

Determinación del Período de Retorno de Oleaje Extremo Mediante Análisis de Salidas de un Modelo Numérico Espectral

Determination of the Return Period of Extreme Wave Height Using Output Analysis of a Spectral Numerical Model

Rosales-Grano, P.¹✉

Datos de Adscripción:

¹✉ Rosales Grano, Pedro
Tecnológico Nacional de México/Instituto Tecnológico de Guaymas. Guaymas,
Sonora, México.
pedro_rg@guaymas.tecnm.mx
ORCID: <https://orcid.org/0009-0006-7911-0780>

Resumen - Con el objetivo de estimar alturas significantes extremas y sus periodos de retorno asociados, se analizó una serie temporal de largo periodo de altura significativa de olas, generada por el modelo espectral WaveWatch III en la costa oriental del Golfo de California. Debido a la falta de mediciones directas extensas en la región de estudio, es que se extrajo la información de un nodo numérico del modelo de oleaje. Se compararon las distribuciones de probabilidad de Rayleigh, Gumbel y Weibull mediante dos enfoques: máximos anuales y el método Peaks Over Threshold (POT), utilizando un umbral de 1.8 m y criterios de independencia temporal entre eventos. La calidad de ajuste de las funciones de distribución se evaluó mediante la prueba de Anderson-Darling y diagramas P-P y Q-Q. Los resultados mostraron que las alturas de ola menores a 1 m predominan durante la mayor parte del año; sin embargo, se identificaron eventos extremos asociados principalmente a huracanes y tormentas tropicales. La distribución de Gumbel presentó el mejor desempeño estadístico, mostrando el menor valor del estadístico de Anderson-Darling y el mejor ajuste visual en los diagramas probabilísticos. Asimismo, el método POT permitió capturar una mayor cantidad de eventos extremos relevantes, incrementando las alturas estimadas para periodos de retorno largos. Se concluye que la selección del método estadístico y de la distribución probabilística influye significativamente en la estimación de oleaje extremo y, por tanto, en el diseño y evaluación de infraestructura costera.

Palabras Clave: Excedencias sobre un umbral, Ingeniería costera, Oleaje extremo, Período de retorno, WaveWatch III.

Abstract - To estimate extreme significant wave heights and their associated return periods, a long-term time series of significant wave height generated by the WaveWatch III spectral model on the eastern coast of the Gulf of California was analyzed. Due to the lack of extensive direct measurements in the study area, data was extracted from a numerical node of the wave model. Rayleigh, Gumbel, and Weibull probability distributions were compared using two

approaches: annual maxima and the Peaks Over Threshold (POT) method, considering a threshold of 1.8 m and temporal independence criteria between events. Goodness-of-fit was evaluated using the Anderson-Darling test together with P-P and Q-Q probability plots. Results indicate that wave heights below 1 m dominate during most of the year; however, extreme events associated with hurricanes and tropical storms were identified. The Gumbel distribution provided the best statistical performance, exhibiting the lowest Anderson-Darling statistic and the best visual agreement in probability plots. Furthermore, the POT method captured a larger number of relevant extreme events, increasing the estimated wave heights for long return periods. It is concluded that both the statistical methodology and the selected probability distribution strongly influence extreme wave estimations and therefore play a critical role in the design and risk assessment of coastal infrastructure.

Keywords: Coastal engineering, Extreme waves, Peak Over Threshold, Return period, WaveWatch III.

I. INTRODUCCIÓN

Las olas generadas por el viento constituyen uno de los agentes dominantes en la ingeniería de costas, al representar la principal fuente de energía y de cargas dinámicas que actúan sobre las estructuras marítimas. Un diseño que no incorpore adecuadamente parámetros fundamentales como la altura, el periodo y la dirección del oleaje puede derivar en fallos estructurales o en una operación ineficiente.

El conocimiento de la dinámica del oleaje es, por tanto, esencial no solo para garantizar la estabilidad estructural, sino también para reducir procesos de erosión y asegurar condiciones adecuadas para la navegación y las actividades portuarias. Estas consideraciones influyen directamente en el diseño integral de las obras, desde la selección de materiales hasta la definición de su geometría.

En este contexto, la estimación de alturas de ola extremas y sus correspondientes periodos de retorno es un componente clave en la ingeniería costera y oceánica. Dichas estimaciones se sustentan en el análisis estadístico de series temporales de oleaje, que permiten modelar su comportamiento probabilístico bajo supuestos de estacionariedad y representatividad.

Asimismo, los eventos extremos de oleaje pueden inducir incrementos en el nivel del mar debido al wave setup y al run-up, lo que incrementa el riesgo de inundaciones por desbordamiento

en zonas costeras. Por ello, el diseño y la gestión de infraestructura marítima requieren estudios de oleaje rigurosos que consideren tanto condiciones medias como extremas.

El objetivo principal de este trabajo es estimar las alturas de ola extremas y sus periodos de retorno en una región de la costa oriental del Golfo de California. Ante la ausencia de registros instrumentales de largo plazo en el sitio de análisis, se emplearon datos provenientes de las salidas del modelo espectral de oleaje, WaveWatch III (Tolman, 1991), correspondientes a un nodo numérico cercano al área de estudio (**Figura 5**). A partir de la serie de tiempo de altura significativa (H_s), y mediante el ajuste de funciones de distribución de probabilidad, se estimaron las alturas de ola extremas y sus periodos de retorno para distintos horizontes temporales.

La determinación de las alturas de ola extremas y su período de retorno (T_r) ha sido estudiada por diversos autores. Martínez-Díaz De León y Coria-Méndez (1993) analizaron registros de altura significativa del oleaje en la Bahía de Todos Santos, México, durante el periodo 1986–1989. Para caracterizar probabilísticamente el oleaje, los autores ajustaron los datos a distribuciones teóricas comúnmente utilizadas en ingeniería costera, incluyendo Fisher–Tippett tipo I, Weibull y Rayleigh. El mejor ajuste se obtuvo con la distribución de Fisher–Tippett tipo I, la cual permitió estimar alturas de ola asociadas a periodos de retorno de 50 y 100 años de aproximadamente 3.41 m y 3.76 m, respectivamente. El trabajo destaca la importancia de seleccionar adecuadamente la distribución de probabilidad para representar los extremos del oleaje, así como el uso de técnicas de ajuste y extrapolación para la estimación de eventos raros, elementos clave en el diseño y evaluación de infraestructura marítima.

