

Función de distribución extremal en fenómenos de tipo meteorológico^(*)

Por ENRIQUE COPEIRO

Dr. Ingeniero de Caminos, C. y P.
Laboratorio de Puertos.
Centro de Estudios y Experimentación
de Obras Públicas.

Se discute la metodología en uso para la determinación de función de distribución extremal en fenómenos de tipo meteorológico. Se analizan las deficiencias y se proponen procedimientos alternativos desde el punto de vista de su utilidad práctica en ingeniería.

1. INTRODUCCION

El análisis estadístico extremal de variables de tipo meteorológico (o geofísicas) como la velocidad del viento, la precipitación, el caudal fluvial o la altura de ola, viene representando cada vez más uno de los elementos básicos en los proyectos de obras civiles que se ven afectados por aquellas variables.

El primer texto dedicado a la estadística de valores extremos fue debido a E. Gumbel (1958). