

shallow-water-waves: Wave Height Distributions on Shallow Foreshores

Technical Overview and Scientific Context

Purpose and Functionality of the Computational Tool

The software tool "shallow-water-waves" (Github: <https://github.com/ricbencar/shallow-water-waves>) provides a computational implementation of the wave height distribution model developed by Jurjen A. Battjes and Heiko W. Groenendijk (Battjes & Groenendijk, 2000).

It also implements a two-tiered "overshoot prevention" logic to ensure results are physically realistic and consistent with established theory (Caires & Van Gent, 2012).

The primary function of this tool is to compute local wave height statistics on shallow foreshores—the gently sloping nearshore regions where wave dynamics are strongly influenced by the seabed.

The model's output provides quantitative data for a wide range of coastal engineering applications, including the design and risk assessment of dikes, breakwaters, and other coastal defense structures where parameters such as wave run-up and overtopping rates are critical design considerations.

The software is available as both a command-line interface (CLI) for integration into automated analysis workflows and a graphical user interface (GUI) for interactive, case-by-case analysis.

Rayleigh Distribution Not Applicable in Shallow Water

The statistical description of wave heights is a foundational element of coastal engineering. **In deep water, where the seabed has negligible influence on wave propagation, the short-term distribution of individual wave heights is described with reasonable accuracy by the Rayleigh distribution** (John William Strutt, 1885), a model first derived for sea waves by M. S. Longuet-Higgins (Longuet-Higgins, 1952).

This distribution is theoretically derived from the assumption that the sea surface elevation can be modeled as a Gaussian random process with a narrow energy spectrum. Under these conditions, all characteristic wave heights (e.g., the significant wave height, the average of the highest 10% of waves) are related by known constants, which simplifies design calculations.

However, as waves propagate from deep water onto a shallow foreshore, the physical processes that govern their behavior change substantially, rendering the fundamental assumptions of the Rayleigh distribution invalid. Two key phenomena are responsible for this departure (Battjes & Groenendijk, 2000):

1. **Nonlinear Shoaling and Triad Interactions:** As the water depth decreases, waves begin to interact with the seabed. This interaction, known as shoaling, increases wave heights and enhances nonlinear effects. Specifically, triad wave-wave interactions transfer energy to higher harmonics, distorting the wave profile into a non-Gaussian shape characterized by peaked crests and shallow, flat troughs. Consequently, the sea surface can no longer be considered a linear Gaussian process.
2. **Depth-Induced Breaking:** The most significant process in shallow water is depth-induced wave breaking. The water depth imposes a physical upper limit on the height a wave can attain. In a random sea state, the largest waves, which would have continued to propagate in deep water, are forced to break, dissipating a significant amount of their energy into turbulence (Battjes & Groenendijk, 2000). This process effectively truncates the upper tail of the wave height distribution, meaning that extremely high waves become far less probable than predicted by the unbounded Rayleigh model.

The accuracy loss of the Rayleigh distribution in the nearshore zone has significant engineering significance. An accurate statistical description of wave heights is essential for the design of coastal structures. Design criteria concerning wave forces, structural stability, wave run-up, and overtopping rates all depend on characteristic wave heights, such as the significant wave height (H_s) or heights with a low probability of exceedance (e.g., $H_{1/100}$ or H_{max}).

Using the Rayleigh distribution in these conditions can lead to a significant over-estimation of the highest waves, resulting in overly conservative and expensive designs, or in some cases, a mischaracterization of the wave energy distribution (Battjes & Groenendijk, 2000).

Point Model Assumptions, Applicability and Principles

The Battjes and Groenendijk model, and by extension this software, is built upon a "point model" philosophy. This is a critical, simplifying assumption that dictates the model's structure, inputs, and domain of applicability. As stated in the original publication:

"The parameterisation is based on the assumption of slow evolution, such that the distribution depends on local parameters only, regardless of the history of the waves in deeper water (a so-called point model)" (Battjes & Groenendijk, 2000).

This means the model predicts the complete wave height distribution at a single, specific point using only the parameters that describe the conditions at that selected point:

- The local water depth, d .
- The local total wave energy, represented by the variance of the surface elevation, m_0 .

- The local bottom slope, $\tan(\alpha)$.

The underlying assumption is that the wave field has propagated over a sufficient distance on a bathymetry of simple, slowly varying geometry, allowing it to reach a state of local equilibrium with the seabed. This approach makes the model computationally efficient and effective for its intended application on plane beaches.

However, this assumption is also the direct source of the model's primary limitation: **the model cannot account for the "memory" of the wave field in situations with complex or rapidly changing bathymetry, such as in the presence of offshore bars and troughs, where the wave height distribution at one point is strongly influenced by conditions at another** (Battjes & Groenendijk, 2000).

