observations from 6 field experiments on barred and unbarred beaches. Using previously calibrated ("default") values of a free parameter γ, all models predict the observations reasonably well (median root-mean-square wave height errors are between 10% and 20%) at all field sites. Model errors can be reduced by roughly 50% by tuning γ for each data record. No tuned or default model provides the best predictions for all data records or at all experiments. Tuned γ differ for the different models and experiments, but in all cases γ increases as the hyperbolic tangent of the deep-water wave height, H_o. Data from 2 experiments are used to estimate empirical, universal curves for γ based on H_o. Using the new parameterization, all models have similar accuracy, and usually show increased skill relative to using default γ.

Keywords: Wave models, surfzone, wave breaking, nearshore

1. Introduction

Numerical modeling increasingly is used to optimize coastal management and protection strategies. Nearshore wave transformation models used to predict currents, setup, and sediment transport range in complexity from wave-resolving, high-order solutions of the extended Boussinesq equations [*Nwogu,* 1993; *Kennedy et al.,* 2000] to wave energy balances using parameterizations of breaking-wave dissipation [*Battjes and Janssen,* 1978; *Thornton and Guza,* 1983]. Here, the accuracy of the parametric models, widely used because they are easy to code and are computationally efficient, is examined.

In all the models examined here, the breaking wave heights are assumed to follow simple probability distributions, and wave-breaking energy dissipation is parameterized using a theory for idealized bores. All the models contain a free parameter γ that can be tuned using wave height observations to provide more accurate predictions of the wave field at spatially dense locations, or to improve wave height forecasts for different time periods or locations.

After the models are outlined (section 2), the observations are described (section 3), and the method of model analysis is explained (section 4). Next, the models are evaluated using the observations, and a new parameterization for γ is developed (section 5). The results are then discussed (section 6), and the conclusions are summarized (section 7).

2. Wave Transformation Models

In all models, the wave field is assumed to be narrow banded in both frequency and direction, and the representative period is assumed to be constant in the cross-shore. The dissipation of wave energy caused by bottom friction is small in the surfzone [*Thornton and Guza,* 1983], and here all dissipation is assumed to result from either wave breaking ($\langle \varepsilon_b \rangle$) or shear stress at the wave-roller interface ($\langle \varepsilon_r \rangle$).

In all except one of the models considered, the cross-shore (x) gradient of the cross-shore wave energy flux, $E_w c_g \cos\theta$, is assumed equal to the local mean rate of energy dissipation in a breaking wave

$$\frac{\partial}{\partial x}\left(E_w c_g \cos\theta\right) = -\langle \varepsilon_b \rangle, \tag{1}$$

where θ is the mean wave angle relative to shore normal. The wave energy, E_w, is found from linear theory as $E_w = \frac{1}{8}\rho g H_{rms}^2$, where ρ is the water density, g is the gravitational acceleration, and H_{rms} is the root-mean-square wave height (defined as $2\sqrt{2}$ times the standard deviation of the sea-surface elevation fluctuations). The linear theory group speed is $c_g = c\left[\frac{1}{2} + \frac{kh}{\sinh(2kh)}\right]$, where c, k, and h are the local wave phase speed, wave number, and water depth, respectively.

Alternatively, in one model [*Lippmann et al.,* 1996] a wave roller is included in which

$$\frac{\partial}{\partial x}(E_r c \cos\theta) + \frac{\partial}{\partial x}(E_w c_g \cos\theta) = -\langle \varepsilon_r \rangle, \tag{2}$$

where the roller energy E_r is

$$E_r = \frac{1}{8}\rho c f \frac{H_{br}^3}{h\tan\sigma}, \tag{3}$$

where f is a representative wave frequency, σ is the slope of the wave front, and H_{br} is the height of the wave at breaking, as described below [i.e., (14)]. Here, σ is held constant at 12.5° (there is little variation in model accuracy for $\sigma > 10°$ [*Lippmann et al.,* 1996]). Energy balances similar to (2) have been used widely [e.g., *Stive and De Vriend,* 1994].

In all models except the roller model, the energy dissipation in a single breaking wave is equated to the dissipation in a hydraulic jump [*Stoker,* 1957; *LeMehaute,* 1962], and a probability distribution function (p.d.f.) is used to describe the fraction of waves that are breaking. The fraction of breaking or broken waves, Q, can be estimated using a Rayleigh wave height distribution truncated discontinuously at some maximum wave height H_m [*Battjes and Janssen,* 1978; *Battjes and Stive,* 1985; and *Nairn,* 1990] (labeled BJ, BS, and Nairn, respectively), yielding

$$\frac{1-Q}{-\ln Q} = \left(\frac{H_{rms}}{H_m}\right)^2, \tag{4}$$

where H_m is found by extending the Miche criterion for the maximum height of periodic waves of constant form [*Miche,* 1951] to

$$H_m = \frac{0.88}{k}\tanh\left(\frac{\gamma}{0.88}kh\right). \tag{5}$$

The free parameter γ is roughly equivalent to the maximum ratio of wave height to water depth, and controls the fraction of breaking waves. The dissipation is then given by

$$\langle \varepsilon_b \rangle = \frac{1}{4}\rho g f B Q H_m^2, \tag{6}$$

where B is of order 1 and controls the level of energy dissipation.

Alternatively, full (i.e., untruncated) Rayleigh distributions and empirical weighting functions can be used to describe the distribution of broken waves [*Thornton and Guza,* 1983; and *Whitford,* 1988] (labeled TG and Whit, respectively), with the corresponding energy dissipation given by

$$\langle \varepsilon_b \rangle = \frac{3\sqrt{\pi}}{16}\rho g f B^3 \frac{H_{rms}^3}{h} M[1 - \frac{1}{(1+(H_{rms}/\gamma h)^2)^{5/2}}], \tag{7}$$

with (TG)

$$M = \left(\frac{H_{rms}}{\gamma h}\right)^2, \tag{8}$$

or (Whit)

$$M = \left\{1 + \tanh\left[8\left(\frac{H_{rms}}{\gamma h}\right) - 1\right]\right\}. \tag{9}$$

For steep beaches where not all waves reach the maximum height and break, BJ can be extended using a full Rayleigh p.d.f. without the depth limitation of nearshore waves [*Baldock et al.,* 1998; and *Ruessink et al.,* 2003] (labeled Bald and Rues, respectively), yielding

$$\langle \varepsilon_b \rangle = \frac{1}{4}\rho g f B\left[-\left(\frac{H_b}{H_{rms}}\right)^2\right]\left(H_b^2 + H_{rms}^2\right), \tag{10}$$

where H_b is the local wave breaking height and can be approximated by H_m [e.g., (5)] (Rues), which in the limit of shallow water (i.e., $kh << 1$) (used by Bald) is

$$H_b = \gamma h. \tag{11}$$

Recently, Bald has been modified to correct for a singularity that can develop in shallow water [*Janssen and Battjes,* submitted] (labeled Jan), giving

$$\langle \varepsilon_b \rangle = \frac{1}{4h} B\rho g f H_{rms}^3 \left[\left(R^3 + \frac{3}{2}R \right) \exp\left(-R^2\right) + \frac{3}{4}\sqrt{\pi}\left(1 - erf\left(R\right)\right) \right], \tag{12}$$

where $R = H_b/H_{rms}$, with H_b determined from (11).

In the roller model (labeled Lipp), the work done by the roller on the surface of the wave, $\langle \varepsilon_r \rangle$, is

$$\langle \varepsilon_r \rangle = \frac{1}{4}\rho g f \frac{H_{br}^3}{h}\cos\sigma, \tag{13}$$

where

$$H_{br}^3 = \frac{3\sqrt{\pi}}{4} M H_{rms}^3 [1 - \frac{1}{(1 + (H_{rms}/\gamma h)^2)^{5/2}}], \tag{14}$$

with M given by (9). Note (13) and (14) are nearly equivalent to (7) if $B = 1$ and $\cos(\sigma) \approx 1$.