En un estudio realizado por Li et al. (2012) se compararon tres metodologías para la estimación de oleaje extremo utilizando un registro de 17 años. En el trabajo se evaluó la eficacia y precisión del método convencional de Goda (Goda, 1988), frente a los métodos de Valor Extremo Generalizado (GEV) y la Distribución de Pareto Generalizada (GPD). Se analizaron registros horarios de altura de ola significativa. El análisis incluyó un enfoque de Peak Over Threshold (POT) que utiliza las distribuciones de Weibull y Gumbel. El modelo GEV se aplicó a los máximos anuales de altura de ola. El modelo GPD se ajustó a los excesos sobre un umbral determinado mediante técnicas de selección de umbral, como el gráfico de vida residual media (Coles et al., 2001). El modelo GEV proporcionó predicciones de altura de ola significativamente más bajas para periodos de retorno superiores a 10 años en comparación con los otros dos métodos. Los autores no aconsejan su uso para este conjunto de datos por ser "poco eficiente en el uso de los datos" (al usar solo máximos anuales) y presentar mayores niveles de incertidumbre. Las predicciones de los métodos de Goda y GPD fueron notablemente similares cuando se seleccionó un umbral adecuado. Las pruebas de bondad de ajuste de Chi-cuadrado favorecieron ligeramente a la distribución GPD para este conjunto de datos específico. La selección del umbral es crítica; un umbral demasiado alto genera sesgos por el tamaño reducido de la muestra, mientras que uno demasiado bajo viola la base asintótica del modelo estadístico. Se recomienda a la comunidad de ingeniería costera adoptar la distribución GPD por sus

ventajas en las técnicas de selección de umbral y su sólido respaldo teórico, aunque el método de Goda sigue siendo un candidato válido si se aplica con un criterio riguroso de selección de umbral y pruebas de bondad de ajuste.

En un estudio reciente (Zhang et al., 2026) se desarrollaron metodologías para la evaluación regional de amenazas inducidas por tifones, viento, oleaje y corrientes aplicables al diseño de estructuras marinas. Acoplaron el modelo de oleaje para aguas someras e intermedias SWAN (Booij et al., 1999; Ris et al., 1999) y el modelo hidrodinámico tridimensional ADCIRC (Luettich y Westerink, 2004) para analizar la sensibilidad de los extremos de cuantiles altos frente a distintos conjuntos de datos de vientos de reanálisis y distribuciones de valores extremos. Encontraron que la elección del conjunto de datos de reanálisis tiene un impacto limitado en los extremos de diseño; sin embargo, la selección del modelo de valores extremos influye de manera significativa en las estimaciones. Los mapas de riesgo para periodos de retorno de 50 años, utilizando distribuciones como la GEV o la GPD y la distribución óptima basada en el criterio de Kolmogorov–Smirnov (Massey Jr., 1951), resultaron, con incertidumbres menores al 10%. En contraste, la distribución de Gumbel mostró desviaciones de hasta el 20%.

En el contexto de la evaluación de riesgos costeros, diversos estudios han destacado la importancia del oleaje extremo generado por ciclones tropicales como principal forzante de la erosión en sistemas playa-duna. En este sentido, Al Azad et al. (2026) analizaron la erosión costera en el noroeste de Florida mediante un enfoque novedoso que incluyó análisis numéricos y probabilísticos, centrado en la determinación del oleaje extremo a partir de ciclones tropicales sintéticos. Estos eventos fueron generados a partir de bases de datos de reanálisis atmosférico, lo que permitió construir un conjunto amplio de condiciones extremas que superan los registros históricos disponibles. A partir de dichos ciclones sintéticos, los autores derivaron parámetros de oleaje (altura significativa, periodo y dirección), los cuales se emplearon como forzamiento en el modelo morfodinámico XBeach (Roelvink et al., 2010) para simular la respuesta del sistema costero. Este enfoque permitió no solo cuantificar volúmenes de erosión, sino también estimar periodos de retorno asociados a eventos extremos, proporcionando una medida probabilística del peligro. Asimismo, el uso de múltiples fuentes de reanálisis evidenció la sensibilidad en la estimación del oleaje extremo, lo que resalta la necesidad de enfoques tipo ensamble para capturar la incertidumbre inherente. En conjunto, este tipo de metodologías integradas, basadas en ciclones sintéticos, reanálisis atmosférico y modelación numérica de alta resolución, representan un avance en la caracterización del oleaje extremo y su impacto en la dinámica costera.

II. PARTE TÉCNICA DEL ARTÍCULO

Para caracterizar los eventos más severos y determinar valores extremos de las olas se utilizan distribuciones de probabilidad. En este contexto, las funciones de densidad de probabilidad (FDP) permiten describir la frecuencia relativa de ocurrencia de distintas alturas de ola y, a partir de ello, estimar niveles asociados a periodos de retorno específicos.

La distribución de Rayleigh ha sido ampliamente empleada para modelar alturas de ola individuales en condiciones de mar

completamente desarrollado y procesos lineales. Su principal ventaja radica en su simplicidad y en que se deriva de supuestos físicos asociados a la superposición de ondas lineales. Sin embargo, presenta limitaciones importantes, ya que no captura adecuadamente condiciones no lineales ni estados de mar complejos, lo que puede conducir a sobreestimaciones o subestimaciones de eventos extremos (Longuet-Higgins, 1952). Por otro lado, la distribución de Gumbel, perteneciente a la familia de distribuciones de valores extremos tipo I, es comúnmente utilizada para modelar máximos anuales o picos de series de tiempo. Su fortaleza reside en su capacidad para representar eventos extremos sin necesidad de conocer la distribución subyacente completa de los datos. No obstante, su uso implica asumir independencia entre eventos extremos y puede no ajustarse adecuadamente en presencia de colas pesadas o datos con alta variabilidad (Gumbel, 1958).

La distribución de Weibull ofrece una mayor flexibilidad debido a la presencia de parámetros que controlan la forma, escala y ubicación de la distribución, lo que permite ajustarla a una amplia variedad de condiciones de oleaje. Esta versatilidad constituye su principal ventaja, ya que puede representar tanto comportamientos moderados como extremos. Sin embargo, su ajuste requiere procedimientos de estimación más complejos y puede ser sensible a la calidad y tamaño de la muestra disponible (Weibull, 1951).

En la práctica, la selección de la distribución más adecuada depende de la naturaleza de los datos, el objetivo del análisis y las condiciones físicas del entorno marino. Es común realizar comparaciones entre diferentes modelos probabilísticos para evaluar su desempeño mediante pruebas de bondad de ajuste, con el fin de reducir la incertidumbre en la estimación de alturas de ola asociadas a períodos de retorno específicos. Este proceso resulta esencial para garantizar diseños seguros y eficientes en ingeniería marítima.