Furthermore, this assumption is the direct cause of the model's documented underperformance on flat or very gently sloping seabeds, a critical constraint discussed in detail by Caires & Van Gent (2012). Understanding this "point model" basis is essential for correctly applying the software and interpreting its results.

Composite Weibull Distribution (CWD) Theoretical Foundation

Two-Population Hypothesis for Shallow-Water Waves

To address the shortcomings of the Rayleigh distribution in shallow water, Battjes and Groenendijk proposed a model based on a realistic physical hypothesis: the wave field on a shallow foreshore can be conceptualized as a mixture of two distinct statistical populations:

- **Non-Breaking Waves:** This group consists of the smaller waves in the spectrum whose heights are not strongly affected by the limited water depth. They propagate relatively undisturbed, and their statistical distribution is assumed to adhere to the Rayleigh model.
- **Breaking/Broken Waves:** This group comprises the larger waves whose heights are actively limited by the local depth. These waves have either already broken or are in the process of breaking, causing their statistical distribution to deviate sharply from the Rayleigh form, with a much lower probability of very high waves.

The Composite Weibull Distribution (CWD) is the mathematical formalization of this two-population concept. It combines two separate Weibull distributions, each representing one of the populations, and matches them at a "transitional wave height," H_{tr} , which acts as the demarcation point between the two regimes (Battjes & Groenendijk, 2000).

Mathematical Formulation of the CWD

The cumulative distribution function (CDF), $F(H)$, which gives the probability that a wave height is less than or equal to a value H , is defined by the CWD as (Battjes & Groenendijk, 2000):

$$F(H) = \begin{cases} F_1(H) = 1 - \exp \left[- \left(\frac{H}{H_1} \right)^{k_1} \right] & H \leq H_{tr} \\ F_2(H) = 1 - \exp \left[- \left(\frac{H}{H_2} \right)^{k_2} \right] & H > H_{tr} \end{cases}$$

The parameters of this distribution were determined through extensive analysis of laboratory wave flume data (Battjes & Groenendijk, 2000). The **shape parameters** k_1 and k_2 are fixed constants:

- $k_1 = 2.0$: This exponent confirms that the distribution for the smaller, non-breaking waves ($H \leq H_{tr}$) is modeled as a Rayleigh distribution, which is a special case of the Weibull distribution where the shape parameter is 2 (Battjes & Groenendijk, 2000).
- $k_2 = 3.6$: This empirically derived shape parameter describes the much steeper decline in probability for the larger, depth-limited waves ($H > H_{tr}$). It reflects the physical reality that wave breaking strongly suppresses the occurrence of very high waves (Battjes & Groenendijk, 2000).

The remaining parameters—the scale parameters H_1 and H_2 , and the transitional wave height H_{tr} —are not fixed but are determined by the local physical conditions. A continuity constraint is imposed such that $F_1(\tilde{H}_{tr}) = F_2(\tilde{H}_{tr})$ (Battjes & Groenendijk, 2000).

Because the shape parameters k_1 and k_2 are not equal, the derivative of the CDF, which is the probability density function (PDF), is discontinuous at the transition point H_{tr} . This is acknowledged as a modeling compromise; it is "physically not realistic but it is nevertheless accepted because all integral statistical properties of the wave heights are well behaved" (Battjes & Groenendijk, 2000).

This approach prioritizes the accurate prediction of integrated quantities used in engineering design (such as $H_{1/100}$ or H_{max}) over maintaining a mathematically smooth PDF.

Empirical Parameterization of Physical Variables

The CWD is grounded in physical reality through empirically derived formulas for its key parameters. These relationships, developed from wave flume experiments, connect the statistical model to measurable properties of the nearshore environment (Battjes & Groenendijk, 2000). Those properties are the following:

Free-surface Variance (m_0) The variance of the **free-surface elevation**, m_0 , represents the total energy in the sea state and is calculated from the spectral significant wave height, H_{m0} using the standard definition (Battjes & Groenendijk, 2000), which is valid both deep and shallow waters:

$$m_0 = \left(\frac{H_{m0}}{4} \right)^2$$

Root-Mean-Square Wave Height (H_{rms}) The **root-mean-square wave height**, H_{rms} , is the fundamental scaling parameter for the entire distribution. In deep water, H_{rms} is directly proportional to the standard deviation of the sea surface elevation ($\sqrt{m_0}$). However, in shallow waters, this relationship is modified by nonlinear effects. The empirically derived formula used in the model is (Battjes & Groenendijk, 2000):

$$\frac{H_{rms}}{\sqrt{m_0}} = 2.69 + 3.24 \frac{\sqrt{m_0}}{d}$$

The parameter $\sqrt{m_0}/d$ is a dimensionless measure of the local wave intensity, or degree of saturation. The choice of the constant 2.69 is a deliberate and crucial feature of the model. For a purely linear, narrow-banded sea state, the theoretical relationship is $H_{rms} = \sqrt{8 \cdot m_0}$ (Battjes & Groenendijk, 2000). However, Battjes and Groenendijk (2000), citing field data analysis by Goda (1979), selected 2.69 as the deep-water limit (i.e., as $d \rightarrow \infty$) to better represent real, broad-banded ocean waves. This decision means that even in deep water, the H_{rms} calculated by this model is approximately 5% lower than the theoretical Rayleigh value.