The peak, mean, or centroidal frequency usually is used as the representative wave frequency in the models. Although all models contain two free parameters, B and γ are interdependent [*Roelvink,* 1993], and can be combined into a single parameter [*Cacina,* 1989]. Here, B is held constant at 1, and γ is varied. Therefore, the tuned values of γ discussed here represent model parameters that implicitly take into account variations in B, and are not necessarily comparable with field observations of H_{rms}/h from previous studies [*Sallenger and Holman,* 1985; *Raubenheimer et al.,* 1996]. Similarly, because the model formulations differ, γ is not expected to have the same numerical value in each model.

The wave models (i.e., Table 1) can be broken into two groups. One group (TG, Whit, and Lipp) incorporates a weighting function in the determination of the dissipation, and one group (BJ, BS, Nairn, Bald, Rues, and Jan) does not. The TG, Whit, and Lipp models differ in the weighting function used [TG uses (8), whereas Whit and Lipp use (9)] and in the addition of a roller in Lipp. The BJ, BS, Nairn, Bald, Rues, and Jan models differ in the p.d.f. used (BJ, BS, and Nairn use a truncated Rayleigh p.d.f., whereas Bald, Rues, and Jan use a full Rayleigh p.d.f.).

The models use different formulations for γ based on calibrations with prior laboratory and field observations (referred to here as the "default" values of γ). Whit, Lipp, and TG use constants of $\gamma = 0.34$, 0.32, and 0.42, respectively. BS, Nairn, Bald, and Jan use functions of the deep-water wave steepness, $S_o = H_o/L_o$, where H_o and L_o are the deep-water wave height and length, respectively, with γ given by

$$\gamma = 0.5 + 0.4\tanh(33S_o) \text{ (BS)}, \tag{15}$$

and

$$\gamma = 0.39 + 0.56\tanh(33S_o) \text{ (Nairn, Bald, and Jan)}, \tag{16}$$

while Rues uses

$$\gamma = 0.76kh + 0.29. \tag{17}$$

Here, a more extensive data set comprised of multiple field experiments is used to evaluate and improve these formulations. See the references given above for the details of each wave model.

3. Observations
3.1 SandyDuck and Duck94: Duck, NC 1997 and 1994

Wave-induced pressures were measured at 2 Hz for 179.2 min starting every 3 hours using pressure gages at 21 (SandyDuck) [*Elgar et al.*, 2001] and 13 (Duck94) [*Raubenheimer et al.*, 1996] cross-shore locations between about 5-m water depth and the shoreline for 90 days during Sep to Nov 1997 (SandyDuck) and for 80 days during Aug to Oct 1994 (Duck94) on a barred beach near Duck, NC (Figures 1A & B, respectively). The 3-hr-long data records were subdivided into 8.5- (SandyDuck) and 17.5- (Duck94) min-long sections to reduce tidally induced depth changes. The bathymetry was surveyed approximately every other day from about 8-m water depth to above the high tide shoreline along cross-shore transects located about 20 m alongshore of the instrumented transects.

Root-mean-square wave heights at the most offshore sensor ($h \approx 5$ m) ranged from 0.20 to 2.10 m (SandyDuck) and from 0.14 to 2.92 m (Duck94). Centroidal frequencies ranged from 0.08 to 0.21 Hz (SandyDuck) and from 0.09 to 0.24 Hz (Duck94). Incident wave angles ranged between ±45° relative to shore normal during both experiments.

3.2 Egmond and Terschelling: The Netherlands, 1994 and 1998

Waves were measured at 2 Hz for approximately 34 min starting every hour for 40 days during Oct and Nov 1998 near Egmond, The Netherlands [*Ruessink et al.*, 2001], and for 34 days during April and May 1994 near Terschelling, The Netherlands [*Ruessink et al.*, 2003; *Ruessink et al.*, 1998] (Figure 1C & D, respectively). On the double-barred beach near Egmond, offshore waves were measured with a directional buoy in 15-m water depth, and wave-induced pressures were measured at 6 cross-shore locations between about 8- and 1-m water depths. On the triple-barred beach near Terschelling, wave-induced pressures were measured at 6 cross-shore locations between 9- and 2-m water depths. The data were processed in 34-min-long records. The bathymetry was surveyed approximately every other day at Egmond. At Terschelling the bathymetry was surveyed only once, but morphological changes during the experiment were negligible [Ruessink, personal communication].

Root-mean-square wave heights at the most offshore sensors ranged from 0.19 to 3.93 m (Egmond) and from 0.12 to 1.84 m (Terschelling). Peak frequencies ranged from 0.08 to 0.26 Hz (Egmond) and from 0.08 to 0.33 Hz (Terschelling). Incident wave angles ranged between ±45° at both experiments.

3.3 NCEX and SwashX: La Jolla, CA 2003 and 2000

Wave-induced pressures were measured at 16 Hz for 51.2 min starting every hour using buried pressure gages at 8 cross-shore locations between about 3.5-m water depth and the shoreline for 23 days during Oct and Nov 2003 (NCEX) [*Thomson et al.*, 2006] and for 14 days during Sep and Oct 2000 (SwashX) [*Raubenheimer*, 2002] on near planar beaches near La Jolla, CA (Figures 1E & F, respectively). The 1-hr-long data records were subdivided into 8.5-min-long sections. The bathymetry was surveyed between about 5-m water depth and the shoreline roughly 5 m alongshore from the instrumented transects approximately every other day.

Root-mean-square wave heights at the most offshore sensor ($h \approx 3.5$ m) ranged from 0.21 to 1.00 m (NCEX) and from 0.19 to 1.05 m (SwashX). Centroidal frequencies

ranged from 0.07 to 0.17 Hz (NCEX) and from 0.09 to 0.20 Hz (SwashX). Incident wave angles were within 5° of shore normal during both experiments.

4. Model Analysis
4.1 Model Procedure

All models were initialized with the rms wave height, representative frequency, mean wave angle, and still water depth observed at the most offshore sensor. The energy-weighted (centroidal) frequency was used as the representative frequency when wave spectra were available (Duck94, SandyDuck, SwashX, and NCEX). The peak frequency was used for Egmond and Terschelling. The wave period was assumed constant for all depths, and the wave angle was interpolated in the cross-shore using Snell's Law. The local water depth was estimated from the measured bathymetric profile, the tidal elevation relative to mean sea level at the offshore sensor, and the setup, $\overline{\eta}$, predicted as [*Longuet-Higgins and Stewart*, 1962 & 1964]

$$\frac{\partial}{\partial x}S_{xx} + \rho g(\overline{\eta} + d)\frac{\partial}{\partial x}\overline{\eta} = 0, \tag{18}$$

where d is the still water level and the wave radiation stress S_{xx} is

$$S_{xx} = E_w\left\{\left[\cos^2(\theta) + 1\right]\frac{c_g}{c} - \frac{1}{2}\right\}. \tag{19}$$

A 1st-order forward-step technique was used to determine the onshore variation of the wave heights. Wave heights were not estimated in depths less than 0.30 m where small errors in the measured bathymetry can lead to significant errors in the modeled wave heights. The nonlinear models, Lipp, BJ, Nairn, and BS, were solved iteratively at each step. Unless noted otherwise, all models were run for all data records in all experiments (Appendix A).

4.2 Model Tuning

For each data record in each experiment, γ in TG, Whit, Lipp, Bald, BJ, and Jan was fit to the observations. The best-fit γ was found by varying γ from 0.10 to 1.00 with a step size of 0.005, and by minimizing the *weighted root-mean-square* (rms) *percent error*

$$\textit{Weighted rms Percent Error} = \sqrt{\sum_n\left[\left((obs_n - pred_n)/obs_n\right)^2 * weight_n\right]} * 100\%, \tag{20}$$

where the weighting function, $weight_n$, is

$$weight_n = \frac{dist_{n-1} + dist_{n+1}}{dist_{tot}}, \tag{21}$$

and $dist_{n-1}$ and $dist_{n+1}$ are the distances from the nth sensor to the neighboring offshore and onshore sensors, and $dist_{tot}$ is the sum of all distances such that the sum of the weights is 1. All interior distances are counted twice, and for the most shoreward sensor $dist_{n+1}$ is assumed equal to $dist_{n-1}$. Since the model was initialized with the most offshore wave height, this value was not used in model tuning and the most offshore distance was counted only once. The 95% exceedence, the median, and the 5% exceedence errors for each model at each experiment were estimated as the smallest rms error that was larger than that calculated for 5%, 50%, and 95% of the records (e.g., Figure 2).