2.1 Distribución de Rayleigh

La distribución de Rayleigh es el modelo probabilístico estándar para describir la variabilidad de las alturas de olas individuales en un estado de mar estacionario a corto plazo, asumiendo un espectro de oleaje de banda estrecha. Matemáticamente, su Función de Densidad de Probabilidad (FDP) se expresa como:

$$f(H) = \frac{2H}{H_{rms}^2} \exp \left[-\left(\frac{H}{H_{rms}} \right)^2 \right] \quad (1)$$

donde: H es la altura de las olas individuales y H_{rms} es la raíz cuadrática media de un número N de olas (H_i), que se calcula mediante:

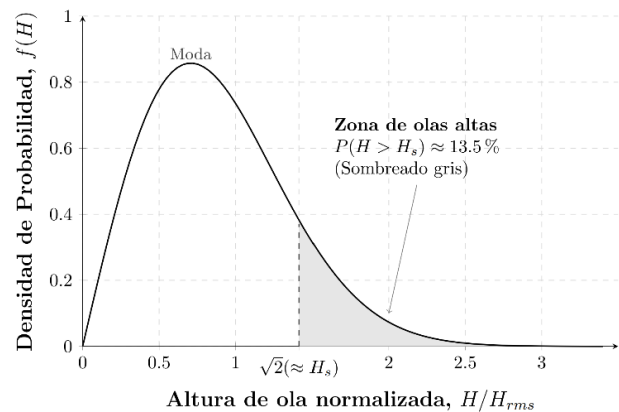
$$H_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N H_i^2} \quad (2)$$

En la **Figura 1** se muestra una gráfica de la distribución; la curva presenta una asimetría; su moda (valor más frecuente) se ubica en $H = H_{rms}/\sqrt{2}$. El área sombreada bajo la cola derecha de la curva representa la probabilidad de excedencia de un umbral específico. En ingeniería de costas este umbral es la altura

significante H_s que es la que se considera como altura de ola de diseño para la mayoría de las obras marítimas. Al evaluar el área correspondiente a $H > H_s$ ($H_s = \sqrt{2}H_{rms}$), es decir, al integrar la FDP desde H_s hasta el infinito el área resultante es $e^{-2} \approx 0.135$. Estadísticamente, esto significa que el 13.5% de las olas individuales de un registro superarán la altura de ola significativa.

Figura 1

Función de distribución de probabilidad de Rayleigh para alturas de olas.



2.2 Distribución de Gumbel

A diferencia del enfoque de Rayleigh, la distribución de Gumbel (Distribución de Valores Extremos Tipo I) se utiliza para el análisis probabilístico a largo plazo. Su propósito principal es modelar la distribución de los valores máximos extraídos de bloques temporales de datos, como las alturas de ola máximas anuales. La FDP de Gumbel está definida por la siguiente ecuación:

$$f(H) = \frac{1}{\beta} \exp \left[-\left(\frac{H - \mu}{\beta} \right) - \exp \left(-\frac{H - \mu}{\beta} \right) \right] \quad (3)$$

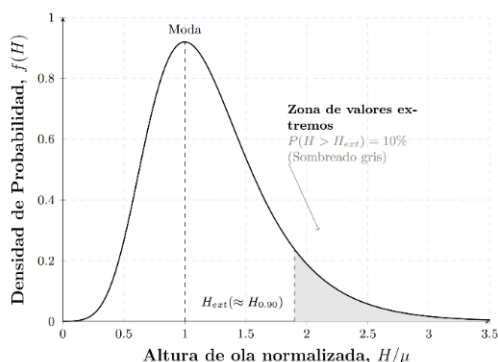
donde μ es el parámetro de localización (que define el desplazamiento de la curva y corresponde a su moda) y $\beta > 0$ es el parámetro de escala (que regula la dispersión de los valores extremos).

La **Figura 2** muestra la FDP de Gumbel, la cual exhibe un decaimiento de doble exponencial. Su característica más importante es la asimetría pronunciada hacia la derecha (una cola derecha extendida). Esta zona es fundamental para la ingeniería, ya que modela la probabilidad de ocurrencia de eventos severos y atípicos. El área sombreada en el extremo derecho de la gráfica representa la probabilidad de excedencia anual, la cual está estrictamente vinculada al período de retorno (T_r) mediante la relación $P(H > H_{ext}) = 1/T_r$.

Esta integración permite determinar las alturas de ola de diseño para garantizar la supervivencia de las estructuras bajo condiciones extremas.

Figura 2

Función de distribución de probabilidad de Gumbel para alturas de olas extremas.



2.3 Distribución de Weibull

La FDP de Weibull se utiliza en el análisis de variables ambientales debido a su gran flexibilidad para modelar diferentes formas de distribución, desde comportamientos altamente sesgados hasta aproximaciones cercanas a la exponencial o normal, dependiendo de sus parámetros de forma y escala. En el contexto del oleaje, esta versatilidad permite representar adecuadamente la variabilidad de alturas significantes de ola, especialmente en el rango de valores extremos. Una de sus principales ventajas es que puede ajustarse con precisión tanto a colas pesadas como ligeras, lo cual resulta importante para estimar niveles de retorno con mayor confiabilidad. Además, su formulación facilita la extrapolación hacia eventos poco frecuentes, como alturas de ola asociadas a periodos de retorno largos (por ejemplo, 50 o 100 años), lo que la convierte en una herramienta robusta para el diseño de estructuras costeras y la evaluación de riesgos. Sin embargo, su desempeño depende de una correcta estimación de parámetros y de la calidad del conjunto de datos.

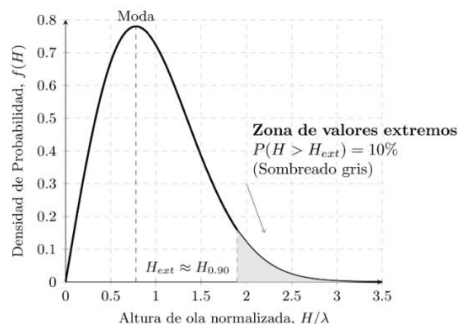
La FDP de Weibull se define como:

$$f(H) = \frac{k}{\lambda} \left(\frac{H}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{H}{\lambda}\right)^k\right] \quad (4)$$

donde k es un parámetro de forma, que determina el comportamiento de la distribución, λ es un parámetro de escala que indica la dispersión de los datos tiene unidades de longitud.

Figura 3

Función de distribución de probabilidad de Weibull para alturas de olas extremas.