While this approach improves realism for broad-banded seas, it also causes the model's dimensional predictions to diverge from pure Rayleigh theory. This divergence was explicitly analyzed by Caires & Van Gent (2012), who demonstrated how **this parameterization causes the model to make predictions that do not smoothly converge to the Rayleigh values in deep water and can lead to physically inconsistent results, a behavior that necessitates a "capping" logic** described in the computational methodology section.

Transitional Wave Height (H_{tr}) The **transitional wave height**, H_{tr} , represents the physical threshold that separates the two wave populations. It is conceptualized as a limiting height for non-breaking waves, influenced by both the local water depth and the steepness of the beach slope. The primary formula for H_{tr} is (Battjes & Groenendijk, 2000):

$$H_{tr} = (0.35 + 5.8 \cdot \tan(\alpha)) \cdot d$$

where d is the local water depth and $\tan \alpha$ is the beach slope (e.g., for a 1:100 slope, $\tan \alpha = 0.01$). The inclusion of the slope term is physically significant. A steeper slope results in a higher value of H_{tr} , which implies that a smaller fraction of the waves are considered to be in the breaking-dominated regime. This accounts for the spatial lag inherent in the breaking process: on a steep slope, a wave may reach a depth where breaking is initiated but has not yet had sufficient time or distance to fully dissipate its energy and reduce its height (Battjes & Groenendijk, 2000).

Computational Methodology and Numerical Implementation

Algorithmic Workflow

The software follows a structured, multi-step algorithm to compute the wave height distribution from a given set of environmental parameters (Battjes & Groenendijk, 2000).

1. **Input Acquisition:** The program obtains the three required input parameters: spectral significant wave height (H_{m0}), local water depth (d), and beach slope denominator (m for a 1:m slope), either from command-line arguments or from interactive user prompts.
2. **Intermediate Parameter Calculation:** It computes the core physical parameters (m_0 , H_{rms} , H_{tr}) using the empirical formulas detailed in the theoretical foundation section.
3. **Dimensionless Transformation:** The key dimensionless shape parameter, $\tilde{H}_{tr} = H_{tr} / H_{rms}$ is calculated. This single value determines the shape of the entire normalized wave height distribution.
4. **Deep-Water Bypass:** The program evaluates if $\tilde{H}_{tr} > 2.75$. If this condition is met, it signifies that depth-limitation effects are negligible. The program then bypasses the CWD solver and directly uses the well-established theoretical ratios for the Rayleigh distribution.
5. **CWD Solution:** If $\tilde{H}_{tr} \leq 2.75$, the program proceeds to solve the system of non-linear equations derived from the CWD to find the dimensionless scale parameters $H1$ and $H2$. From these, it computes the required statistical wave height ratios, e.g., $H(1/3)$, $H(1/10)$, $H(1/50)$, $H(1/100)$, etc.
6. **Dimensional Conversion:** The calculated dimensionless ratios are multiplied by the dimensional H_{rms} value to obtain the final wave heights in meters.
7. **Physical Consistency Capping:** The final dimensional wave heights are capped at their theoretical Rayleigh limits (e.g., $H_{1/3}$ is capped at H_{m0}). This step ensures the output remains physically plausible and conservative (Caires & Van Gent, 2012).
8. **Report Generation:** All inputs, intermediate values, dimensionless ratios, and final dimensional results are formatted into a comprehensive

report and written to the output file `report.txt`.

Numerical Solution of the CWD Governing Equations

The core numerical task of the software is to determine the dimensionless scale parameters, \tilde{H}_1 and \tilde{H}_2 , for a given value of \tilde{H}_{tr} . These two unknowns are found by solving a system of two coupled, non-linear equations that enforce the mathematical consistency of the CWD (Battjes & Groenendijk, 2000):

1. **Continuity Constraint:** The probability must be continuous at the transitional height H_{tr} . In dimensionless form this becomes (with $k_1 = 2$ and $k_2 = 3.6$):

$$\left(\frac{\tilde{H}_{tr}}{\tilde{H}_1} \right)^{k_1} = \left(\frac{\tilde{H}_{tr}}{\tilde{H}_2} \right)^{k_2}$$

