

Selección del modelo extremal para la altura de ola

JAVIER MARTINEZ ARANZABAL (*)

RESUMEN. Ciertos métodos de análisis extremal del oleaje, como es el de los máximos anuales, requieren la obtención de la distribución límite de la variable altura de ola. La importancia de la elección de esta distribución límite se hace tanto mayor cuanto más elevados sean los periodos de retorno cuyas alturas de ola asociadas se quieran determinar.

Basándose en el hecho de que en el límite cualquier función de distribución pertenece al dominio de atracción de las conocidas distribuciones asintóticas (Gumbel, Fréchet y Weibull), y teniendo en cuenta ciertas características de las mismas así como las limitaciones físicas propias del fenómeno, se van a exponer una serie de conclusiones encaminadas a que, dentro del inevitable grado de incertidumbre propio de este tipo de problemas, la distribución límite seleccionada sea la más adecuada para cada caso particular.

SELECTION OF MODEL FOR EXTREME WAVE HEIGHTS

ABSTRACT. Certain methods employed in extreme wave analysis, i.e. Annual Maxima, require determining the limit distribution of the wave heights. The choice of the limit distribution becomes more significant as the return period of the corresponding wave height increases.

This paper presents some approaches to improve the choice of the distribution limit for specific cases. These approaches are based on the fact that any distribution function tends asymptotically to the domain of attraction of one of the well-known three asymptotic distributions (Gumbel, Fréchet and Weibull). They also take into account the characteristics of these distributions as well as the physical limitations of the phenomenon.

1. INTRODUCCION

La altura de ola se encuentra entre las variables físicas cuya función de distribución original es desconocida. Esta desagradable hipótesis hace que aquellos estudios o proyectos de ingeniería marítima que requieran la obtención del régimen de temporales, es decir, la distribución extremal de la altura de ola, se enfrenten habitualmente al problema de elegir una función de distribución que sea capaz de predecir, con la mayor fiabilidad posible, el comportamiento del fenómeno cuando éste se extrapola a periodos de retorno suficientemente elevados.

Dado que no hay una justificación teórica concluyente para decantarse por un modelo en detrimento de otro, es habitual seleccionar la función de distribución extremal basándose en ciertos criterios de buena práctica, siendo uno de los más frecuentes el probar varias distribuciones, entre las comúnmente aceptadas para el oleaje, y optar por una u otra en base a la bondad del ajuste a los datos de partida. Como se observará en algunos ejemplos, los valores de retorno que, para una misma muestra, proporcionan distintas distribuciones pueden presentar diferencias considerables.

2. PLANTEAMIENTO Y OBJETIVOS

Los distintos métodos utilizados para obtener el régimen extremal de oleaje pueden clasificarse, en una primera aproximación, en dos categorías: los que emplean la distribución límite de la variable (el de los máximos anuales es uno de ellos) y aquellos que estiman la distribución original y la extrapolan a los periodos de retorno solicitados (el de la distribución inicial pertenece a este grupo). Aquí se va a prestar atención a los primeros, sin entrar en otras consideraciones que las relativas a la elección de la función de distribución.

En este artículo no se pretende establecer una norma definitiva sobre qué modelo debe emplearse en cada caso, sino más bien exponer una serie de consideraciones teóricas y prácticas que sirvan de ayuda a la hora de abordar un problema de oleaje extremal, aún a sabiendas de que, inevitablemente, siempre existirá un cierto grado de incertidumbre.

3. DOMINIOS DE ATRACCION DE LAS FUNCIONES ASINTOTICAS

La obtención del régimen extremal de oleaje, a efectos de diseño, es un problema de máximos en el cual se desconoce la función de distribución original, debiéndose por tanto estimar la distribución límite a partir de datos muestrales.

(*) Licenciado en Ciencias Físicas. Director de Programa del CEPYC-CEDEX.

Sin embargo, el conocimiento de esa distribución límite puede simplificarse partiendo del hecho importante de que en el límite cualquier función de distribución pertenece al dominio de atracción de una de las tres distribuciones que se indican a continuación (Fisher-Tippett, 1928).

Distribución de Gumbel (Fisher-Tippett tipo I o Asintota I)

Se deduce a partir de distribuciones de la variable ilimitadas en la cola superior, lo cual tiene una forma exponencial que se puede escribir como:

$$F(x) = 1 - e^{-g(x)} \quad [1]$$

donde $g(x)$ es una función creciente de x .

La función de distribución tiene la forma siguiente:

$$F(x) = e^{-e^{-\frac{x-A}{B}}} \quad , -\infty < x < \infty, \quad B > 0 \quad [2]$$

Distribución de Fréchet (Fisher-Tippett tipo II o Asintota II)

Esta asintota surge a partir de distribuciones ilimitadas en la cola de interés, que tiende asintóticamente a la forma

$$F(x) = 1 - B \left(\frac{1}{x} \right)^C \quad , x \geq 0 \quad [3]$$

estando limitadas en el otro extremo por el valor cero (Gumbel propuso una expresión más general incluyendo un límite inferior para la variable).

La cola de la distribución de Fréchet es más larga que la de Gumbel, lo que hace que los valores de retorno correspondientes a sus extrapolaciones sean más elevados.

La función de distribución se puede expresar como:

$$F(x) = e - \left(\frac{B}{x} \right)^C \quad , x \geq 0 \quad [4]$$

Distribución de Weibull (Fisher-Tippett tipo III o Asintota III)

Se deduce a partir de distribuciones limitadas en la cola de interés. Cuando el interés se centra en los valores máximos, como es el presente caso, existe un límite superior A en cuyas proximidades las funciones de distribución pertenecientes al dominio de atracción de Weibull tienen la forma:

$$F(x) = 1 - B(A - x)^C \quad , x \leq A, \quad C > 0 \quad [5]$$

La función de distribución de Weibull se puede escribir entonces del siguiente modo:

$$F(x) = e - \left(\frac{A - x}{B} \right)^C \quad , x \leq A \quad [6]$$

Las asintotas II y III están relacionadas con la distribución de Gumbel mediante transformaciones logarítmicas.

Para determinar el dominio de atracción de una distribución cualquiera existen varios métodos entre los que se encuentran los siguientes: Piekands (1975), Galambos (1980), el método de la curvatura (Castillo y Galambos, 1986) y el del papel probabilístico (Castillo, 1988).