La **Figura 3** muestra la FDP de Weibull, la relación $H_{ext} \approx H_{0.90}$ corresponde a una interpretación probabilística basada en percentiles. El valor $H_{0.90}$ representa el percentil 90 de la distribución, es decir, aquel valor de altura de ola tal que el 90% de las observaciones son menores o iguales a éste. De forma equivalente, se cumple que:

$$P(H \leq H_{0.90}) = 0.90 \Rightarrow P(H > H_{0.90}) = 0.10 \quad (5)$$

por lo tanto, al definir $H_{ext} \approx H_{0.90}$, se está estableciendo un umbral de valores extremos tal que únicamente el 10% de las olas más altas superan dicho valor.

El área sombreada en la cola derecha de la distribución (**Figura 3**) representa precisamente esta probabilidad de excedencia. Matemáticamente, corresponde al área bajo la función de densidad de probabilidad desde H_{ext} hasta infinito:

$$\int_{H_{ext}}^{\infty} f(H) dH = P(H > H_{ext}) = 0.10 \quad (6)$$

2.4 Funciones de distribución acumulativas

La estimación de periodos de retorno se basa en la modelación probabilística de las alturas de ola significantes máximas mediante funciones de distribución acumulativa (FDA), las cuales permiten describir la probabilidad de que una variable aleatoria no exceda un determinado valor. Entre las distribuciones más empleadas destacan Rayleigh, Gumbel y Weibull, cada una con fundamentos teóricos y dominios de aplicación específicos.

La distribución de Rayleigh se utiliza comúnmente para describir la estadística del oleaje bajo condiciones de mar completamente desarrollado y procesos lineales, donde las componentes espectrales de la superficie libre son independientes y gaussianas. Su FDA proporciona una aproximación adecuada para alturas de ola individuales o significantes en estados de mar regulares; su capacidad para representar eventos extremos es limitada, ya que tiende a subestimar la probabilidad de ocurrencia de valores muy altos.

Por otro lado, la distribución de Gumbel, derivada de la teoría de valores extremos, se enfoca específicamente en modelar máximos anuales o eventos extremos independientes. Su FDA es particularmente adecuada cuando se emplea el enfoque de máximos por bloque, donde se selecciona el valor máximo de oleaje en intervalos de tiempo definidos (por ejemplo, anuales). Esta distribución permite una extrapolación robusta hacia periodos de retorno largos, siendo utilizada debido a su simplicidad y fundamento estadístico.

En contraste, la distribución de Weibull ofrece una mayor flexibilidad gracias a su parámetro de forma, lo que le permite adaptarse a diferentes regímenes de datos, incluyendo tanto valores medios como extremos. Su FDA es especialmente útil en el enfoque de excedencias sobre un umbral, donde se consideran múltiples eventos extremos que superan un cierto nivel crítico. Esta capacidad de ajuste le permite caracterizar el comportamiento del oleaje en una amplia gama de condiciones, desde estados moderados hasta eventos severos.

En conjunto, estas funciones de distribución acumulativa (**Tabla 1**) son herramientas útiles para la estimación de periodos de retorno del oleaje extremo. La elección adecuada entre ellas depende del tipo de datos disponibles, el método de análisis empleado y el nivel de precisión requerido en la extrapolación de eventos raros.

Tabla 1
Funciones de distribución acumulativa.

Funciones de distribución acumulativas (FDA)		
Rayleigh	$F(H) = 1 - \exp\left(-\frac{H^2}{2H_{rms}^2}\right)$	(7)
Gumbel	$F(H) = \exp\left[-\exp\left(-\frac{H-\mu}{\beta}\right)\right]$	(8)
Weibull	$F(H) = 1 - \exp\left[-\left(\frac{H}{\lambda}\right)^k\right]$	(9)

2.5 Determinación del periodo de retorno y altura de ola asociada a partir de la función de distribución acumulativa.

Sea $F(H)$ la función de distribución acumulativa (FDA) de la altura de ola H , definida como la probabilidad de no excedencia: $F(H) = P(X \leq H)$; donde X representa la variable aleatoria que corresponde a la altura significativa. La probabilidad de excedencia está dada por: $P(X > H) = 1 - F(H)$.

El periodo de retorno T_r se define como el inverso de la probabilidad de excedencia por evento:

$$T_r = \frac{1}{P(X > H)} = \frac{1}{1 - F(H)} \quad (10)$$

cuando se trabaja con series temporales con resolución Δt , es necesario introducir la tasa de ocurrencia n (número de datos por año):

$$n = \frac{(365)(24)}{\Delta t}$$

por ejemplo, para un intervalo de muestreo de $\Delta t = 6$ horas, se tiene $n = 1460$ eventos/año.

Por lo que el periodo de retorno se expresa como (Massel, 2017):

$$T_r = \frac{1}{n(1 - F(H))} \quad (11)$$

La altura de ola asociada a un periodo de retorno dado $H(T_r)$ se obtiene invirtiendo las FDA:

$$F(H) = 1 - \frac{1}{nT_r} \quad (12)$$

$$H(T_r) = F^{-1}\left(1 - \frac{1}{nT_r}\right) \quad (13)$$

Las expresiones específicas para determinar las alturas significativas en función del periodo de retorno se muestran en la **Tabla 2** para las distribuciones de Rayleigh, Gumbel y Weibull.

Tabla 2

Expresiones para determinar la altura de la ola (H) en función del periodo de retorno (T_r).

$$\text{Rayleigh} \quad H(T_r) = H_{rms} \sqrt{-2 \ln\left(1 - \frac{1}{nT_r}\right)} \quad (14)$$

$$\text{Gumbel} \quad H(T_r) = \mu - \beta \ln\left[-\ln\left(1 - \frac{1}{nT_r}\right)\right] \quad (15)$$

$$\text{Weibull} \quad H(T_r) = \alpha \left[-\ln\left(1 - \frac{1}{nT_r}\right)\right]^{1/k} \quad (16)$$

Con el propósito de verificar las ecuaciones y los códigos implementados en computadora, se utilizó la serie de datos hipotética de alturas significativas anuales que se muestra en la **Tabla 3**.