2. **Normalization Constraint:** The mean square of the normalized wave heights (the second moment of the probability density function) must equal one. This is expressed using incomplete gamma functions:

$$\sqrt{\tilde{H}_1^2 \cdot \gamma \left(1 + \frac{2}{k_1}, \left(\frac{\tilde{H}_{tr}}{\tilde{H}_1} \right)^{k_1} \right) + \tilde{H}_2^2 \cdot \Gamma \left(1 + \frac{2}{k_2}, \left(\frac{\tilde{H}_{tr}}{\tilde{H}_2} \right)^{k_2} \right)} = 1$$

where $\gamma(a, x)$ and $\Gamma(a, x)$ are the unnormalized lower and upper incomplete gamma functions, respectively.

The software employs a numerical root-finding algorithm, such as a Newton-Raphson matrix method, to simultaneously solve this system for \tilde{H}_1 and \tilde{H}_2 . Given an initial guess of these two variables $\tilde{H}_1^{(0)}$ and $\tilde{H}_2^{(0)}$, the next iteration is found by solving the linear system:

$$J(\tilde{H}_1^{(i)}, \tilde{H}_2^{(i)}) \begin{pmatrix} \Delta \tilde{H}_1 \\ \Delta \tilde{H}_2 \end{pmatrix} = - \begin{pmatrix} F_1(\tilde{H}_1^{(i)}, \tilde{H}_2^{(i)}) \\ F_2(\tilde{H}_1^{(i)}, \tilde{H}_2^{(i)}) \end{pmatrix}$$

where JJ is the Jacobian matrix of partial derivatives:

$$J = \begin{pmatrix} \frac{\partial F_1}{\partial \tilde{H}_1} & \frac{\partial F_1}{\partial \tilde{H}_2} \\ \frac{\partial F_2}{\partial \tilde{H}_1} & \frac{\partial F_2}{\partial \tilde{H}_2} \end{pmatrix}$$

The solution is updated as:

$$\tilde{H}_1^{(i+1)} = \tilde{H}_1^{(i)} + \Delta\tilde{H}_1 \quad \text{and} \quad \tilde{H}_2^{(i+1)} = \tilde{H}_2^{(i)} + \Delta\tilde{H}_2$$

This process is repeated until the values of F1 and F2 are close to zero.

Once \tilde{H}_1 and \tilde{H}_2 are known, any desired statistical property of the distribution can be calculated (Battjes & Groenendijk, 2000).

Dimensionless Wave-Height Ratios ($\tilde{H}N$ and $\tilde{H}1/N$)

The dimensionless wave-height ratios are critical outputs of the model. The calculation involves solving a system of two non-linear equations derived from the Composite Weibull distribution, ensuring that the normalized H_{rms} of the distribution equals one. This is achieved using a Newton-Raphson matrix method for simultaneous root-finding.

The core of this calculation is finding the values of $\tilde{H}1$ and $\tilde{H}2$ that satisfy simultaneously the normalized H_{rms} equation (Equation 7.11 from Groenendijk, 1998) and the continuity condition between the two Weibull distributions (Equation 3.4). Once \tilde{H}_1 and \tilde{H}_2 (the normalized scale parameters of the first and second Weibull distributions, respectively) are determined, two types of dimensionless wave heights can be calculated:

- **$\tilde{H}N$ (Wave Height with $1/N$ Exceedance Probability):** This is the wave height (H) such that the probability of a wave exceeding it is $1/N$. It is calculated by first determining a candidate $\tilde{H}N$ from the first part of the distribution. If this candidate is less than \tilde{H}_{tr} , then $\tilde{H}N$ is taken from the first part. Otherwise, it is taken from the second part of the distribution.
 - If $\tilde{H}N, candidate < \tilde{H}_{tr}$: $\tilde{H}N = \tilde{H}1 \cdot (\ln(N))^{1/k_1}$
 - If $\tilde{H}N, candidate \geq \tilde{H}_{tr}$: $\tilde{H}N = \tilde{H}2 \cdot (\ln(N))^{1/k_2}$
- **$\tilde{H}_{1/N}$ (Mean of the Highest $1/N$ -part of Wave Heights):** This represents the average height of the highest N -th fraction of waves (e.g., $H_{1/3}$ for significant wave height). The calculation depends on whether \tilde{H}_N (from the previous step) falls within the first or second part of the Composite Weibull distribution.