Aquí se va a prestar atención a este último, ya que por su sencillez permite determinar el dominio de atracción a partir de una muestra, sin necesidad de conocer la función de distribución original, lo que hace que sea de gran utilidad para el caso de la altura de ola.

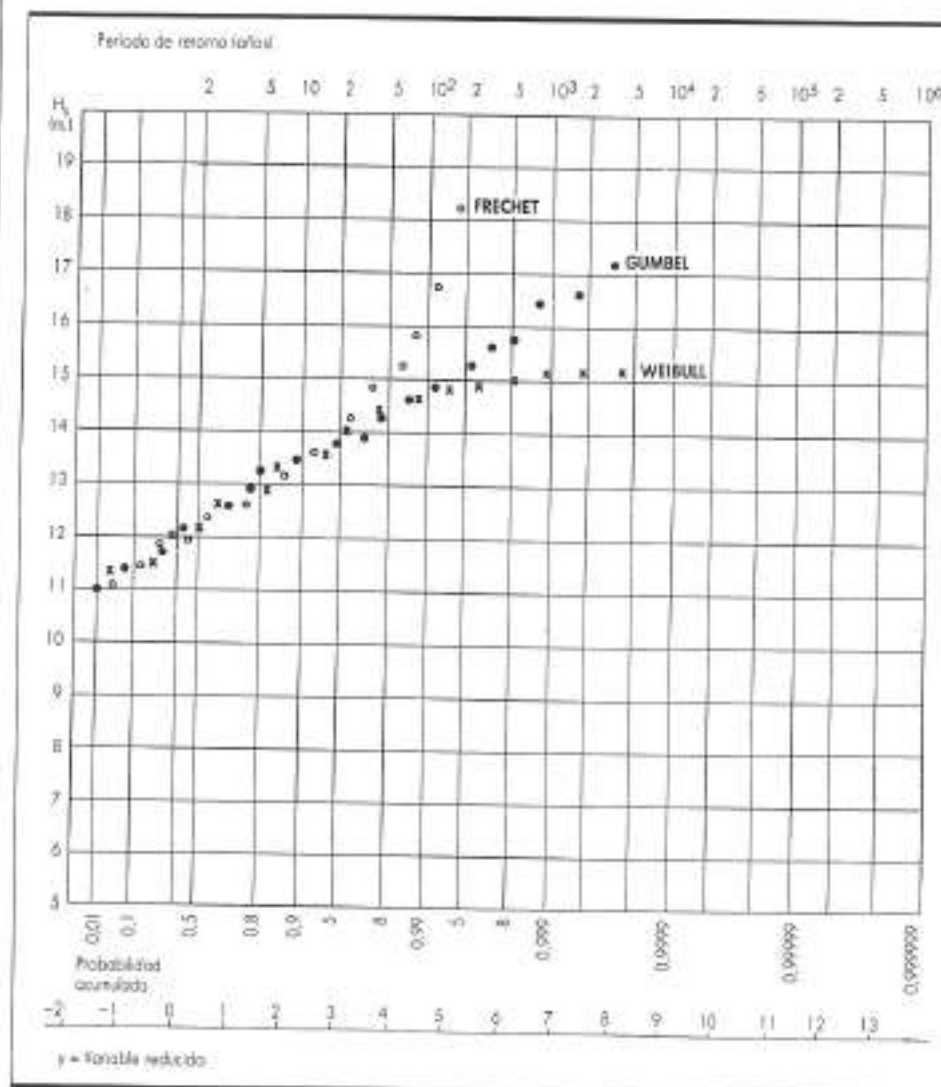
A grandes rasgos el método consiste en representar los datos muestrales en papel probabilístico de Gumbel, se recomienda el uso de la fórmula de punteo dada por Gringorten (1963) y observar la curvatura de la cola superior de la nube de puntos representada. Las colas superiores de distribuciones pertenecientes al dominio de atracción de Weibull presentarán la concavidad hacia abajo, en concordancia con la existencia de un límite superior de la variable. Las que pertenezcan a Fréchet, de acuerdo con el hecho de admitir valores infinitos de la variable, tendrán la concavidad hacia arriba. Por último, las colas superiores de aquellas distribuciones pertenecientes al dominio de atracción de Gumbel mostrarán una tendencia rectilínea. En la figura 1 se esquematiza el método que se acaba de exponer. A continuación (tabla 1) se indican los dominios de atracción, para máximos, de algunas de las distribuciones más comúnmente empleadas (Castillo y Sarabia, 1992).

DISTRIBUCION	DOMINIO DE ATRACCION
NORMAL	GUMBEL
LOGNORMAL	GUMBEL
EXPONENCIAL	GUMBEL
RAYLEIGH	GUMBEL
GUMBEL _M	GUMBEL
GUMBEL _m	GUMBEL
PARETO	FRECHET
CAUCHY	FRECHET
WEIBULL _M	WEIBULL
WEIBULL _m	GUMBEL
UNIFORME	WEIBULL
FRECHET _M	FRECHET
FRECHET _m	GUMBEL

TABLA 1. Dominios de atracción para máximos.

M₀ Para máximos.
m₀ Para mínimos.

FIGURA 1. Caracterización gráfica de los dominios de atracción para máximos.



4. AJUSTES MUESTRALES

En este apartado se van a exponer una serie de casos prácticos con el propósito de resaltar las diferencias que pueden surgir según se seleccione una u otra distribución límite. De esta forma se tratan de clarificar las principales peculiaridades de cada una de estas distribuciones, que a fin de cuentas serán las que decidan la elección de un modelo determinado. En la tabla 2 aparecen conjuntamente las muestras correspondientes a los distintos casos que van a ser objeto de análisis. Por alejarse del propósito de este artículo se prescinde del cálculo de los parámetros de cada distribución, pudiéndose obtener estos directamente a partir de su relación con la recta de ajuste (Martínez, J., y Martín, M. J., 1990).

Caso 1. Se considera la muestra representada en la figura 2, que corresponde a los máximos anuales registrados durante un período de veinte años. Estos valores se

han representado en el papel probabilístico de Gumbel asignando a cada uno de ellos la frecuencia de representación dada por la fórmula de puntuo de Gringorten

$$f' = \frac{m - 0,44}{N + 0,12} \quad [7]$$

donde m corresponde al número de orden, ordenados los máximos de forma creciente, y N es el total de datos.

Según lo explicado en el apartado anterior, parece que esta muestra pertenece al dominio de atracción de Weibull (concavidad hacia abajo); no obstante, se va a realizar el ajuste a las tres distribuciones límite con el propósito de analizar las diferencias que hay entre ellas.

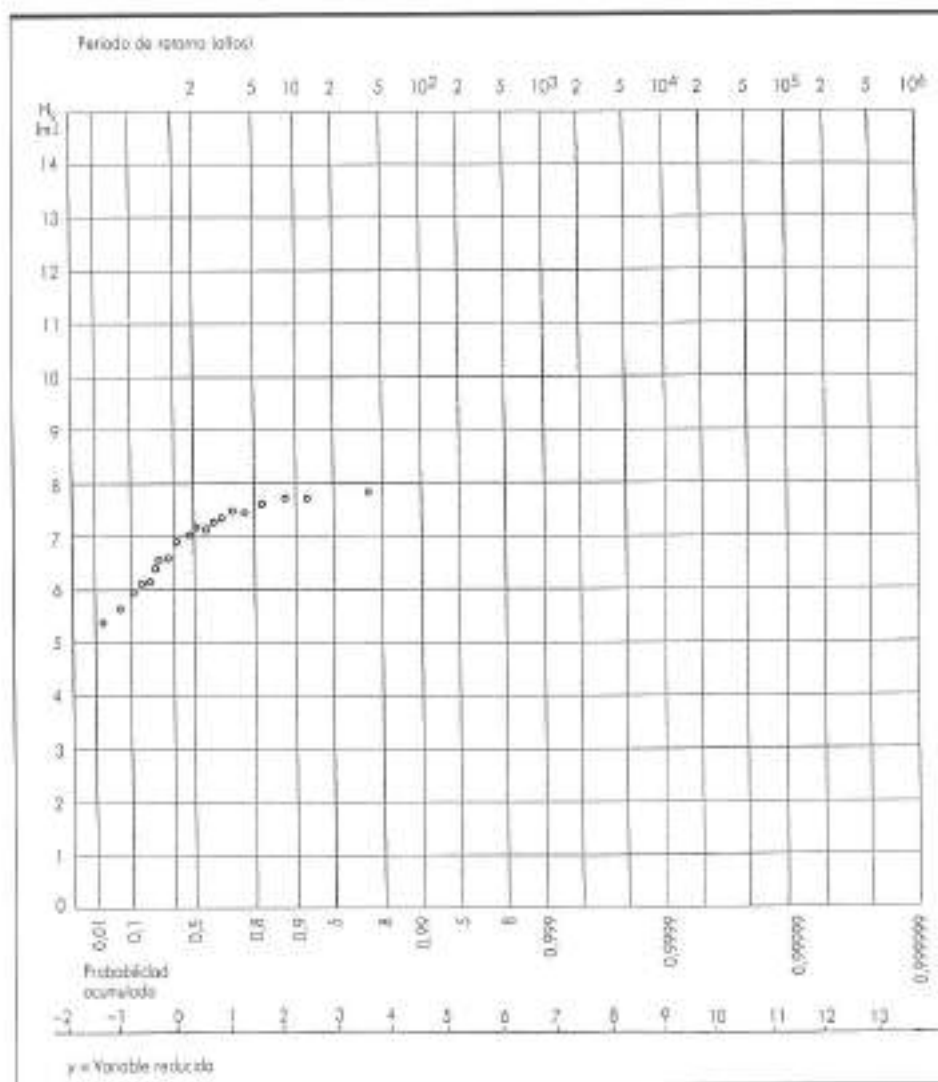
En la figura 3 se muestra el ajuste de los puntos a las asíntotas II y III, ya que el papel probabilístico es el

ALTURAS DE OLA (m)			FRECUENCIA DE PRESENTACION	
CASOS 1 y 2	CASO 3	CASO 4	CASOS 1, 2 y 3	CASO 4
5,4	5,9	5,9	0,0278	0,0293
5,7	6,2	6,2	0,0775	0,0816
6,0	6,2	6,2	0,1272	0,1339
6,2	6,4	6,4	0,1769	0,1862
6,2	6,6	6,6	0,2266	0,2385
6,5	6,7	6,7	0,2763	0,2908
6,7	6,9	6,9	0,3260	0,3431
6,7	6,9	6,9	0,3757	0,3954
6,9	7,1	7,1	0,4254	0,4477

7,0	7,2	7,2	0,4751	0,5000
7,2	7,2	7,2	0,5249	0,5523
7,2	7,3	7,3	0,5746	0,6046
7,3	7,4	7,4	0,6243	0,6569
7,4	7,6	7,6	0,6740	0,7092
7,5	7,8	7,8	0,7237	0,7615
7,5	7,9	7,9	0,7734	0,8138
7,6	7,9	7,9	0,8231	0,8661
7,7	8,2	8,2	0,8728	0,9184
7,7	8,5	8,5	0,9225	0,9707
7,8	10,7	-	0,9722	-

TABLA 2. Muestras utilizadas en los casos prácticos.

FIGURA 2. Representación de las alturas de ola en papel probabilístico de Gumbel. Casos 1 y 2.



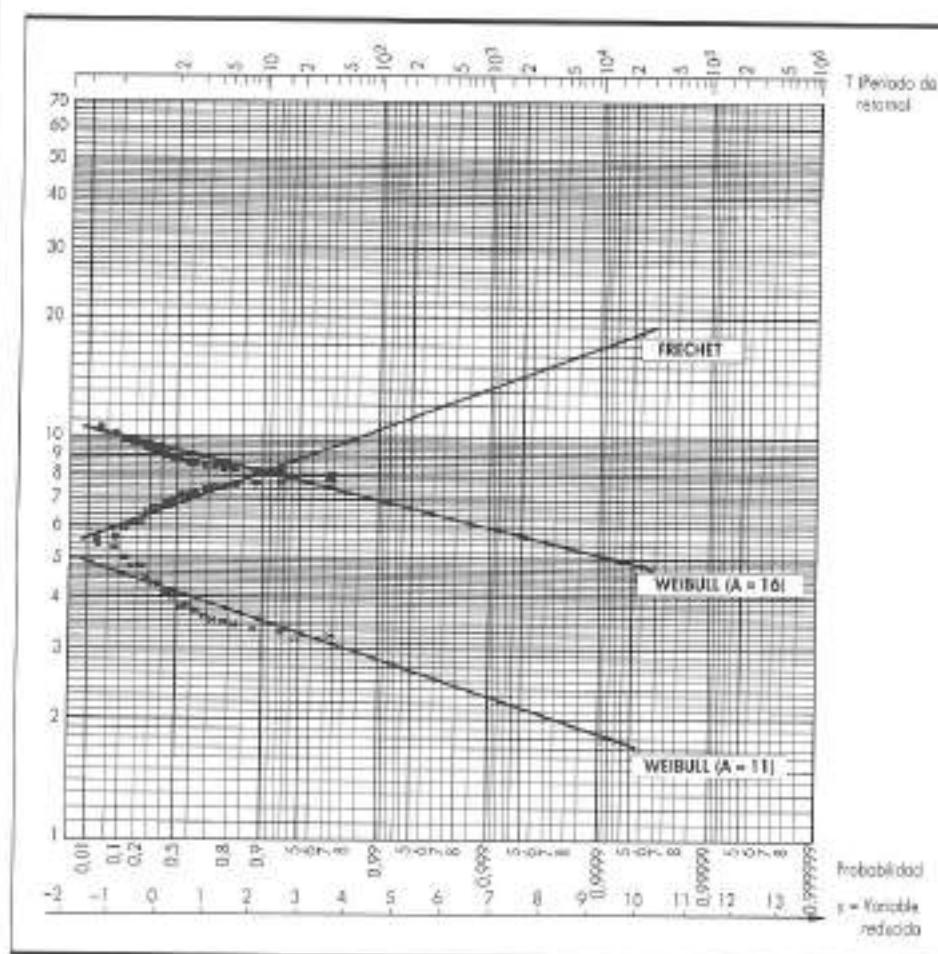


FIGURA 3. Ajuste a las asintotas II y III. Casos 1 y 2.

mismo para ambas (Copeiro, 1978) con la particularidad de que la recta correspondiente a la distribución de Weibull tiene pendiente negativa, y representa valores de la variable $A - x$.

En el caso del ajuste a la distribución de Weibull, se ha tomado un valor del parámetro de posición $A = 16$.

Con objeto de comparar las predicciones extremas correspondientes a las tres distribuciones, se han representado éstas conjuntamente en papel probabilístico de Gumbel (figura 4). Igualmente, en la tabla 3, se muestran las alturas de ola asociadas a ciertos períodos de retorno típicos en ingeniería marítima.

Como puede apreciarse, las diferencias son bastante significativas. Sin embargo, en este caso, y vista la tendencia que sigue la cola superior de la distribución, no parece que a la hora de proyectar una estructura portuaria sea excesivamente arriesgado elegir los valores de diseño proporcionados por la distribución de Weibull, con el consiguiente ahorro económico que ello supondría frente a las otras distribuciones.

Caso 2. Este caso es esencialmente el anterior, con una modificación cual es considerar un nuevo valor del pará-

T_R (años)	GUMBEL	FRECHET	WEIBULL
50	9,2	9,8	8,6
100	9,6	10,5	9,0
500	10,7	12,3	9,8

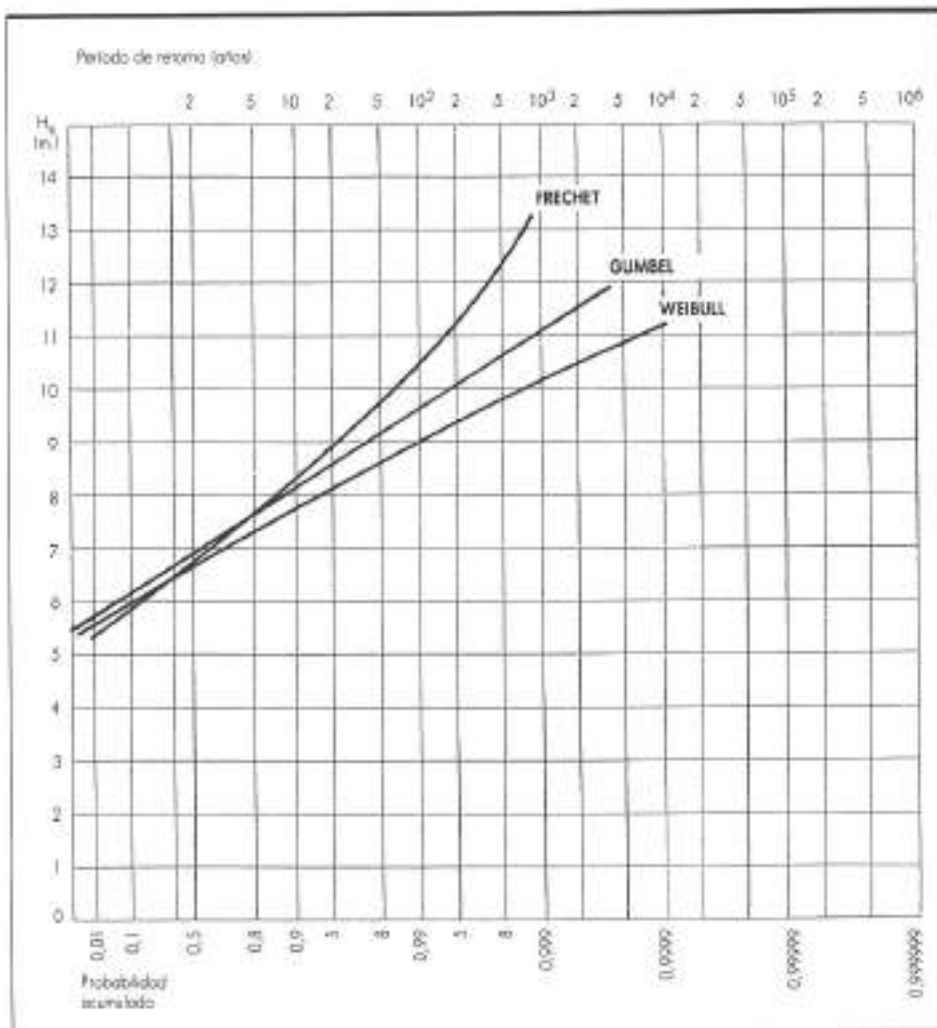
TABLA 3. Valores de retorno para el caso 1.

metro de posición en la distribución de Weibull, $A = 11$. De esta forma se obtiene el ajuste representado en la figura 3, que como puede apreciarse es bastante más pobre que el resultante de considerar el valor $A = 16$.

Las diferencias entre los valores de retorno, según se considere uno u otro valor de A , se muestran en la tabla 4.

En este caso la figura 2 pone de manifiesto que la distribución pertenece al dominio de atracción de Weibull, pero, a pesar de la diferente bondad de los ajustes, no hay razones objetivas que aseguren cual es el límite superior de la variable, pudiendo ser alguno de los ensa-

FIGURA 4. Regímenes extremos en papel probabilístico de Gumbel. Caso 1.



T_a (años)	$A = 16$	$A = 11$
50	8,6	8,0
100	9,0	8,2
500	9,8	8,7

TABLA 4. Influencia del parámetro de posición de la Asintota III.

yados o cualquier otro valor con la única condición de ser mayor que la máxima altura registrada.

Caso 3. En este caso la muestra corresponde también a una serie de veinte máximos anuales que, de manera análoga al caso anterior, se representan en papel probabilístico de Gumbel (figura 5). Ahora se observa que la cola superior tiene una tendencia contraria a la vista en el caso 1, presentando la concavidad hacia arriba, lo que hace suponer que pertenece al dominio de Fréchet.

En la figura 6 se muestran los ajustes correspondientes a las asíntotas II y III, habiéndose tomado para esta última el valor del parámetro de posición $A = 18$.

La representación conjunta, en papel probabilístico de Gumbel (figura 7), ofrece unos valores de retorno que se resumen en la tabla 5.

En este caso, contrariamente a lo ocurrido en los anteriores, es la distribución de Fréchet la que ofrece un mejor ajuste, tal como se aprecia en la figura 6, por lo que, al margen de otras consideraciones no menos importantes y que se tratarán posteriormente, sería ella la distribución elegida.

Como puede observarse en las representaciones gráficas, a medida que crecen los periodos de retorno aumentan las discrepancias entre las alturas pronosticadas por las tres asíntotas. Esto hace que, antes de decidirse por un modelo de distribución límite, haya que comprobar que la función de distribución original converge con suficiente rapidez a la asíntota elegida (en el

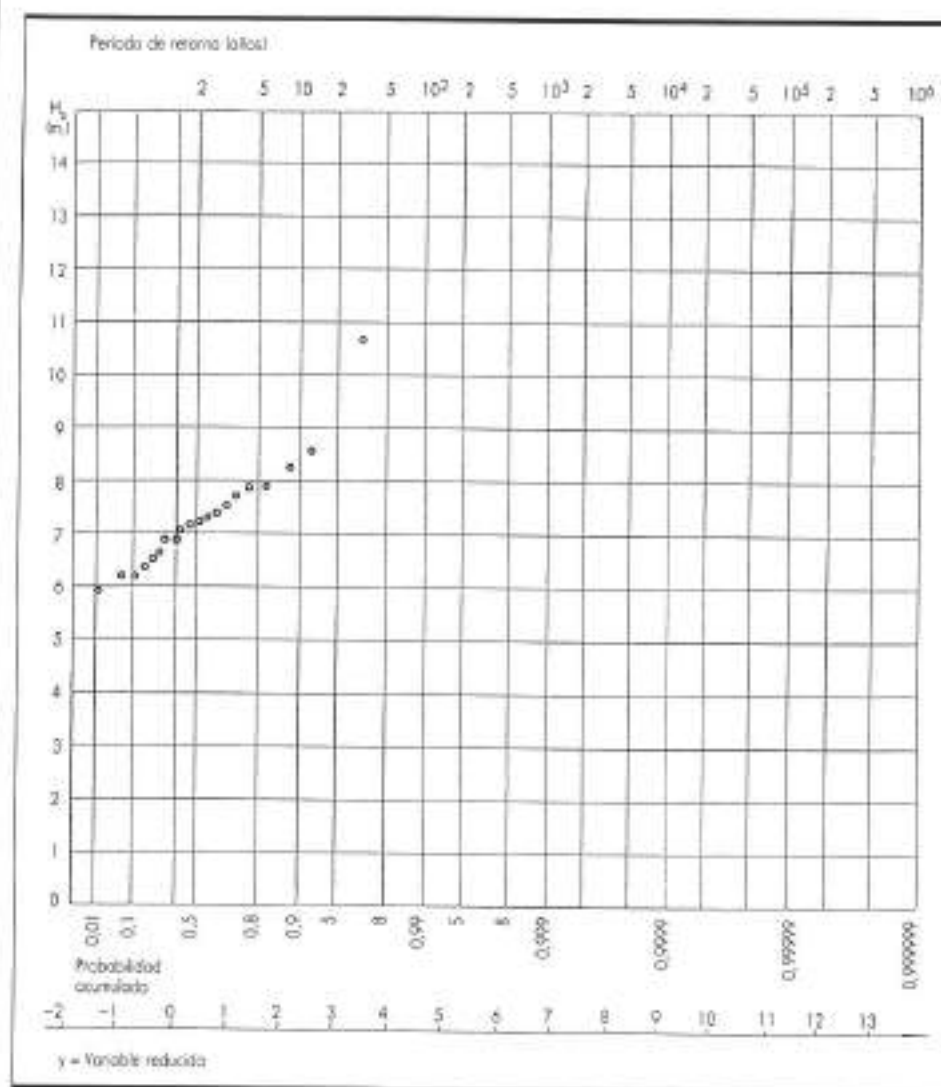


FIGURA 5. Representación de las alturas de ola en papel probabilístico de Gumbel. Caso 3.

caso de oleaje, al ser desconocida la distribución original, hay que hacer esta comprobación a partir de la muestra).

Caso 4. En este caso se trata de poner de manifiesto la importancia de los «outliers», que son puntos que se alejan considerablemente de la media muestral.

La presencia de un «outlier» puede deberse a causas bien distintas, por ejemplo un error de medida (en el caso de los datos visuales es relativamente frecuente), presencia de ruido en la señal registrada, errores en el proceso de cálculo (análisis de registros, hindcast, etc.) o también puede corresponder a un suceso excepcional (podría ocurrir que en un tiempo de medida habitual, por ejemplo diez años, tuviera lugar un suceso cuyo periodo de retorno real fuera muy superior a ese tiempo). En cualquier caso estos valores pueden alterar las muestras hasta el punto de hacerlas perder representa-

tividad de la población a la que pertenecen dando lugar a estimas sobredimensionadas.