Tabla 3

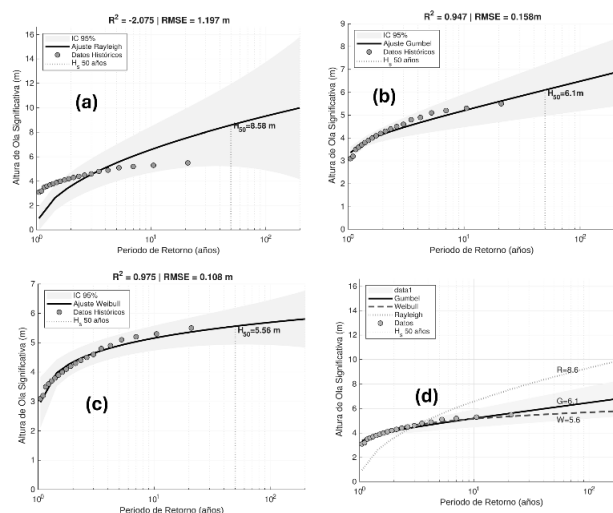
Máximos de alturas significativas H_s anuales en metros, utilizados como prueba de los códigos computacionales (caso hipotético).

3.2	4.5	3.8	5.1	4.2	3.5	4.8	5.5	3.9	4.1
4.3	3.7	5.2	4.6	3.1	4.4	4.9	5.3	3.6	4.0

En la **figura 4** se muestran las gráficas de los periodos de retorno determinados en base a las ecuaciones de la **Tabla 2**, mientras más estrecha sea el área sombreada se tiene mayor certeza en la estimación. En este caso hipotético, la distribución de Weibull presenta la zona sombreada más estrecha, lo que sugiere que es esta la distribución adecuada para predecir las alturas significativas durante eventos extremos y sus periodos de retorno asociados. La distribución que presenta el área sombreada más amplia es la distribución de Rayleigh. Se observa cómo los datos están dispersos alrededor de la curva teórica.

Figura 4

Periodo de retorno determinado en base a las ecuaciones que se muestran en la **Tabla 2**; los paneles (a), (b) y (c) corresponden a las FDA de Rayleigh, Gumbel y Weibull respectivamente, el panel (d) muestra las FDA y los datos utilizados en el cálculo.



III. RESULTADOS Y DISCUSIÓN

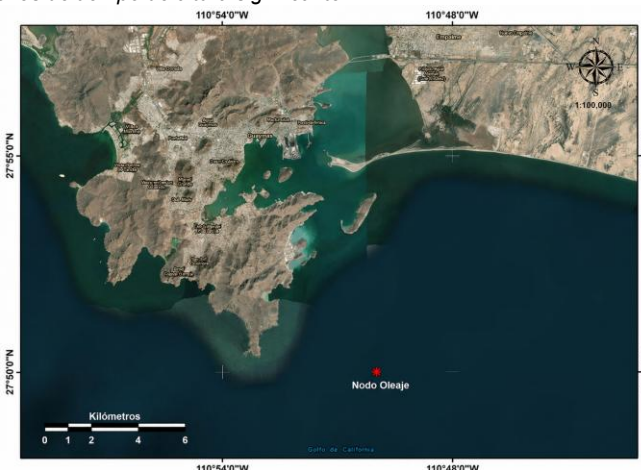
Para determinar el período de retorno y la altura de ola significativa asociada durante eventos meteorológicos extremos, se utilizaron datos del modelo numérico espectral de oleaje WaveWatch III (Tolman, 1991) disponibles en el portal del Instituto Francés de Investigación para la Explotación del Mar (Ifremer, s/f). La base de datos abarca de 1990 a 2018, con excepción de los años 1992 y 1993, en los cuales no se cuenta con información. El formato original de los archivos fue NetCDF, y su procesamiento se realizó mediante rutinas desarrolladas en MATLAB. Los registros de altura significativa se extrajeron de un nodo numérico localizado en la costa oriental del Golfo de California, en las coordenadas geográficas 27°50'0.22" N y 110°49'59.53" O. La ubicación del nodo analizado se ubica en las cercanías de la bahía de Guaymas, en la costa de Sonora, México, como se muestra en la **Figura 5**.

La resolución espacial en latitud y longitud del modelo numérico fue de 0.166 grados que equivale aproximadamente a 18.5 km, mientras que la resolución temporal fue de 3 horas, este intervalo de tiempo es por lo general el que se utiliza cuando se programa instrumentación oceanográfica como sensores de presión o boyas con acelerómetros para el registro del fenómeno del oleaje, y corresponde a la variabilidad que pueden experimentar los trenes de olas generados por vientos locales o durante tormentas.

Es pertinente mencionar que una limitación importante es que el análisis se realizó utilizando exclusivamente información proveniente del modelo espectral WaveWatch III, pues no existen registros instrumentales de largo plazo para efectuar una validación local de las alturas significantes simuladas. Aunque WaveWatch III ha sido ampliamente validado a escala global y regional y constituye una de las principales fuentes de información para estudios de oleaje extremo, las estimaciones obtenidas pueden presentar incertidumbres derivadas de la resolución espacial del modelo, de los campos de viento utilizados como forzamiento y de la representación de procesos físicos como la disipación por rompimiento, interacciones no lineales y generación por viento. En zonas costeras semicerradas, como el Golfo de California, estas incertidumbres pueden incrementarse debido a la influencia de la batimetría, la geometría del golfo y la limitada resolución espacial del modelo global, factores que no siempre son adecuadamente representados. En consecuencia, las alturas extremas estimadas deben interpretarse como aproximaciones consistentes con el comportamiento regional del oleaje, más que como una representación exacta de las condiciones locales. No obstante, el empleo de bases de datos de reanálisis constituye actualmente una práctica ampliamente aceptada cuando no existen registros instrumentales suficientemente largos para realizar análisis de valores extremos, debido a que proporcionan series temporales continuas de varias décadas indispensables para la estimación de periodos de retorno de 50 o 100 años. En este contexto, los resultados obtenidos deben considerarse una aproximación robusta para la caracterización probabilística del oleaje extremo en la región de estudio.

Figura 5

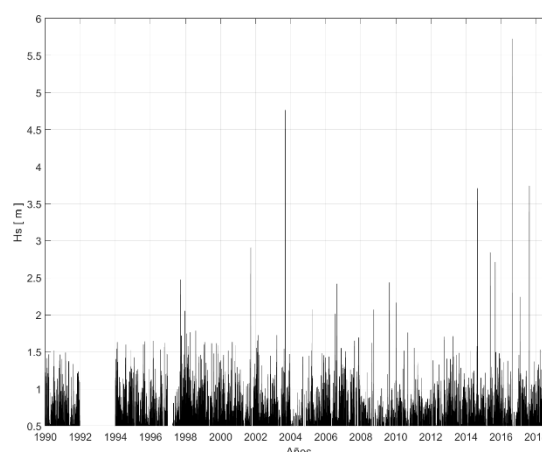
Nodo numérico de oleaje analizado, se localiza en las coordenadas geográficas 27°50'0.22"N 110°49'59.53"O, de donde se extrajeron las series de tiempo de altura significativa.