Case 1: $\tilde{H}N < \tilde{H}_{tr}$ (The wave height with $1/N$ exceedance probability is smaller than the transitional wave height). This scenario implies that the integration for $\tilde{H}_{1/N}$ spans both parts of the Composite Weibull distribution. The formula used is (Groenendijk 1998, Equation A.10):

$$\tilde{H}_{1/N} = N \cdot H_1 \cdot \left[\Gamma \left(\frac{1}{k_1} + 1, \ln(N) \right) - \Gamma \left(\frac{1}{k_1} + 1, \left(\frac{H_{tr}}{H_1} \right)^{k_1} \right) \right] +$$

$$N \cdot H_2 \cdot \Gamma \left(\frac{1}{k_2} + 1, \left(\frac{H_{tr}}{H_2} \right)^{k_2} \right)$$

where $\Gamma(a, x)$ is the unnormalized upper incomplete gamma function.

Case 2: $\tilde{H}N \geq \tilde{H}tr$ (The wave height with $1/N$ exceedance probability is greater than or equal to the transitional wave height). In this case, the integration for $\tilde{H}_{1/N}$ only involves the second part of the Composite Weibull distribution. The formula used is (Groenendijk 1998, Equation A.17):

$$\tilde{H}_{1/N} = N \cdot \tilde{H}_2 \cdot \Gamma \left(\frac{1}{k_2} + 1, \ln(N) \right)$$

Supporting Mathematical Functions

The core calculations rely on precise implementations of fundamental mathematical functions:

- **Complete Gamma Function ($\Gamma(z)$):** This is a generalization of the factorial function to real and complex numbers. In the implementation, `std::tgamma` is used. For calculating the logarithm of the complete gamma function ($\ln(\Gamma(a))$), `std::lgamma` is employed for improved numerical stability, especially for large values of a .

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad (a > 0)$$

- **Unnormalized Lower Incomplete Gamma Function ($\gamma(a, x)$):** This function is computed using a hybrid numerical approach for stability and accuracy. For small values of x (specifically, $x < a + 1.0$), a series expansion is used. For larger values of x , a continued fraction expansion is employed. This adaptive strategy ensures robust and precise computation across different input ranges.

$$\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$$

- **Unnormalized Upper Incomplete Gamma Function ($\Gamma(a, x)$):** This is calculated as $\Gamma(a) - \gamma(a, x)$.

$$\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$$

Consistency Checks and Physical Constraints

To ensure the model's predictions remain physically realistic and consistent with established theory, the software implements a two-tiered "overshoot prevention" logic. This logic is a direct and necessary consequence of the model's empirical formulation, particularly its parameterization of H_{rms} (Battjes & Groenendijk, 2000; Caires & Van Gent, 2012).

The $\tilde{H}_{tr} > 2.75$ Threshold Switch The first safeguard is a check on the dimensionless transitional height, \tilde{H}_{tr} . If this value exceeds 2.75, the program bypasses the CWD solver entirely and defaults to using standard Rayleigh distribution statistics. This is not an arbitrary choice but a computationally efficient shortcut based on the model's documented behavior. The lookup table provided by Battjes and Groenendijk (2000, Table 2) shows the numerical solutions for various statistical wave height ratios as a function of \tilde{H}_{tr} .

An analysis of this table reveals that **for all values of \tilde{H}_{tr} greater than approximately 2.75, the solutions of the CWD converge to and become numerically indistinguishable from the theoretical values of the Rayleigh distribution (e.g., $\tilde{H}_{1/3} \approx 1.416$)** (Caires & Van Gent, 2012). Therefore, this threshold identifies the regime where depth-limitation effects are negligible. By applying the known Rayleigh solution directly, the software avoids unnecessary computation while remaining true to the model's behavior.

Capping of Statistical Parameters The second safeguard is applied after the CWD solution has been found and converted to dimensional wave heights. Each calculated statistical wave height is compared to its theoretical maximum value under the Rayleigh distribution, and capped if it exceeds this limit. For example, the calculated $H_{1/3}$ is not allowed to exceed the input H_{m0} . This capping is necessary to correct for potential inconsistencies arising from the model's specific parameterization of H_{rms} , a behavior analyzed by Caires & Van Gent (2012).

As previously discussed, the model's H_{rms} can differ from the theoretical Rayleigh H_{rms} for the same sea state. This can lead to situations where the CWD predicts a dimensional wave height that is physically implausible (e.g., an average of the top third of waves being larger than the spectrally defined significant wave height). This final check ensures that the software's output remains conservative and physically consistent (Caires & Van Gent, 2012).

Model Applicability, Limitations, and Broader Context

Domain of Validity Based on Calibration Data

The Battjes and Groenendijk model was developed, calibrated, and validated using an extensive set of laboratory wave flume data. The experimental setup consisted of waves propagating over gently sloping, plane beaches with slopes

ranging from 1:20 to 1:250 (Battjes & Groenendijk, 2000). The software's predictions are most reliable and accurate when applied to physical environments that closely match these conditions.