Percent errors were used [e.g., (20)] to give extra weight to the smaller wave heights near the shoreline, and the distance weighting [e.g., (21)] was used to give roughly

equal weight across the instrumented transects. However, the conclusions are not changed if unweighted or absolute error metrics are used (Appendix B).

5. Results
5.1 Model Evaluation
5.1.1 Default Models

Using default (e.g., previously calibrated) values for γ, the eight models (e.g., TG, Whit, Lipp, BS, Nairn, Bald, Rues, and Jan in Table 1) show reasonable agreement with the observations (e.g., Table 2 and Figure 3). However, the predicted wave heights from different models can differ significantly for a single data record (e.g., Figure 3), and the prediction errors for any one model can change significantly between different data records from one experiment (e.g., note the range between the 95% and 5% exceedence errors in Table 2). No single default model predicts the observations best for all data records (not shown), or for all experiments (for example, Bald has the lowest median error of any model at NCEX, but the third highest at Egmond and Terschelling). Although the models are based on the assumption that the waves are narrow banded in frequency, model errors are uncorrelated with spectral width, possibly because several other sources of error exist, and possibly because wave heights have a Rayleigh p.d.f. even for relatively broad spectra [*Elgar et al.*, 1984]. Model errors also are not correlated with the (usually small) difference between the centroidal and peak frequencies.

5.1.2 Tuned Models

Tuning the models improves model-data accuracy (e.g., errors are smaller in Table 3 than in Table 2). The tuned BJ, BS, and Nairn models are identical, and only BJ is given in subsequent figures and tables. The percent error reduction owing to model tuning is estimated from the Brier Skill Score (BSS) [*Murphy and Epstein*, 1989; *Ruessink et al.*, 2003]

$$BSS = \left[1 - \frac{Error(\gamma_{tuned})}{Error(\gamma_{untuned})} \right] * 100\%. \tag{22}$$

Tuning reduces the median errors by 25-50% (SandyDuck), 27-45% (Duck94), 67%-77% (Egmond), 32%-52% (Terschelling), 41-75% (NCEX), and 22-37% (SwashX) relative to the errors estimated from the default models. Similar reductions are found for the 95% and 5% exceedence errors (not shown). No single tuned model predicts the observations best for all data records or at all experiments.

5.2 Parameterization of γ
5.2.1 Best-fit γ

Each model shows a large spread in the best-fit γ for each experiment (e.g., Figure 4). In addition, the weighting functions [i.e., (8) and (9)] used by TG, Whit, and Lipp and the new dissipation formulation (12) used in Jan result in best-fit γ's that are smaller than those found for BJ and Bald. Therefore, the empirical relationships for γ developed for one model are not necessarily appropriate for other models. For example, the formulations developed for BJ or Bald [i.e., (15)-(17)] are not appropriate for use with TG, Whit, Lipp, or Jan. Although Jan used (15) in their model, the model predictions were not calibrated, and smaller values of γ result in improved model performance. The mean best-fit γ in each

model also varies between the six field experiments, from 0.30 to 0.51 (TG), 0.21 to 0.37 (Whit), 0.20 to 0.37 (Lipp), 0.37 to 0.61 (Bald), 0.41 to 0.66 (BJ), and 0.26 to 0.56 (Jan), suggesting that γ likely changes with wave conditions and details of the bathymetry. For the data for which spectra were available, the best-fit γ are not correlated with the difference between the centroidal and peak frequencies. To avoid possible ambiguities in selecting a representative frequency for wave fields with multiple spectral peaks, the centroidal frequency is used here when possible.

5.2.2 Correlation with S_o and H_o

The two Duck experiments (i.e., SandyDuck and Duck94), which had the most sensors distributed evenly across the surfzone, are used to develop a universal empirical relationship between γ and the incident wave field. This relationship is then tested at the other four experiments.

Previous studies [*Battjes and Stive,* 1985; *Nairn*, 1990] showed that γ depends on the deep-water wave steepness, $S_o=H_o/L_o$. At all experiments, the deep-water wave height H_o was estimated by unshoaling the observations from the deepest sensor assuming conservation of wave energy flux. Here, the correlations between the best-fit γ and H_o at SandyDuck and Duck94 (average correlations of the unbinned data are 0.67 and 0.66, respectively) are about 50% larger than those between best-fit γ and S_o. The dependence of the best-fit γ on the offshore wave height may be related to variations in B (which relates the dissipation in a hydraulic jump to that in a breaking wave) or to other parameterized processes, such as infragravity wave generation and differences in breaking-wave dissipation in spilling and plunging waves.

For all models, γ increases almost linearly with increasing H_o for small waves, then becomes nearly constant for large H_o (e.g., Figure 5). Standard deviations of binned values of γ range between 0.01 and 0.10 and are larger for small H_o (Figure 6). The relationship between γ and H_o is described well by a hyperbolic tangent curve

$$\gamma = a + b\left[\tanh\left(cH_o\right)\right], \tag{23}$$

where *a, b,* and *c* are determined (using a least squares fit) for each model and experiment. Correlations between the binned values of best-fit γ and (23) usually are greater than 0.9, and average (for the two experiments) correlations between the unbinned data and the universal curves are 0.77 (TG), 0.60 (Whit), 0.67 (Lipp), 0.81(Bald), 0.73 (BJ), and 0.85 (Jan) (e.g., Figure 6). Best-fit γ's also are correlated with the inverse Iribarren number, $\dfrac{1}{\xi} = \dfrac{\sqrt{S_o}}{\beta_{av}}$, where β_{av} is the surfzone averaged beach slope [*Raubenheimer et al.*, 2001], for the four US experiments. However, on beaches with large or multiple bars (e.g., Egmond and Terschelling), β_{av} is poorly defined owing to its dependence on the definition of the offshore boundary of the surfzone. Furthermore, if the offshore boundary of the surfzone is estimated from the predicted wave energy dissipation, the location depends on γ.

The empirical curves for γ based on H_o differ slightly for all models at the two experiments (Figure 6, compare the grey dashed curve with the solid black curve in each panel). Universal, experiment-averaged curves for γ (Figure 7) are obtained for each model

by averaging the curves from the two Duck experiments, rather than by fitting to the raw data, because the two experiments had different numbers of data points. The coefficients for the universal curves are given in Table 4.

5.2.3 Application of the γ Curves

Application of the universal curves for each model at SandyDuck and Duck94 reduces the prediction errors by 7 to 36% (*mean* = 18%) and 7 to 28 % (*mean* = 17%), respectively, relative to the default models. To examine the applicability of the new parameterization of γ to other sites, the universal curves are applied at the other four experiments.

5.2.3a Comparison with Egmond and Terschelling

Using the universal curves instead of default γ improves the results for five (TG, Whit, Bald, BJ, and Jan) of the wave models at Egmond and for all six of the wave models at Terschelling (Table 5). Although the wave models were initialized far offshore of the surfzone at these two experiments, including bottom stress estimates [e.g., *Thornton and Guza*, 1983] in the models has little effect on either the predicted wave heights or the best-fit γ. For unknown reasons, at Egmond the skill of the Lipp model decreased using the universal curve relative to using the default γ.

Similar to previous results [Ruessink, personal communication], use of γ = 0.42 in the TG model for large deep-water waves (H_o > 1.5 m) causes too much dissipation in the outer surfzone, and thus the predicted waves are smaller than the observed waves. For data records at Egmond when H_o > 1.5 m, using γ determined from the universal curves with five of the wave models (i.e., excluding the Lipp model) reduces median errors by 35% relative to using default values for γ, with the largest improvement (60%) for the TG model and a slight reduction in accuracy (-3%) for the BJ model. Thus, use of the default value of γ = 0.42 in the TG model may result in significant underprediction of surfzone wave heights when H_o is large.