La **Figura 6** muestra la serie de tiempo analizada de las salidas del modelo WaveWatch III (Tolman, 1991), se observa la presencia de eventos extremos asociados a huracanes, en especial durante los años 2016 a 2018. De la serie de tiempo se extrajeron los máximos de altura significativa por año, con esta información se realizó el análisis del periodo de retorno. Adicionalmente, para realizar un análisis complementario se estableció un umbral de altura significativa de 1.8 m y un intervalo mínimo de 48 horas entre registros, con el fin de asegurar la independencia de los eventos. En consecuencia, únicamente se consideraron aquellas observaciones de $H_s > 1.8$ que estuvieran separadas al menos por 48 horas.

Figura 6

Alturas significantes extraídas del nodo numérico de oleaje analizado



La **Tabla 4** muestra las alturas significantes para periodos de retorno de 10, 25, 50 y 100 años calculados mediante las distribuciones de probabilidad, considerando máximos anuales y un intervalo de confianza IC al 95%, estos periodos de retorno

son lo que se consideran por lo general para el diseño de obras marítimas. En la **Figura 7** se muestran las alturas significantes en función de los periodos de retorno, en la figura, mientras más estrecha sea la zona sombreada se tiene más certeza en la estimación, caso contrario una zona sombreada amplia presenta mayor incertidumbre, especialmente en valores extremos con periodos de retorno largos, se observa que la distribución de Rayleigh, panel a) de **Figura 7** muestra la zona sombreada más amplia para periodos de retorno largos. Este resultado es importante pues permite evaluar el riesgo de subestimar o sobre estimar alturas de ola extremas. En la **Figura 7** se observa que para el caso donde se consideran máximos anuales la distribución de Gumbel (**Figura 7** panel b) presenta la zona sombreada más estrecha, lo que sugiere que esta distribución es la que mejor representa la variabilidad de los datos. De acuerdo con los datos presentados en la **Tabla 4**, la distribución de Gumbel mostró una altura significativa mayor a las otras dos para periodos de retorno 50 y 100 años. Para periodos de retorno menores a 25 años, los datos históricos están cercanos a las curvas teóricas y las bandas sombreadas son estrechas para las 3 distribuciones analizadas.

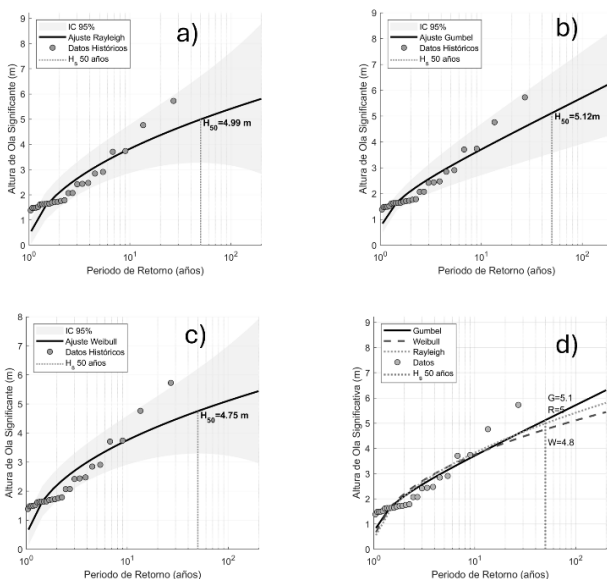
Tabla 4

Periodos de retorno (T_r) en años y alturas significantes (H_s) en metros para diferentes FDP, obtenidos de los máximos anuales.

	Rayleigh	Gumbel	Weibull
10	3.83	3.71	3.75
25	4.53	4.52	4.36
50	4.99	5.12	4.75
100	5.42	5.72	5.11

Figura 7

Período de retorno determinado en base a las ecuaciones que se muestran en la **Tabla 2**. Los paneles a) b) y c) corresponden a las distribuciones de Rayleigh, Gumbel y Weibull, el panel d) muestra las 3 distribuciones. El cálculo se realizó considerando los máximos anuales como datos de entrada.



En un análisis complementario se utilizó el método POT (Peak Over Threshold), descrito por Coles et al. (2001). En este contexto, la selección del umbral constituye un aspecto esencial, dado que determina el equilibrio entre sesgo y varianza en la estimación. La elección de un umbral demasiado bajo conduce a la inclusión de observaciones que no pertenecen al régimen asintótico de valores extremos, lo que introduce sesgo en los resultados. Por el contrario, un umbral excesivamente alto limita el número de excedencias disponibles y, en consecuencia, incrementa la incertidumbre asociada a la estimación de los parámetros. En este trabajo se evaluaron distintos valores de umbral mediante el análisis del número de excedencias independientes y de la estabilidad de los parámetros de ajuste. Se observó que valores inferiores a 1.8 m producían un incremento importante en el número de eventos, incorporando estados de mar moderados que no corresponden a condiciones extremas. Por el contrario, umbrales superiores reducían considerablemente el tamaño de la muestra, incrementando la incertidumbre estadística de las distribuciones ajustadas. El valor de 1.8 m representa un balance adecuado entre ambos criterios, ya que conserva un número suficiente de eventos independientes para el análisis estadístico y, simultáneamente, selecciona únicamente condiciones de oleaje asociadas a tormentas tropicales y huracanes en el Golfo de California.

Se seleccionaron los datos de entrada a las funciones de densidad de probabilidad considerando una altura de ola significativa umbral de 1.8 m; es decir, se incluyeron todos los registros de H_s superiores a 1.8 m sin importar el año de ocurrencia. Para garantizar la independencia de los eventos, se estableció que los datos no provinieran de un mismo huracán o tormenta, aplicando un intervalo mínimo de 48 horas entre sucesos consecutivos. El ajuste de las distribuciones Rayleigh, Weibull y Gumbel fue evaluado mediante la prueba de bondad de ajuste de Anderson-Darling (Anderson y Darling, 1954) y mediante inspección gráfica con diagramas de probabilidad acumulada P-P y diagramas de cuantiles comparados Q-Q (Gan y Koehler, 1990). La prueba de Anderson-Darling determina dos valores (A^2 y p) que miden la discrepancia entre la distribución empírica de los datos y la distribución teórica que se está evaluando. A^2 se calcula integrando la diferencia entre la función de distribución acumulada empírica y la teórica, con mayor peso en las colas. Un valor pequeño de A^2 indica que las curvas empírica y teórica son muy similares, el modelo se ajusta bien a los datos. Un valor grande de A^2 refleja mayores discrepancias por lo que el modelo se ajusta mal. A partir de A^2 se obtiene el parámetro p . Cuanto más grande sea A^2 , más pequeño será el valor de p , aumentando la probabilidad de rechazar la hipótesis nula (que el modelo es adecuado). Si A^2 es pequeño, el valor de p tiende a ser grande, lo que significa que no hay evidencia suficiente para rechazar el modelo. Se estableció un valor de significancia de 0.05 como umbral de decisión. Si el valor de p es menor a 0.05 se rechaza la hipótesis nula, es decir el modelo propuesto no se ajusta bien a los datos, caso contrario si el valor de p es mayor a 0.05 no existe evidencia estadística suficiente para rechazar ninguna de ellas como modelo plausible para los datos analizados.