Identified Constraints: Underperformance on Flat Seabeds

While effective within its intended domain, subsequent research has identified important limitations and conditions under which the model's predictions should be used with caution. The most significant limitation identified in the literature is the model's performance on flat or very gently sloping seabeds. **The research of Caires & Van Gent (2012) demonstrated that when the Battjes and Groenendijk model is applied to flat-bottom conditions—an environment outside its original calibration range—it systematically underestimates high wave heights by as much as 15%.**

This underperformance is a direct consequence of the model's physical assumptions. The parameterization of the transitional wave height, H_{tr} , is fundamentally tied to the beach slope, $\tan \alpha$. The slope dictates the rate of depth change, which drives the continuous, gradual energy dissipation process assumed by the model.

On a flat bottom, where $\tan \alpha$ is zero, the model predicts a low H_{tr} and consequently a heavily truncated wave height distribution, implying intense wave breaking. However, the physical reality is different; on a flat bottom, the ratio of significant wave height to depth can be higher, and the breaking process is more intermittent rather than continuous (Caires & Van Gent, 2012).

The model's core assumption about the dissipation mechanism does not hold, leading to a predictable and non-conservative underestimation of extreme wave heights. This has critical implications for engineering design on shallow continental shelves or over large shoals, where the risk of extreme waves may be significantly higher than the model predicts (Caires & Van Gent, 2012).

Comparative Performance with Other Models

More recent, extensive field data analyses have placed the Battjes and Groenendijk's model within the broader context of other shallow-water wave height distributions. Studies have shown that the model provides a good fit for the most severe, depth-limited sea states (typically where the relative wave height $H_s/d > 0.45$) on sloping beaches (Karpadakis et al., 2020).

However, for moderately energetic conditions or different bathymetries, other models may provide more accurate results. **Users should be aware that no single model has been found to be universally superior across all shallow-water conditions, and the choice of model may depend on the specific characteristics of the sea state being analyzed (Karpadakis et al., 2020).**

The **Rayleigh's 1885 model**, is based on the assumption of a narrow-banded, linear Gaussian sea surface. It performs optimally in deep water conditions where there is no interaction with the seabed. However, this model has significant limitations, as it fails completely in shallow water environments. Its applicability is restricted to deep waters, and it is known to be unbounded, as highlighted by Longuet-Higgins in 1952.

The **Glukhovskiy's 1966 model** relies on the Weibull distribution with a depth-dependent shape parameter to characterize wave heights. It is best suited for moderately energetic shallow-water conditions, where the wave dynamics are influenced by the seabed. Nonetheless, it tends to overestimate extreme wave heights, a limitation documented by van Vledder in 1991.

Battjes and Groenendijk's 2000 model employs a composite Weibull distribution, representing two populations, and is referred to as a "Point model." It is most effective in scenarios involving severely depth-limited sea states, specifically when the ratio of significant wave height to water depth exceeds 0.45, on plane or sloping beaches with slopes ranging from 1:20 to 1:250. However, this model tends to underestimate wave heights on flat bottoms, which can lead to non-conservative predictions. Additionally, it neglects the influence of wave history, as discussed by Caires and Van Gent in 2012.

The **Mendez and colleagues' 2004 model** is based on the concept of wave energy propagation with breaking. It serves as an alternative approach for conditions characterized by moderate energy levels. Despite its usefulness, it has limitations, including a bounded shape parameter for low Iribarren numbers, as described by Mendez et al. in 2004.

Finally, the **Karmpadakis et al.' 2022 model** is founded on a continuous distribution that incorporates nonlinearity and wave directionality. It is particularly suitable for intermediate and shallow waters over flat or horizontal seabeds. Nonetheless, this model was developed for short-crested seas and may require upscaling from laboratory scales to real-world applications, which could introduce additional uncertainties.

Building and Running

This guide details how to compile and use the various programs in the shallow water wave model suite.

Command-Line Interface (CLI)

The CLI application calculates shallow-foreshore wave-height distribution parameters. It's available in both C++ and Fortran.

C++ Version (shallow-water-waves_cli.cpp) Compilation Instructions (g++):

```
g++ -O3 -march=native -std=c++17 -Wall -Wextra -pedantic -Wconversion \
-Wsign-conversion -static -static-libgcc -static-libstdc++ -o \
shallow-water-waves_cli shallow-water-waves_cli.cpp
```

Usage: You can run the application by providing parameters as command-line arguments or by entering them interactively.

- **With command-line arguments (e.g., $H_{m0}=2.5$, $d=5$, $\text{slope}M=100$):**

```
./shallow-water-waves_cli 2.5 5 100
```

- **Interactive input:**

```
./shallow-water-waves_cli
```

The program will then prompt you for the required values.