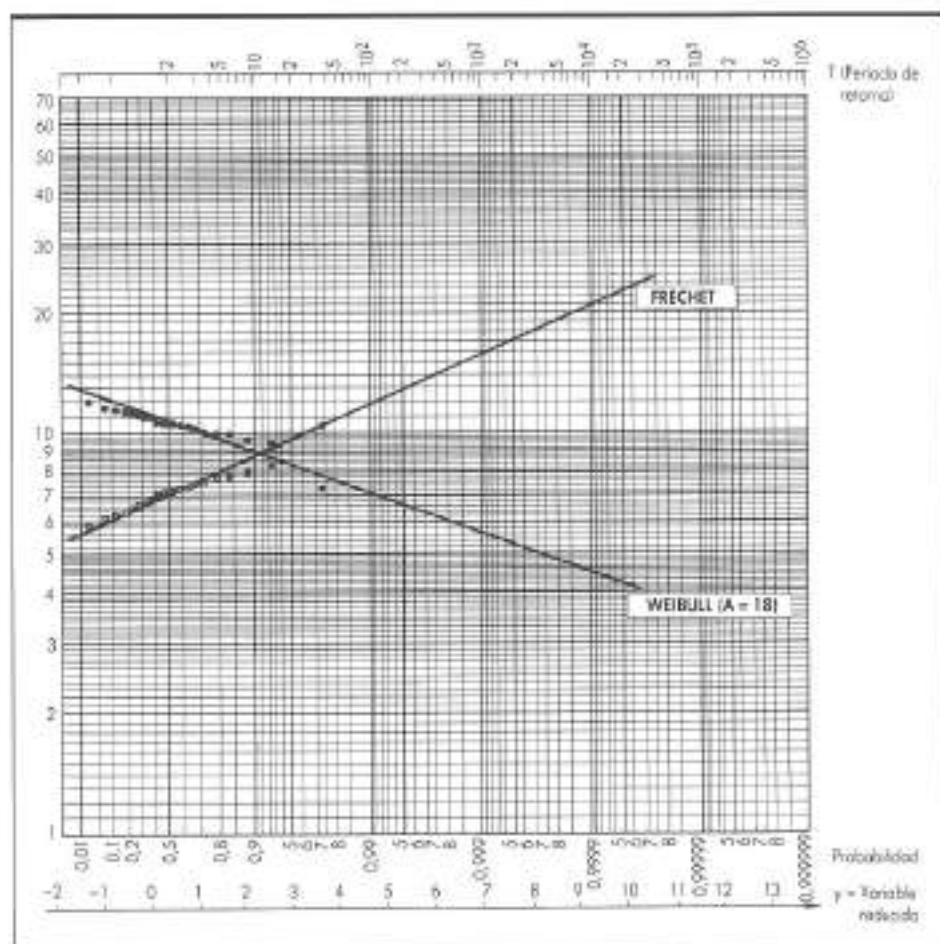
La muestra de partida para estudiar esta circunstancia es la misma que la del ejemplo anterior, pero suponiendo que el mayor valor, $H = 10,7$ m, es un «outlier» que se elimina, es decir, queda una muestra de diecinueve máximos, cuya representación gráfica en papel probabilístico de Gumbel, con frecuencias dadas por la fórmula de Gringorten, como en los casos anteriores, corresponde a la figura 8.

La figura 9 muestra los ajustes a las distribuciones de Fréchet y Weibull, manteniéndose para esta última el valor del parámetro de posición, $A = 18$.

Representando conjuntamente, en papel probabilístico de Gumbel, los ajustes a las tres distribuciones (figura 10) se aprecian los valores de retorno pronosticados por cada distribución. La tabla 6 resume estos valores.

En este caso la observación de la curvatura de la

FIGURA 6. Ajuste a las asintotas II y III. Caso 3.



T_R (años)	GUMBEL	FRECHET	WEIBULL
50	10,2	11,1	10,2
100	10,9	12,0	10,8
500	12,3	14,5	11,9

TABLA 5. Valores de retorno para el caso 3.

cola superior (figura 8) no resulta determinante para decidir el dominio de atracción de la distribución, salvo para descartar Fréchet, por lo que habría que utilizar alguno de los otros métodos antes mencionados. Otra posible solución sería elegir directamente la distribución de Gumbel, ya que no presenta grandes diferencias con la de Weibull y se queda del lado de la seguridad sin suponer un incremento considerable del coste de la obra.

5. RESULTADOS

De la comparación de los casos 1 y 2 se desprende la importancia del parámetro de posición en la distribución

T_R (años)	GUMBEL	FRECHET	WEIBULL
50	9,0	9,5	8,7
100	9,4	10,1	9,0
500	10,2	11,7	9,7

TABLA 6. Valores de retorno para el caso 4.

de Weibull, cuya existencia implica un límite superior para la altura de ola, debiéndose por tanto estimar dicho parámetro a partir de los datos muestrales.

Si se comparan los resultados de los dos últimos ejemplos, se observa que la inclusión del posible «outlier» (caso 3) proporciona, para una misma distribución asintótica, estimas notablemente superiores a las obtenidas cuando se elimina dicho valor (caso 4). Estas sobrestimaciones aumentan con el período de retorno, así mientras que para $T_R = 50$ años son del orden del 15 %, para $T_R = 500$ años superan el 20 %.

Sin embargo, con ser problemático el que para una misma distribución se obtengan estas diferencias, mu-