5.2.3b Comparison with SwashX and NCEX

Using the universal curves reduces model errors relative to the errors using default γ for all six of the wave models at NCEX, but for unknown reasons only one (Jan) of the wave models at SwashX (Table 5).

6. Discussion
6.1 Universal Curves

Compared with default γ, using the universal curves reduces prediction errors most for TG, Bald, and Jan (e.g., highest BSS values in Table 5) and least for Whit, Lipp, and BJ. For the models using a constant default γ (i.e., TG, Whit, and Lipp), TG may show more improvement owing to its wider range of best-fit γ (Figure 4, compare the solid curve with the dotted and dashed-dotted curves, and Figure 6A-C, compare the maximum value of γ reached for each of the curves). For the models using a variable default γ (i.e., Bald, BJ, and Jan), BJ may show the smallest improvement because the default values [i.e., (15) and (16)] were determined using BJ, and therefore do not take into account changes in γ owing to the modifications of the dissipation in Bald and Jan.

The universal curves for γ typically overestimate the best-fit values for SandyDuck, Egmond, Terschelling, and NCEX, and underestimate the values for SwashX and Duck94 (e.g., Figure 8). The larger spread at Egmond and Terschelling (Figure 8, cyan and yellow curves) likely occurs because the lack of sensors in shallow water decreases best-fit γ for small H_o. The spread in γ owing to using the universal curves rather than the best-fit values is roughly similar for the other five models (not shown).

Including the NCEX data in developing the universal curves does not change the curves significantly. Inclusion of the data from Egmond and Terschelling reduces the values of γ estimated for small H_o (see above), but does not change significantly the values estimated for $H_o > 1.5$ m.

When calibrated with the same procedure [e.g. the 3 free parameters (a,b,c)] all models have similar accuracy (Table 6).

6.2 Number of Sensors Needed for Tuning

The usefulness of model tuning depends on the number and locations of the observations used. Here, the rms error between the observed wave heights and the TG model predictions at all sensor locations increases when fewer than 4 locations spanning the surfzone are used in tuning the model (Figure 9A). However, the accuracy of the model tuned with data from only two locations is higher than for a constant default γ of 0.42 (Figure 9A, see the data point at 0 tuning locations).

To optimize predictions of the cross-shore distribution of wave heights, data are needed from at least two sensors spanning the surfzone, which changes in width and location with changing wave conditions and tidal levels. At SandyDuck, three sensors are needed close to the shore to ensure that at least two sensors are located in the surfzone during all tidal stages for small waves, and at least one (or ideally two) sensors are needed in deeper water to span the width of the wider surfzone during large waves (not shown). The scatter in best-fit γ increases relative to that calculated using all sensors (i.e., Figure 4) for small waves when only two tuning locations are located near the shoreline (Figure 9B, compare the dashed with the solid curve), and for most wave heights when only one offshore tuning location is used (Figure 9B, compare the dashed-dotted with the solid curve). Similar results are found using the Bald model.

6.3 Model Accuracy as a Function of Water Depth

Model accuracy decreases with decreasing water depth, partially owing to the accumulation of errors with increasing distance from the location of model initialization. However, the model predictions show similar deviations from the observations at all experiments even though the bathymetries and offshore initialization depths are different.

The mean and rms errors using the default and universal γ's with the TG model are small for $h > 2$ m (Figure 10), with rms errors increasing with decreasing depth for $h < 2$ m (Figure 10B, D, F, and H). Except at NCEX (Figure 10G), using the universal curves results in slight overprediction of the observed wave heights for roughly $h > 1$ m (Figure 10A, C, and E), and underprediction of the wave heights for $h < 1$ m. Differences between the patterns of under- and overprediction at each experiment are at least partly related to the value of γ used. For example, using γ = 0.32 instead of γ = 0.42 in the TG model at NCEX, which has a smaller best-fit γ than SwashX (not shown), results in similar mean and rms errors in $h = 1$ m for both experiments. Furthermore, the pattern of over- and

underprediction across the surfzone may be caused by cross-shore variations in the observed ratio of wave height to water depth (γ_{obs}). For example, the underprediction of wave heights in shallow water may be related to the shoreward increase in γ_{obs} [*Raubenheimer et al.,* 1996].

At Duck94 and SandyDuck mean errors at a given water depth are larger on the shoreward edge of the bar trough than at more offshore locations (Figure 10, the solid black curves have smaller mean errors than the dashed black curves for all water depths in A and C). Thus, although the pattern of over- and underprediction (Figure 10A, C, E, and G, black curves) is similar at different experiments, it is sensitive to cross-shore location on barred bathymetries. Previous studies on multi-barred bathymetries [*Ruessink et al.,* 2003] found a similar increase in overprediction of wave heights in bar troughs.

Using universal γ decreases both the mean and rms errors in most water depths (Figure 10, compare black with grey curves). Although using best-fit γ further decreases the errors, no tuned, universal, or default model has the smallest mean or rms errors in all water depths for all experiments (not shown). Egmond and Terschelling were not used in this analysis because sensors rarely were located in $h < 2$ m.

7. Conclusions

Several parametric models for the transformation of wave heights across the surfzone were tested and calibrated with observations collected along cross-shore transects at 6 experiments on barred and unbarred beaches. Models using previously calibrated (default) values for the free parameter γ predict the cross-shore distribution of the observed wave heights with median rms errors between 10% and 20%. Tuning the free parameter in each model reduces the errors by approximately 50%. Root-mean-square errors for all models are small in water depths $h > 2$ m, and increase with decreasing depth for $h < 2$ m. To tune the models accurately, data must span the surfzone, which may require at least three to five sensors depending on tidal and wave height ranges. No tuned or default model provides the best predictions for all data records or at all experiments.

Best-fit γ are correlated with the deep-water wave height, H_o. Relative to using the default values of γ, estimating γ using universal curves based on H_o from two experiments at Duck, NC usually reduces errors for all models.

Appendix A: Poorly Defined Fits

To tune the models, the rms error must have a well-defined minimum at some γ (e.g., Figure A1A). However, when only one wave height observation is located within the surfzone the minimum in the rms error curve often becomes broad (e.g., Figure A1B),

resulting in a poorly defined best-fit γ. In this case, small errors in the measured bathymetry or wave heights can change the estimated value of γ significantly.

At SandyDuck, γ always was well defined. However, at Duck94, instruments often were sparsely spaced both over the sandbar and on the steep foreshore, where there was a shore break (and narrow surfzone) during periods of high tides and small waves. Thus, there are a significant number of data records for which only one sensor was located in the surfzone. To avoid using poorly fit γ's in determining the universal curves, Duck94 data records with broad minima in the rms error curve were excluded. Visual examination of the results (not shown) indicates a reasonably clear delineation between the regimes of well and poorly defined fits.

All data records from Egmond and Terschelling were used. The data records with small H_o were retained to ensure that the models did not dissipate too much energy over the relatively long distance between the location of model initialization and the outer edge of the surfzone.

At NCEX and SwashX, γ becomes poorly defined at low tides when the five shallowest sensors were above mean sea level. Therefore, data records from these experiments were included in the analysis only when at least one of the nearshore sensors was submerged.

Appendix B: Error Calculation

The sensitivity of the results to the error metric is evaluated by calculating the root-mean-square (rms) error in four ways: percent (a) weighted [i.e., (20)] and (b) unweighted error, and absolute (c) weighted and (d) unweighted error, given as

$$Unweighted\ Percent\ Error = \sqrt{mean\left[\left((obs_n - pred_n)/obs_n\right)^2\right]} *100\%, \qquad (B1)$$

$$Weighted\ Absolute\ Error = \sqrt{\sum_n \left[(obs_n - pred_n)^2 * weight_n\right]}, \qquad (B2)$$

and

$$Unweighted\ Absolute\ Error = \sqrt{mean\left[(obs_n - pred_n)^2\right]}. \qquad (B3)$$

Model errors and best-fit γ are similar for all error metrics. For example, the mean best-fit γ values for the TG and Bald models for SandyDuck (Figure B1A & B1B) are similar for all four error metrics. Although γ varies slightly for individual data records, the histograms of best-fit γ are not significantly different for any of the error metrics (Figure B1C & B1D). Similar results were found for the other models, and at Duck94.