Fortran Version (shallow-water-waves_cli.f90) Compilation Instructions (gfortran):

```
gfortran -O3 -march=native -std=f2018 -Wall -Wextra -pedantic \
-fno-underscoring -o shallow-water-waves_cli_f shallow-water-waves_cli.f90
```

Usage: The Fortran version runs identically to the C++ version.

- **With command-line arguments (e.g., $H_{m0}=2.5$, $d=5$, $\text{slope}M=100$):**

```
./shallow-water-waves_cli_f 2.5 5 100
```

- **Interactive input:**

```
./shallow-water-waves_cli_f
```

Graphical User Interface (GUI)

The GUI application provides a native Windows interface for the same calculations.

Compilation Instructions (g++ for Windows):

```
g++ -O3 -march=native -std=c++17 -Wall -Wextra -pedantic \
-Wconversion -Wsign-conversion -municode shallow-water-waves_gui.cpp -o \
shallow-water-waves_gui -mwindows -static -static-libgcc -static-libstdc++
```

Usage: Run the compiled executable (shallow-water-waves_gui.exe). A window will appear where you can input the H_{m0} , d , and Beach slope m values in the text fields and click "Compute" to see the results. The report is also saved to report.txt.

Jupyter Notebook Interface (shallow-water-waves.ipynb)

This notebook provides a detailed implementation and explanation of the wave height distribution model. It is structured to be both an interactive computational tool and an educational document.

Prerequisites Before running the notebook, ensure you have the required Python libraries installed by running the following command in your terminal or command prompt:

```
pip install numpy scipy pandas notebook matplotlib
```

Usage

1. **Follow the Notebook Sequentially:** Each section builds upon the last, explaining the theory and then providing the code that implements it.
2. **Run the Code Cells:** Execute the code cells in order as you progress through the notebook.
3. **Modify Inputs:** In the final section, change the input parameters to match your scenario and run the last cell to see the results.

Carvalho (2025) Table Generator

This program computes and tabulates normalized wave height parameters over a range of conditions (H_{tr}/H_{rms}), generating the file `carvalho2025_table.txt`. It does not require user input.

C++ Version (carvalho2025_table.cpp) Compilation Instructions (g++):

```
g++ -O3 -march=native -std=c++17 -Wall -Wextra -pedantic -static \
-o carvalho2025_table carvalho2025_table.cpp
```

Usage:

```
./carvalho2025_table
```

Fortran Version (carvalho2025_table.f90) Compilation Instructions (gfortran):

```
gfortran -O3 -march=native -std=f2008 -Wall -Wextra -pedantic \
-fno-underscoring -o carvalho2025_table_f carvalho2025_table.f90
```

Usage:

```
./carvalho2025_table_f
```

Software Usage and Output Interpretation

Required Input Parameters

The software requires three input parameters to define the local environmental conditions. The first parameter is the Significant Wave Height, denoted as H_{m0} , which represents the spectral significant wave height and is defined as $4\sqrt{m_0}$. Its units are meters (m). The second parameter is the Water Depth, symbolized by (d), indicating the local still water depth at the point of interest, also measured in meters (m). The third parameter is the Beach Slope, represented as (m), which is the denominator of the beach slope expressed as a ratio 1:(m), and is dimensionless.

Execution via Command-Line and Graphical Interfaces

The software can be run from the command line or through a graphical user interface.

Command-Line Interface (CLI) The CLI application (`shallow-water-waves_cli`) can be executed in two modes:

- **Argument-based execution:** Provide the three input parameters as command-line arguments in the order H_{m0} , d , m .
`./shallow-water-waves_cli 2.0 5.0 100`
- **Interactive execution:** Run the executable without arguments. The program will then prompt the user to enter each value interactively.
`./shallow-water-waves_cli`

Graphical User Interface (GUI) The GUI application (`shallow-water-waves_gui`) provides a user-friendly window with input fields for the three parameters. After entering the values, clicking the "Calculate" button will perform the computation and display the results directly in the interface.

Analysis of the Generated Report File

Both the CLI and GUI applications generate a detailed text file named `report.txt` in the execution directory. This file contains a comprehensive summary of the calculation.

The report is organized into several sections. The **Inputs** section lists the parameters H_{m0} , d , and the slope, echoing the user-provided values for verification. The **Intermediate Values** section presents key physical and dimensionless parameters such as H_{m0} , H_{rms} , H_{tr} , and \tilde{H}_{tr} , with \tilde{H}_{tr} being the most critical value that determines the shape of the wave height distribution.

The **Dimensionless Ratios** section includes ratios like $\tilde{H}_{1/3}$, $\tilde{H}_{1/10}$, etc., representing the normalized shape of the wave height distribution. The **Final**

Wave Heights section provides primary dimensional outputs such as $H1/3$, $H1/10$, etc., measured in meters for engineering applications.