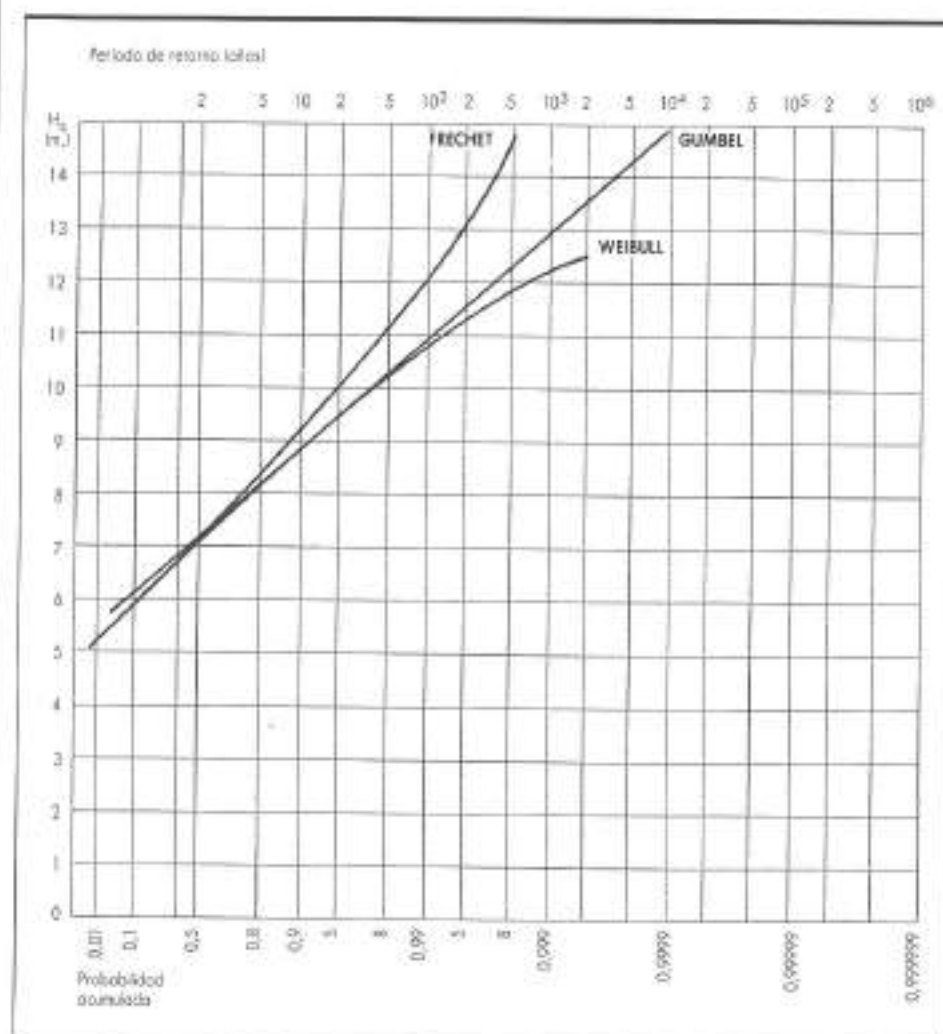


FIGURA 7. Regímenes extremos en papel probabilístico de Gumbel, Caso 3.

cho más grave es que el hecho de considerar o no el posible «outlier» determine la elección de una u otra distribución límite, ya que entonces las diferencias aumentan sensiblemente. Así, mientras que si se incluye el «outlier» la muestra parece pertenecer al dominio de atracción de Fréchet, obteniéndose para un período de retorno de 500 años una altura de ola $H_{500} = 14,5$ m, su no inclusión lleva a elegir la distribución de Gumbel cuya altura de ola asociada a ese período de retorno es $H_{500} = 10,2$ m.

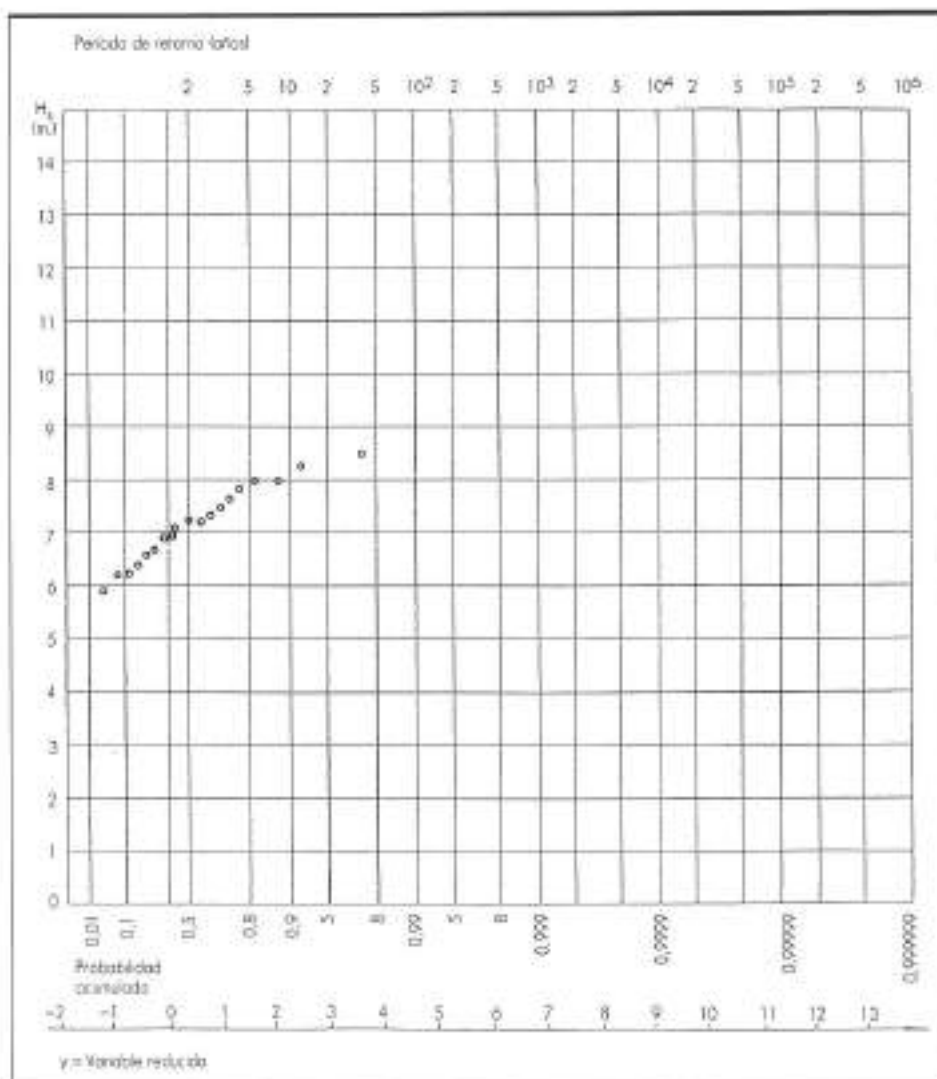
Esto hace que haya que tener en cuenta características propias de la naturaleza del fenómeno que se está estudiando, en este caso el oleaje, junto con algunas consideraciones relativas a las distribuciones asintóticas. En efecto, es sabido que el uso de la distribución de Fréchet sólo se justifica teóricamente para variables ilimitadas en la cola de interés, por lo que no parece muy apropiada para modelizar el oleaje que, por razones físicas, es una variable limitada.