Los resultados de la prueba Anderson-Darling indicaron que las tres distribuciones presentan valores de significancia superiores

a 0.05 (**Tabla 5**). Sin embargo, la distribución Gumbel presentó el menor valor del estadístico A^2 y el mayor valor de p , lo que sugiere un mejor desempeño relativo frente a las distribuciones Rayleigh y Weibull.

Tabla 5
Prueba de Anderson-Darling

Distribución	A^2	p
Rayleigh	1.566	0.159
Weibull	1.268	0.235
Gumbel	1.059	0.332

Este resultado es consistente con el análisis anterior donde se consideró un evento por año. Sin embargo, los resultados de la prueba de bondad de ajuste de Anderson-Darling son sensibles a el valor del umbral seleccionado. En este caso la determinación del umbral para la elección de los datos de entrada a las FDP se estableció en base a las condiciones físicas del sitio de estudio en este caso el Golfo de California, una ola de 1.80 m solo se presenta durante eventos de tormenta como huracanes, tormentas o depresiones tropicales, el Golfo de California es un mar semicerrado de aproximadamente 1000 km de longitud por 150 km de ancho, por lo que la generación de olas se ve limitada por las características geográficas del mismo, aun cuando el viento sea de considerable magnitud y sople durante un intervalo largo de tiempo.

En las gráficas tipo P-P (paneles superiores) y Q-Q (paneles inferiores) de la **Figura 8**, el mejor ajuste se determina cuando los puntos se acercan a la línea diagonal. Visualmente, Gumbel presenta una alineación más consistente, especialmente frente a Rayleigh, donde se observa mayor curvatura y desviación sistemática. Weibull muestra un ajuste intermedio, pero con desviaciones más marcadas en algunos cuantiles.

Figura 8
Gráficas de probabilidad acumulada (P-P) paneles superiores y diagramas de cuantiles (Q-Q) paneles inferiores, para la 3 FDP analizadas.

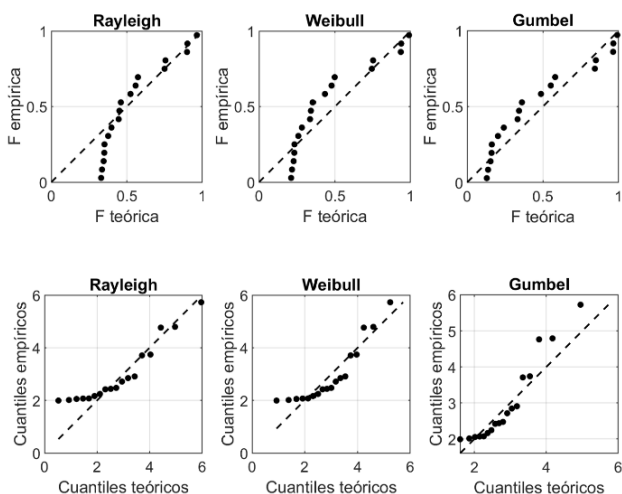


Tabla 6
Períodos de retorno (T_r) en años y alturas significantes (H_s) en metros para diferentes FDP, con un umbral de 1.8 m.

T_r /FDP	Rayleigh	Gumbel	Weibull
10	4.35	3.75	4.18
25	5.29	4.41	4.81
50	5.9	4.90	5.20
100	6.4	5.38	5.54

La **Tabla 6** muestra los periodos de retorno para 10, 25, 50 y 100 años, así como sus alturas significantes asociadas para las tres FDP analizadas, se observa que la distribución de Rayleigh presenta los mayores valores de altura significativa lo que hace suponer que está sobreestimando el valor de H_s mientras que la distribución de Gumbel calcula valores menores que parecen ser más coherentes. Este resultado se ve respaldado por los gráficos de probabilidad. En los diagramas P-P (**Figura 8**), la distribución Gumbel muestra una mayor proximidad de los puntos respecto a la línea de referencia, indicando una mejor concordancia entre las probabilidades empíricas y teóricas. En los diagramas Q-Q (**Figura 8**), aunque se observan desviaciones en los valores extremos, la distribución Gumbel reproduce de manera más adecuada la tendencia general de los cuantiles observados. En contraste, Rayleigh evidencia una desviación sistemática más pronunciada, particularmente en la región central y superior de los datos, mientras que Weibull presenta un comportamiento intermedio.

La presencia de algunas desviaciones en los extremos puede atribuirse al tamaño de la muestra, a la variabilidad natural de los datos o a la influencia de valores extremos. No obstante, dado que la distribución Gumbel está asociada al modelado de valores extremos, su mejor desempeño resulta coherente cuando el fenómeno analizado presenta máximos, eventos extremos o valores de alta magnitud.

IV. CONCLUSIONES

Con base en la prueba de bondad de ajuste de Anderson-Darling y en la evaluación gráfica mediante diagramas P-P y Q-Q, se concluye que la distribución Gumbel proporciona el mejor ajuste relativo a los datos analizados. Aunque Rayleigh y Weibull no son rechazadas estadísticamente, Gumbel presenta el menor estadístico A^2 , el mayor valor del parámetro p y una mejor alineación visual con la línea teórica de referencia. Por tanto, se recomienda emplear la distribución Gumbel como modelo probabilístico principal para describir el comportamiento de los datos en este estudio.