The **Diagnostic Ratios** section features ratios which are used to compare the model output against theoretical Rayleigh values.

References

- **Battjes, J. A., & Groenendijk, H. W. (2000).** Wave height distributions on shallow foreshores. *Coastal Engineering*, 40(3), 161-182. URL: https://data-ww3.ifremer.fr/BIB/Battjes_Groenendijk_CE2000.pdf
- **Caires, S., & Van Gent, M. R. A. (2012).** Wave height distribution in constant and finite depths. *Coastal Engineering Proceedings*, 1(33), 15. URL: <https://doi.org/10.9753/icce.v33.waves.15>
- **Goda, Y. (1979).** A review on statistical interpretation of wave data. *Report of the Port and Harbour Research Institute, Japan*, 18(1), 5-32. URL: <https://www.pari.go.jp/PDF/vol018-no01-01.pdf>
- **Groenendijk, H. W. (1998).** *Shallow foreshore wave height statistics*. M.Sc.-thesis, Delft University of Technology, Department of Civil Engineering, Section Fluid Mechanics, The Netherlands. URL: <http://resolver.tudelft.nl/uuid:fe03dda9-40d9-4046-87fb-459f01fcd3d3>
- **Groenendijk, H. W., & Van Gent, M. R. A. (1998).** *Shallow foreshore wave height statistics; A predictive model for the probability of exceedance of wave heights*. Technical Report H3351, WL | delft hydraulics, The Netherlands. URL: <http://dx.doi.org/10.13140/RG.2.2.14180.68486>
- **Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. D. E., Enke, K.,... & Walden, H. (1973).** *Measurements of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP)*. *Ergänzungsheft zur Deutschen Hydrographischen Zeitschrift Reihe, A* (8), 95. URL: https://www.researchgate.net/publication/256197895_Measurements_of_wind-wave_growth_and_swell_decay_during_the_Joint_North_Sea_Wave_Project_JONSWAP
- **James, J. P., & Panchang, V. (2022).** Investigation of Wave Height Distributions and Characteristic Wave Periods in Coastal Environments. *Journal of Geophysical Research: Oceans*, 127, e2021JC018144. URL: <https://doi.org/10.1029/2021JC018144>
- **Karmpadakis, I., Swan, C., & Christou, M. (2020).** Assessment of wave height distributions using an extensive field database. *Coastal Engineering*, 157, 103630. URL: <https://doi.org/10.1016/j.coastaleng.2019.103630>
- **Karmpadakis, I., Swan, C., & Christou, M. (2022).** A new wave height distribution for intermediate and shallow water depths. *Coastal Engineering*, 175, 104130. URL: <https://doi.org/10.1016/j.coastaleng.2022.104130>
- **Klopman, G. (1996).** *Extreme wave heights in shallow water*. WL | delft hydraulics, Report H2486, The Netherlands.

- **Klopman, G., & Stive, M. J. F. (1989).** *Extreme waves and wave loading in shallow water*. Paper presented at the E&P Forum Workshop in Paris, Delft Hydraulics, The Netherlands. URL: <https://repository.tudelft.nl/record/uuid:84ef8201-1130-418f-954a-79a80c626173>
- **Longuet-Higgins, M. S. (1952).** On the statistical distribution of heights of sea waves. *Journal of Marine Research*, 11(3), 245-266. URL: https://elischolar.library.yale.edu/journal_of_marine_research/774/
- **Longuet-Higgins, M. S. (1980).** On the distribution of the heights of sea waves: Some effects of nonlinearity and finite band width. *Journal of Geophysical Research*, 85(C3), 1519-1523. URL: <https://doi.org/10.1029/JC085iC03p01519>
- **Naess, A. (1985).** On the distribution of crest to trough wave heights. *Ocean Engineering*, 12(3), 221-234. URL: [https://doi.org/10.1016/0029-8018\(85\)90014-9](https://doi.org/10.1016/0029-8018(85)90014-9)
- **Tayfun, M. A. (1990).** Distribution of large wave heights. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 116(6), 686-707. URL: [https://doi.org/10.1061/\(ASCE\)0733-950X\(1990\)116:6\(686\)](https://doi.org/10.1061/(ASCE)0733-950X(1990)116:6(686))
- **Thornton, E. B., & Guza, R. T. (1982).** Energy saturation and phase speeds measured on a natural beach. *Journal of Geophysical Research*, 87(C12), 9499-9508. URL: <https://doi.org/10.1029/JC087iC12p09499>
- **Thornton, E. B., & Guza, R. T. (1983).** Transformation of wave height distribution. *Journal of Geophysical Research*, 88(C10), 5925-5938. URL: <https://doi.org/10.1029/JC088iC10p05925>