Todo esto hace que el problema de los «outliers» sea de gran importancia en el análisis extremal del oleaje, por lo que su detección y tratamiento constituyen un tema prioritario en dicho análisis.

6. CONCLUSIONES

- Métodos frecuentemente empleados para calcular el régimen extremal de oleaje requieren la obtención de la distribución límite, la cual puede reducirse a una de las tres distribuciones asintóticas: Gumbel, Fréchet y Weibull.
- Razones físicas, relativas a la propia naturaleza del oleaje, desaconsejan el uso de la distribución de Fréchet para el cálculo del régimen extremal. Un buen ajuste a esta distribución podría deberse a la presencia de «outliers».
- La distribución de Weibull implica la determinación de un valor máximo de la altura de ola (parámetro

FIGURA 8. Representación de las alturas de ola en papel probabilístico de Gumbel. Caso 4.



de posición), cuya estima, caso de ser éste el dominio de atracción de la variable y habida cuenta de la sensibilidad de los resultados frente a dicho parámetro, debe realizarse con mucha precisión. Para el caso de profundidades reducidas puede suponerse igual a la altura de ola máxima por limitación de fondos.

- La distribución de Gumbel puede aplicarse a todo el campo de valores de la variable, y presenta la particularidad de que puede ser aproximada por las otras dos asíntotas.
- Cuando la convergencia a la asíntota correspondiente sea muy lenta debe descartarse el empleo de la distribución límite para el cálculo del régimen extremo. En estos casos se recomienda utilizar otra metodología de cálculo, por ejemplo, el método de la distribución inicial, que parte de estimar la distribu-

ción original y extrapolarla al período de retorno que se desee.

AGRADECIMIENTOS

Este artículo se ha realizado contando con la ayuda del Sector de Estudios Portuarios, y en especial de María J. Martín Soldevilla, a quien se agradece especialmente por su colaboración y comentarios.

REFERENCIAS

- CASTILLO, E., y GALAMBOS, J. (1986). «Determining the domain of attraction of an extreme value distribution». *Technical Report, Temple University, Philadelphia*, pp. 1-12.
- CASTILLO, E. (1988). «Extreme value theory in engineering». Academic Press, San Diego, California.

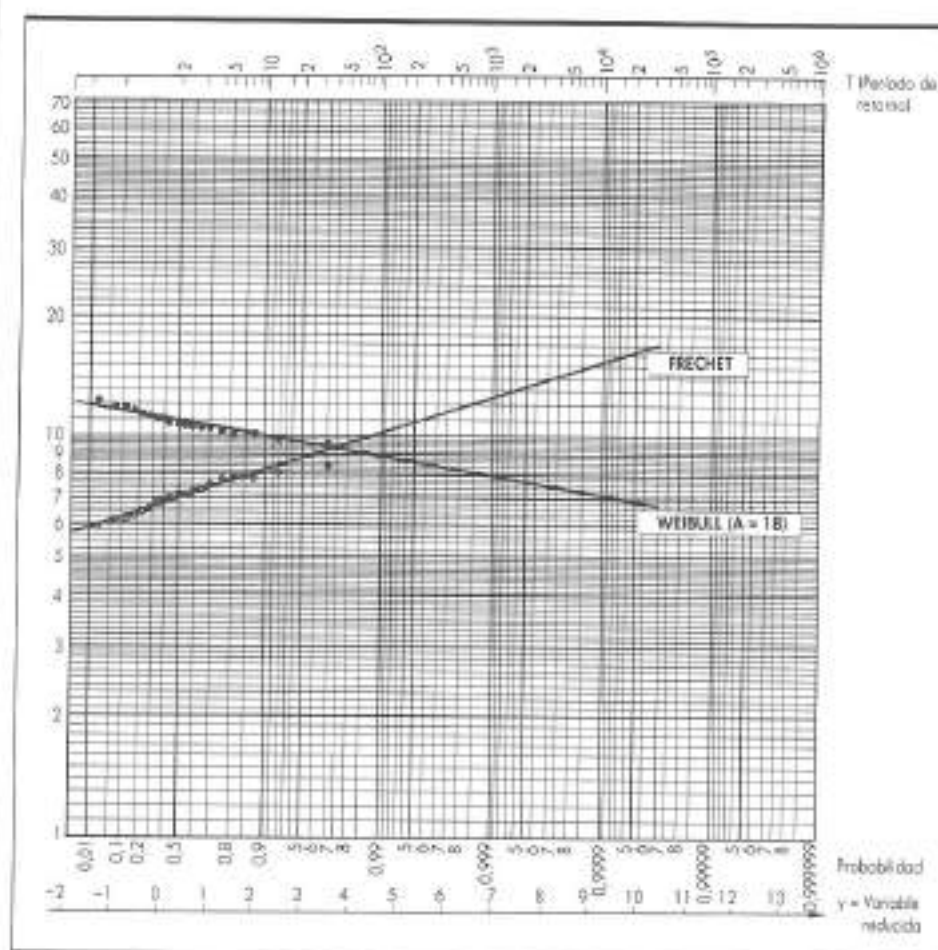


FIGURA 9. Ajuste a las asintotas II y III, Caso 4.

CASTILLO, E., y SARABIA, J. M. (1992). «Engineering analysis of extreme value data. Selection of models». *Journal of Waterways, Port, Coastal and Ocean Eng.*, vol. 118, núm. 2, pp. 129-146.

COPEIRO, E. (1978). «Análisis extremal de variables geofísicas». Laboratorio de Puertos Ramón Iribarren, CEDEX, publ. núm. 15.

GALAMBOS, J. (1980). «A statistical test for extreme value distributions». *Colloquia Mathematica Societatis János Bolyai*, pp. 221-229.

FISHER, R. A., y TIPPETT, L. H. C. (1928). «Limiting forms of the frequency distribution of the largest or smallest member of a sample». *Proc. Cambridge Phil. Soc.*, vol. 24, pp. 180-190.

MARTINEZ, J., y MARTIN, M. J. (1990). «Descripción de las funciones de distribución más utilizadas en estudios de clima marítimo». Monografías CEDEX, vol. M11.

PICKANDS, J. III (1975). «Statistical inference using extreme order statistics». *The Annals of Statistics*, Univ. of Pennsylvania, vol. 3, núm. 1, pp. 119-131.

FIGURA 10. Regímenes extremos en papel probabilístico de Gumbel, Caso 4.