Using a percent error metric [i.e., (20)], which gives more weight to the smaller wave heights in shallower water, minimizes the mean and rms errors in shallow water at the expense of larger errors in deeper water. However, the difference in errors between error metrics is significantly smaller than the difference in errors between tuned and default models.

References
Baldock, T.E., P. Holmes, S. Bunker, and P. Van Weert, Cross-shore hydrodynamics within an unsaturated surf zone, *Coastal Eng., 34*, 173-196, 1998.

Battjes, J.A. and J.P.F.M. Janssen, Energy loss and setup due to breaking of random waves, International Conference on Coastal Engineering, ASCE, 569-587, 1978.

Battjes, J.A. and M.J.F. Stive, Calibration and verification of a dissipation model for random breaking waves*, J. Geophys. Res., 90*, 9159-9167, 1985.

Cacina, N., Test and evaluation of surf forecasting model, M.Sc. Thesis, Naval Postgraduate School, Monterey, CA, 41 pp., 1989.

Elgar, S., R.T. Guza, and R.J. Seymour, Groups of waves in shallow water, *J. Geophys. Res. 89*, 3623-3634, 1984.

Elgar, S., R.T. Guza, W.C. O'Reilly, B. Raubenheimer, and T.H.C. Herbers, Wave energy and direction observed near a pier, *ASCE J. Waterway, Port, Coastal, and Ocean Engineering, 127*, 2-6, 2001.

Janssen, T.T. and J.A. Battjes, A note on wave energy dissipation over steep beaches, *Coastal Eng.,* Submitted.

Kennedy, A.B, Q. Chen, J.T. Kirby, and R.A. Dalrymple, Boussinesq modeling of wave transformation, breaking and runup. I: One dimension. *ASCE J. Waterway, Port, Coastal, and Ocean Engineering*, *126*, 39-47, 2000.

LeMehaute, B., On non-saturated breakers and the wave run-up, International Conference on Coastal Engineering, ASCE, 77-92, 1962.

Lippmann, T.C., A.H. Brookins, and E.B. Thornton, Wave energy transformation on natural profiles, *Coastal Eng., 27*, 1-20, 1996.

Longuet-Higgins, M.S. and R.W. Stewart, Radiation stress and mass transport in gravity waves, with application to 'surf-beats,' *J. Fluid Mech., 13*, 481-504, 1962.

Longuet-Higgins, M.S. and R.W. Stewart, Radiation stresses in water waves: A physical discussion with applications, *Deep Sea Res., 11*, 529-562, 1964.

Miche, M. Le Pouvoir Reflechissant des Ouvrages Maritimes Exposes a l'Action de la Houle *Annals des Ponts et Chaussess, 121e* Annee, 285-319, 1951. (Translated by Lincoln and Chevron, University of California, Berkeley, Wave Research Laboratory, Series 3, Issue 363, June 1954.)

Murphy, A.H. and E.S. Epstein, Skill scores and correlation coefficients in model verification, *Mon. Weather Rev., 117,* 572-581, 1989.

Nairn, R.B, Prediction of cross-shore sediment transport and beach profile evolution, Ph.D thesis, Imperial College, London, 391 pp., 1990.

Nwogu, O., Alternative form of Boussinesq equations for nearshore wave propagation, *ASCE J. Waterway, Port, Coastal, and Ocean Engineering, 119*, 618-638, 1993.

Raubenheimer, B., Observations and predictions of fluid velocities in the surf and swash zones, *J. Geophys. Res., 107*, 3190, doi:10.1029/2001JC001264, 2002.

Raubenheimer, B., R.T. Guza, and S. Elgar, Wave transformation across the inner surf zone, *J. Geophys. Res., 101*, 25589-25597, 1996.

Raubenheimer, B., R.T. Guza, and S. Elgar, Field observations of wave-driven setdown and setup, *J. Geophys. Res., 106,* 4629-4638, 2001.

Roelvink, J.A., Dissipation in random wave groups incident on a beach, *Coastal Eng., 19,* 127-150, 1993.

Ruessink, B.G, K.T. Houwman, and P. Hoekstra, The systematic contribution of transporting mechanisms to the cross-shore sediment transport in water depths 3 to 9 m, *Mar. Geol. 152*, 295-324, 1998.

Ruessink, B.G., J.R. Miles, F. Feddersen, R.T. Guza, and S. Elgar, Modeling the alongshore current on barred beaches, *J. Geophys. Res., 106*, 22451-22463, 2001.

Ruessink, B.G., D.J.R. Walstra, and H.N. Southgate, Calibration and verification of a parametric wave model on barred beaches, *Coastal Eng., 48*, 139-149, 2003.

Sallenger, A.H. and R.A. Holman, Wave energy saturation on a natural beach of variable slope, *J. Geophys. Res., 90*, 11939-11944, 1985.

Stive, M.J.F. and H.J. De Vriend, Shear stress and mean flow in shoaling and breaking waves, International Conference on Coastal Engineering, ASCE, 594-608, 1994.

Stoker, J.J., *Water Waves*, Interscience, New York, 1957.

Thomson, J., S. Elgar, T.H.C. Herbers, B. Raubenheimer, and R.T. Guza, Tidal modulation of infragravity waves via nonlinear energy losses in the surfzone, *Geophysical Research Letters, 33*, L05601, doi:10.1029/2005GL025514, 2006.

Thornton, E.B. and R.T. Guza, Transformation of wave height distribution, *J. Geophys. Res., 88*, 5925-5938, 1983.

Whitford, D.J., Wind and wave forcing of longshore currents across a barred beach, Ph.D Thesis, Naval Postgraduate School, Monterey, CA, 205 pp., 1988.
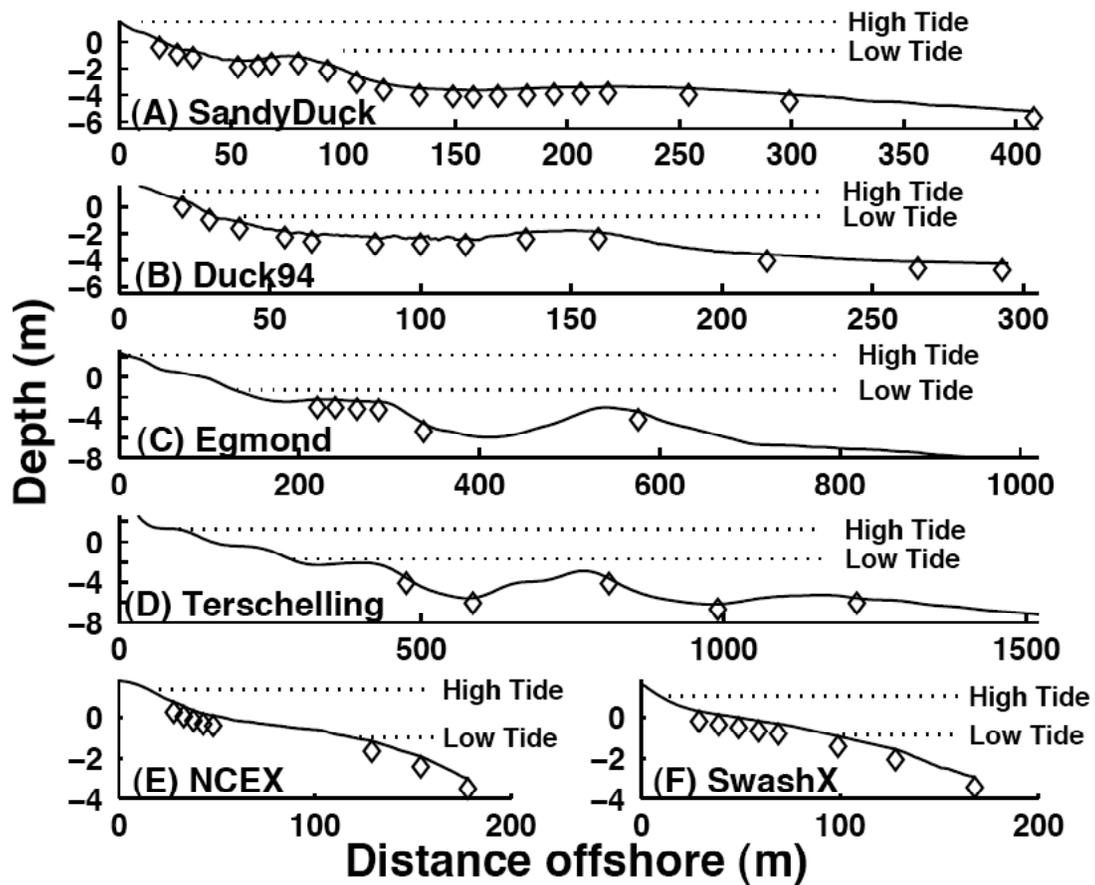
Figure 1: Water depth (relative to mean sea level) (solid curves), pressure sensor locations (diamonds), and tidal levels (dotted lines) *versus* distance offshore for (A) SandyDuck, (B) Duck94, (C) Egmond, (D) Terschelling, (E) NCEX, and (F) SwashX. The most offshore sensors for Egmond (15 m depth) and Terschelling (9 m depth) are not shown.
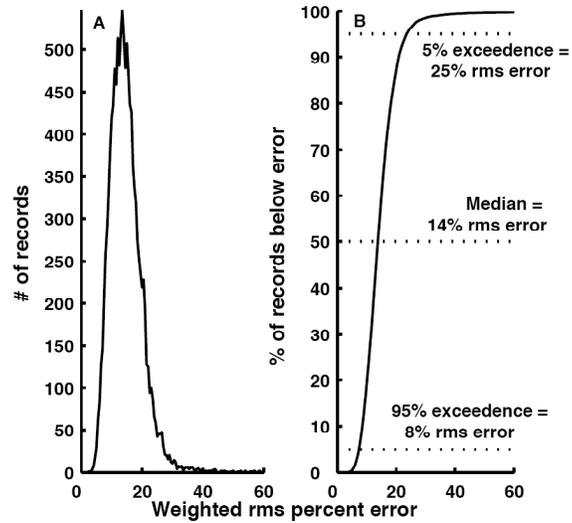
Figure 2: (A) Number of records (e.g., histogram) and (B) cumulative sum of the number of records (as a percent of the total number of records) *versus* the weighted rms percent errors for the default Whit model for the SandyDuck experiment.
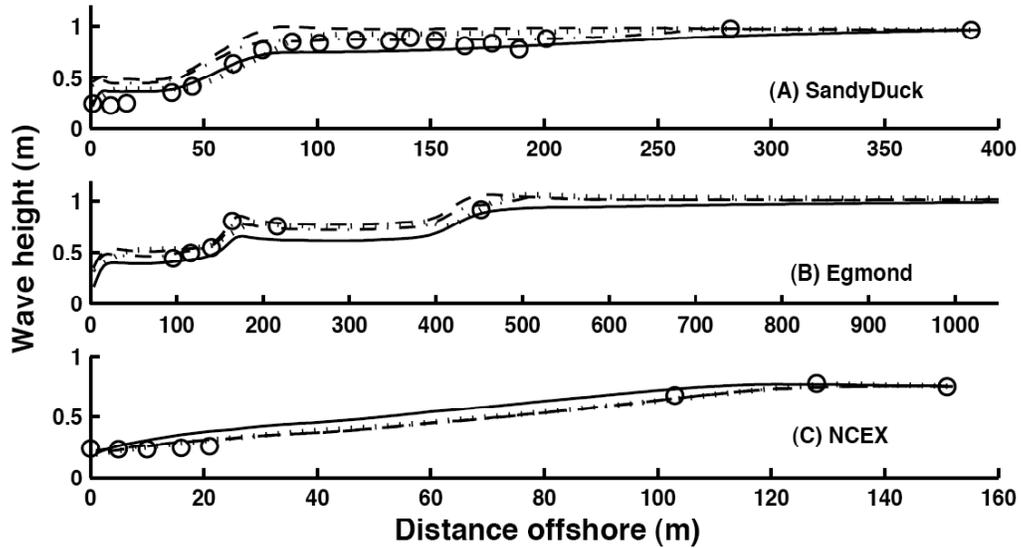
Figure 3: Observed (circles) and predicted [TG (solid curves), Bald (dashed curves), Nairn (dashed-dotted curves), and Lipp (dotted curves) models] wave heights *versus* distance offshore for (A) SandyDuck Sep 27 19:51, (B) Egmond Nov 4 15:00, and (C) NCEX Nov 10 08:34 (local standard time).
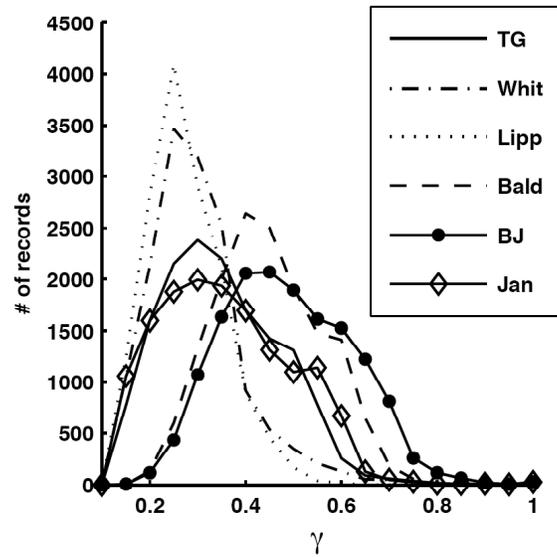
Figure 4: Number of records (histograms) *versus* the best-fit γ for the TG (solid curve), Whit (dashed-dotted curve), Lipp (dotted curve), Bald (dashed curve), BJ (solid curve with circles), and Jan (solid curve with diamonds) models for SandyDuck. Histograms of the best-fit γ at the other experiments are similar.
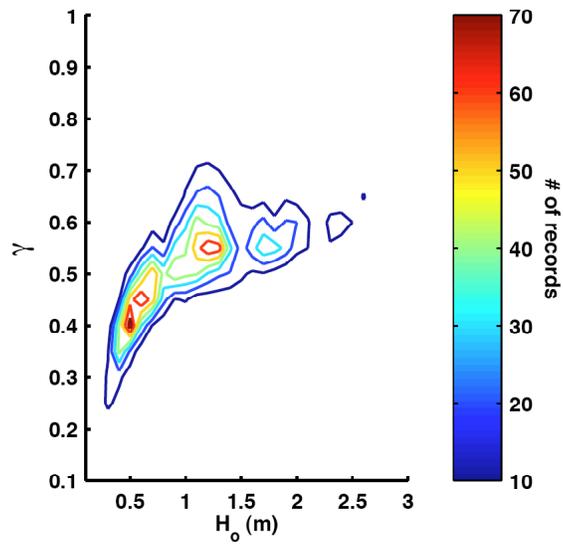
Figure 5: The number of data records (contours) as a function of the best-fit γ and the deep-water wave height H_o for the TG model at Duck94. Red contours are the largest number of data records and dark blue contours are the smallest number of data records (color scale is on the right hand side). The pattern observed between the best-fit γ and H_o is similar for the other models and experiments.
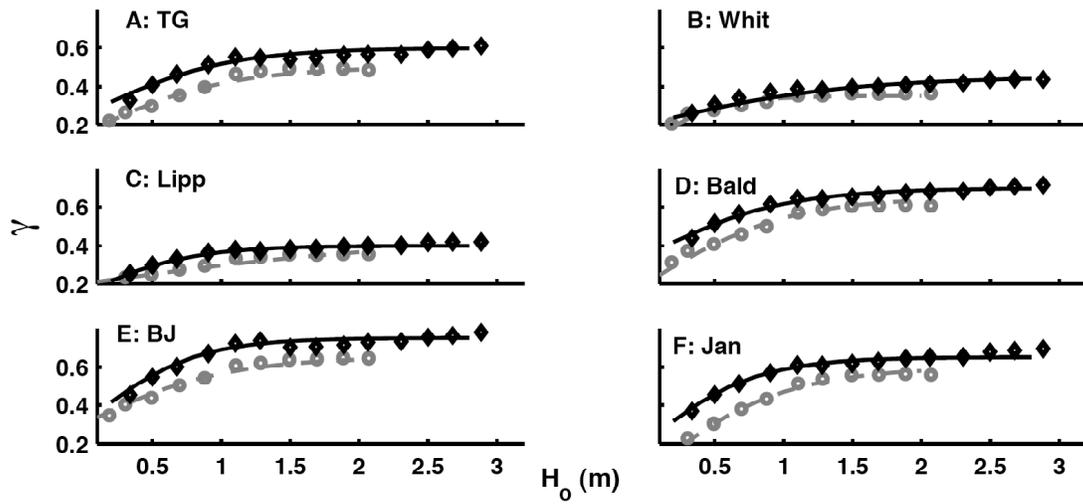
Figure 6: Empirical hyperbolic tangent curves fit to data from SandyDuck (grey dashed curve) and Duck94 (black solid curve) for the (A) TG, (B) Whit, (C) Lipp, (D) Bald, (E) BJ, and (F) Jan models. Symbols (diamonds, Duck94; and circles, SandyDuck) and vertical bars represent the mean values and standard deviations of the best-fit γ in bins of $H_o \pm 0.10$ m.
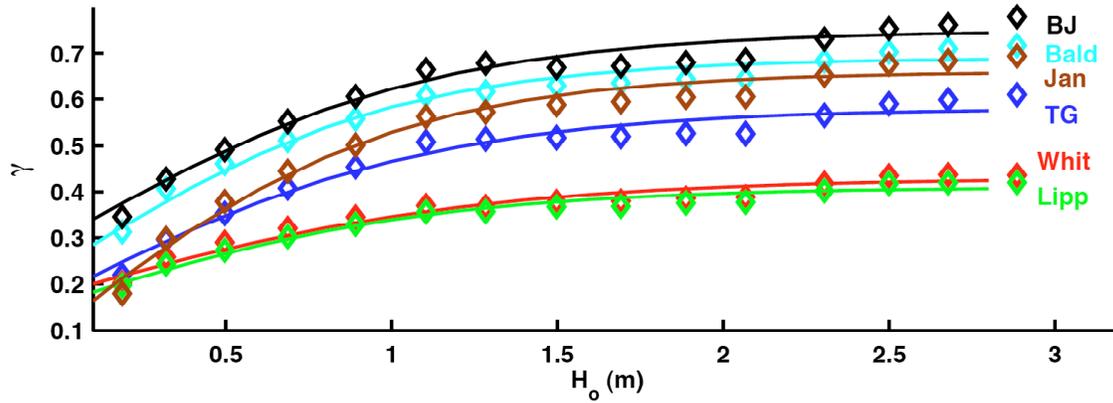
Figure 7: Universal empirical curves averaged over the two Duck experiments for the TG (blue), Whit (red), Lipp (green), Bald (cyan), BJ (black), and Jan (brown) models. The diamonds are the average of the binned values of the best-fit γ from Duck94 and SandyDuck (i.e., Figure 6, diamonds and circles) in bins of $H_o \pm 0.10$ m. Values of the coefficients for the universal curves are given in Table 4.
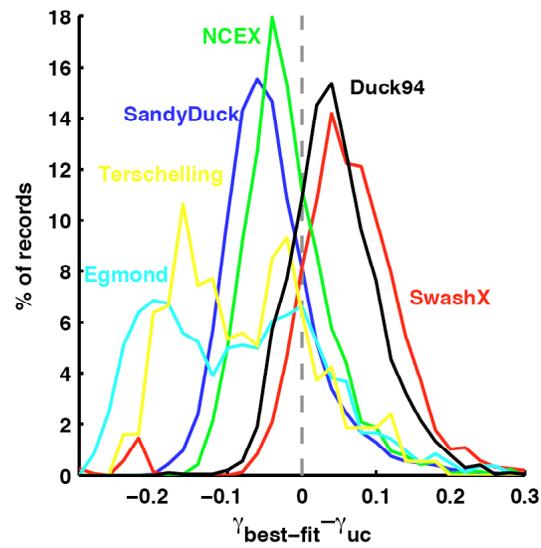
Figure 8: Percent of data records for the TG model *versus* the difference between the best-fit γ ($\gamma_{\text{best-fit}}$) and the γ based on the universal curves (γ_{uc}, e.g., Figure 7) for SandyDuck (blue curve), Duck94 (black curve), SwashX (red curve), NCEX (green curve), Egmond (cyan curve), and Terschelling (yellow curve). The vertical grey dashed line represents perfect agreement, and values less (greater) than 0 occur when the universal curve over-(under-) estimates the best-fit γ.
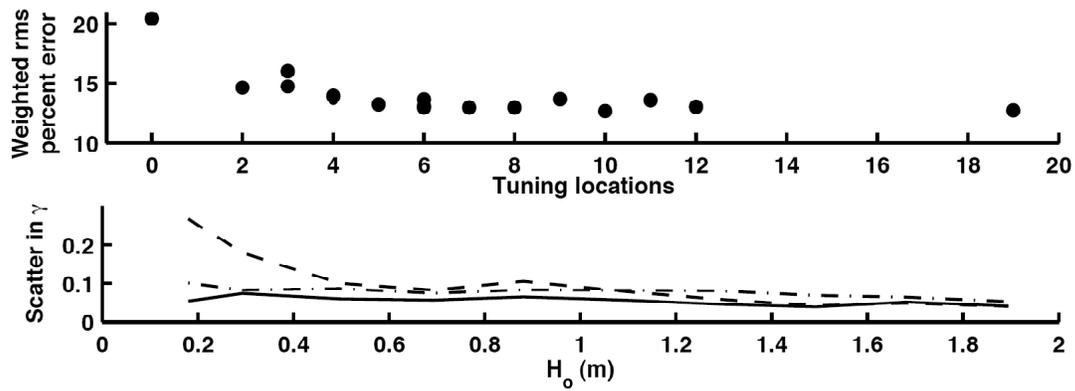
Figure 9: (A) Weighted rms percent error *versus* the number of tuning locations used with the TG model at SandyDuck, and (B) the scatter (defined as one standard deviation of the best-fit γ) of the best-fit γ *versus* the deep-water wave height H_o when the TG model is tuned using data from all sensors (solid curve), 2 offshore sensors and 2 shallow sensors (dashed curve), and 3 shallow sensors, but only 1 offshore sensor (dashed-dotted curve). In (A) two different sets of sensors were tested (i.e., there are 2 symbols in the vertical) for 3, 4, and 6 tuning locations.
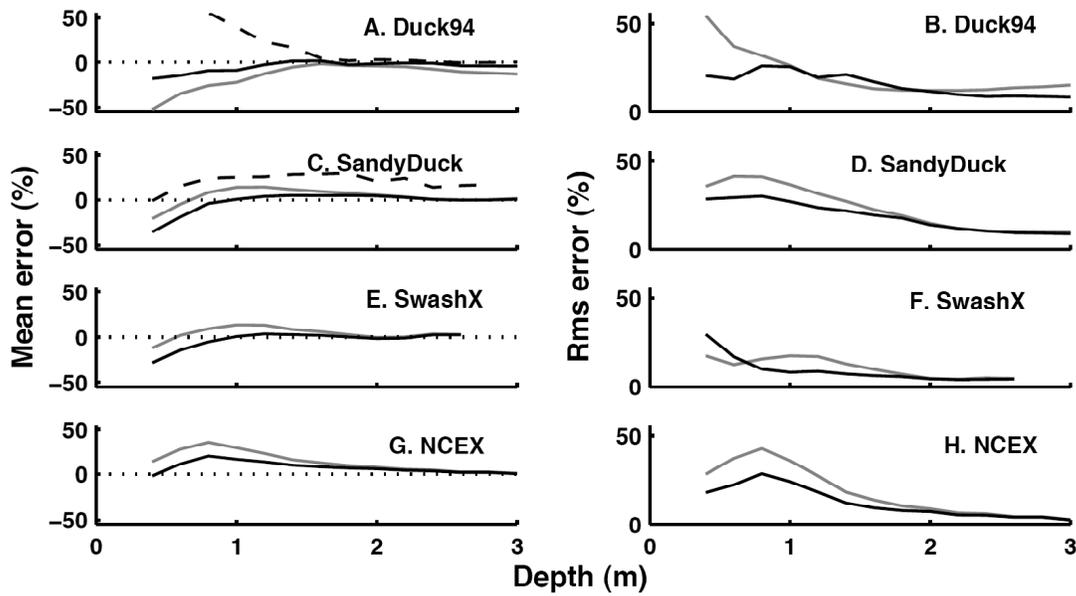
Figure 10: Mean (A, C, E, & G) and rms (B, D, F, & H) prediction errors for the TG model with γ = 0.42 (default, grey curves) and γ from the universal curve (black curves) *versus* water depth for Duck94 (A & B), SandyDuck (C & D), SwashX (E & F), and NCEX (G & H). Positive (negative) mean errors correspond to overprediction (underprediction). Mean errors for Duck94 and SandyDuck (A and C, respectively) were calculated for locations shoreward (dashed black curves) and seaward (solid black curves) of the sandbar trough.
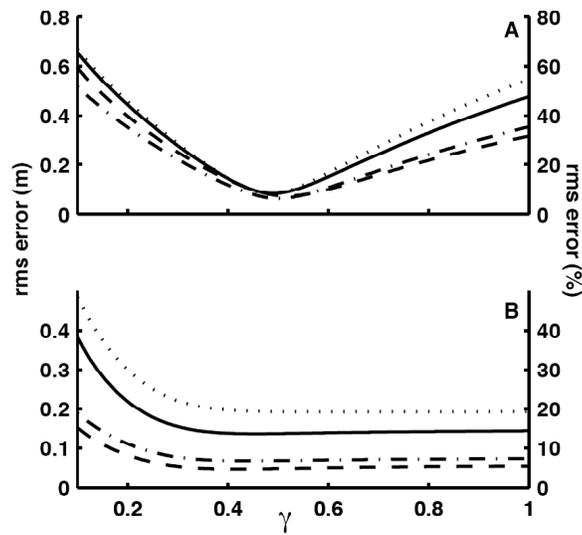
Figure A1: Root-mean-square (rms) errors *versus* γ for a data record from Duck94 with (A) a well defined best-fit γ and (B) a poorly defined best-fit γ. The four curves correspond to the four error metrics: weighted rms percent error (solid curves, right axes), percent error (dotted curves, right axes), weighted rms absolute error (dashed curves, left axes), and absolute error (dashed-dotted curves, left axes).
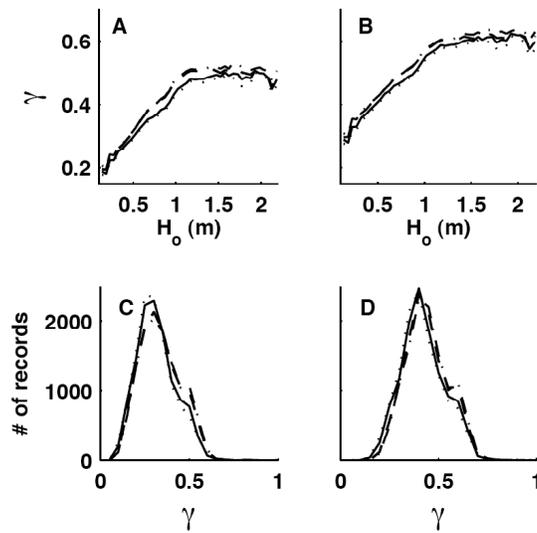
Figure B1: The mean best-fit γ (±0.1 m bins in H_o) *versus* deep-water wave height H_o at SandyDuck for (A) TG and (B) Bald models for the four different error metrics (solid curve is eqn (20), dotted curve is eqn (B1), dashed curve is eqn (B2), and dashed-dotted curve is eqn (B3)) and (C & D) the corresponding number of records (i.e., histograms) *versus* best-fit values of γ at SandyDuck.