La comparación entre la **Tabla 4** y la **Tabla 6** evidencia que la estimación de alturas significantes extremas depende de manera importante del conjunto de datos utilizado y del criterio de selección aplicado. En la **Tabla 4**, obtenida a partir de máximos anuales, los valores estimados para los distintos periodos de retorno presentan una variación moderada entre las funciones de densidad de probabilidad (FDP), destacando que la distribución de Gumbel proporciona las mayores alturas para periodos de

retorno altos (50 y 100 años), mientras que Weibull tiende a generar las menores estimaciones. Por otro lado, en la **Tabla 6**, elaborada considerando un umbral de 1.8 m, se observa un incremento generalizado de las alturas extremas estimadas, especialmente para las distribuciones Rayleigh y Weibull. Este comportamiento sugiere que el método POT incorpora un mayor número de eventos extremos relevantes respecto al enfoque de máximos anuales, capturando con mayor sensibilidad la variabilidad del oleaje severo. Asimismo, las diferencias entre las FDP se amplifican para periodos de retorno largos, lo que pone de manifiesto la incertidumbre inherente a la extrapolación estadística de eventos raros y la necesidad de seleccionar cuidadosamente tanto la metodología de muestreo como la distribución probabilística más adecuada para estudios de diseño y riesgo costero.

La determinación de oleaje extremo mediante distribuciones de densidad de probabilidad es un análisis de equilibrio entre la teoría matemática y el conocimiento físico de la zona de estudio. La elección del umbral en el método POT (Coles et al., 2001) o la suposición de estacionariedad en las series de tiempo es una decisión que requiere un criterio experto. En el contexto actual de cambio climático, donde la intensidad y frecuencia de los ciclones tropicales y extra tropicales muestran variaciones significativas, la estadística de valores extremos debe evolucionar hacia modelos no estacionarios que consideren las tendencias de largo plazo, garantizando así que la infraestructura costera del futuro sea resiliente ante un océano cada vez más dinámico.

Las estimaciones obtenidas dependen de la calidad de la base de datos utilizada. La incorporación futura de mediciones instrumentales permitirá validar localmente las alturas significantes simuladas por el modelo WaveWatch III y reducir la incertidumbre en la estimación de periodos de retorno.

V. AGRADECIMIENTOS

Al Tecnológico Nacional de México y al Instituto Tecnológico de Guaymas, por proporcionar los medios para el desarrollo de esta investigación.

VI. REFERENCIAS

- Al Azad, A. A., Marsooli, R., Jamous, M., Appendini, C. M., Ruiz Salcines, P., y Duran, R. (2026). Coastal erosion hazards in northwest Florida due to extreme wind waves generated by tropical cyclones. *Coastal Engineering*, 203, 104872. <https://doi.org/10.1016/J.COASTALENG.2025.104872>
- Anderson, T. W., y Darling, D. A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268), 765–769. <https://doi.org/10.1080/01621459.1954.10501232>
- Booij, N., Ris, R. C., y Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions: 1. Model description and validation. *Journal of geophysical research: Oceans*, 104(C4), 7649–7666.

- Coles, S., Bawa, J., Trenner, L., y Dorazio, P. (2001). An introduction to statistical modeling of extreme values (Vol. 208). Springer.
- Gan, F. F., y Koehler, K. J. (1990). Goodness-of-Fit Tests Based on P-P Probability Plots. *Technometrics*, 32(3), 289–303. <https://doi.org/10.1080/00401706.1990.10484692>
- Goda, Y. (1988). On the methodology of selecting design wave height. En *Coastal Engineering 1988* (pp. 899–913).
- Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press.
- Ifremer. (s/f). WAVEWATCH III – wind waves in the Earth system. Recuperado <https://iowaga.ifremer.fr/Products/WAVEWATCH-III>
- Li, F., Bicknell, C., Lowry, R., y Li, Y. (2012). A comparison of extreme wave analysis methods with 1994–2010 offshore Perth dataset. *Coastal Engineering*, 69, 1–11. <https://doi.org/10.1016/j.coastaleng.2012.05.006>
- Longuet-Higgins, M. S. (1952). On the statistical distribution of the heights of sea waves. *Journal of Marine Research*, 11, 245–266.
- Luetlich, R. A., y Westerink, J. J. (2004). Formulation and numerical implementation of the 2D/3D ADCIRC finite element model version 44. XX (Vol. 20). R. Luetlich Chapel Hill, NC, USA.
- Martinez-Diaz De Leon, A., y Coria-Mendez, P. (1993). Probability distribution of wave height in Todos Santos Bay, B.C., Mexico. *Ciencias Marinas*, 19(2), 203–218. <https://doi.org/10.7773/cm.v19i2.928>
- Massel, S. R. (2017). *Ocean surface waves: their physics and prediction* (Vol. 45). World scientific.
- Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253), 68–78.
- Ris, R. C., Holthuijsen, L. H., y Booij, N. (1999). A third-generation wave model for coastal regions: 2. Verification. *Journal of Geophysical Research: Oceans*, 104(C4), 7667–7681.
- Roelvink, D., Reniers, A., Van Dongeren, A. P., de Vries, J., Lescinski, J., y McCall, R. (2010). XBeach model description and manual. Unesco-IHE Institute for Water Education, Deltares and Delft University of Technology. Report June, 21(2010), 2.

- Tolman, H. L. (1991). A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. *Journal of Physical Oceanography*, 21(6), 782–797.
- Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, 18(3), 293–297.
- Zhang, Y. B., Cheng, H., Sheng, C., y Dai, K. S. (2026). Regional assessment of typhoon wind, wave, and current multi-hazards for offshore structure design using SWAN + ADCIRC: Sensitivity to reanalysis wind forcings and extreme value distributions. *Weather and Climate Extremes*, 52. <https://doi.org/10.1016/j.wace.2026.100899>

No.	Rol	Descripción
Nota 1: favor de eliminar los corchetes una vez colocado su nombre Nota 2: elimine el todo texto de apoyo que se proporcionó en los asteriscos de cada fila y debajo de cada nombre.		
1	Conceptualización [Conceptualization]	Pedro Rosales Grano
2	Curación de datos [Data curation]	Pedro Rosales Grano
3	Análisis formal [Formal analysis]	Pedro Rosales Grano
4	Obtención de fondos [Funding acquisition]	Pedro Rosales Grano
5	Investigación [Investigation]	Pedro Rosales Grano
6	Metodología [Methodology]	Pedro Rosales Grano
7	Administración del proyecto [Project administration]	Pedro Rosales Grano
8	Recursos [Resources]	Pedro Rosales Grano
9	Software [Software]	Pedro Rosales Grano
10	Supervisión [Supervision]	Pedro Rosales Grano
11	Validación [Validation]	Pedro Rosales Grano
12	Visualización [Visualization]	Pedro Rosales Grano
13	Redacción - borrador original [Writing – original draft]	Pedro Rosales Grano
14	Redacción - revisión y edición [Writing – review & editing]	Pedro Rosales Grano