Table 1: Wave Model Summary

Wave Model	Energy Eqn.	Dissipation	γ
TG	(1)	(7) and (8)	0.42 or tuned
Whit	(1)	(7) and (9)	0.34 or tuned
Lipp	(2) and (3)	(13) and (14)	0.32 or tuned
BJ	(1)	(4) - (6)	Tuned
BS	(1)	(4) - (6)	(15)
Nairn	(1)	(4) - (6)	(16)
Bald	(1)	(10) and (11)	(16) or tuned
Rues	(1)	(5) and (10)	(17)
Jan	(1)	(11) and (12)	(16) or tuned

Table 2: 95% exceedence (minimum), median, and 5% exceedence (maximum) percent errors for the eight default wave models for all six experiments.

		TG	Whit	Lipp	Bald	Rues	Nairn	BS	Jan
SandyDuck	Min	8	8	7	7	7	7	8	8
	Med	14	14	14	14	15	12	16	18
	Max	26	25	26	33	28	26	30	37
Duck94	Min	6	9	8	5	5	5	5	5
	Med	14	15	13	10	11	9	9	11
	Max	22	25	23	23	19	18	19	27
SwashX	Min	6	4	3	4	3	3	5	6
	Med	13	16	14	12	15	13	12	12
	Max	23	34	26	21	23	21	21	23
NCEX	Min	7	5	3	2	3	2	7	7
	Med	16	12	10	7	11	8	15	16
	Max	43	28	27	21	25	22	38	41
Egmond	Min	5	4	4	5	2	3	3	7
	Med	17	18	19	19	17	16	17	23
	Max	45	49	70	49	47	46	46	52
Terschelling	Min	6	6	6	6	6	6	5	6
	Med	16	17	17	17	19	16	17	20
	Max	33	41	44	44	40	38	39	42

Table 3: Median percent errors for the six tuned models for all six experiments. Results for BJ, BS, and Nairn are identical, and are listed as BJ.

	TG	Whit	Lipp	Bald	BJ	Jan
SandyDuck	9	9	9	9	9	9
Duck94	8	11	9	6	7	6
SwashX	9	12	11	9	9	7
NCEX	4	7	5	3	4	4
Egmond	5	6	6	5	5	5
Terschelling	8	12	10	10	8	10

Table 4: Coefficients for the universal empirical curves $\gamma = a + b\left[\tanh\left(cH_o\right)\right]$ for the six models.

	TG	Whit	Lipp	Bald	BJ	Jan
a	0.18	0.18	0.16	0.24	0.30	0.11
b	0.40	0.25	0.25	0.45	0.45	0.55
c (m^{-1})	0.90	0.80	0.90	1.00	0.90	1.00

Table 5: The BSS values (i.e., percent error reduction) using the universal empirical curves instead of the default γ at the 4 experiments not used in calibration. Here the tuned BJ model is compared with the default Nairn because (16) was obtained using more data than (15). Positive values represent model improvement.

	TG	Whit	Lipp	Bald	BJ	Jan
Egmond	27	5	-20	18	3	25
Terschelling	16	6	6	9	3	18
NCEX	48	1	2	22	1	53
SwashX	-7	-49	-49	-39	-8	13

Table 6: Median percent errors for the six tuned models using the universal empirical γ curves.

	TG	Whit	Lipp	Bald	BJ	Jan
SandyDuck	11	12	11	11	11	12
Duck94	10	13	11	9	9	9
NCEX	8	12	10	6	8	8
SwashX	14	23	21	17	14	10
Egmond	12	17	23	16	15	17
Terschelling	13	16	16	16	15	16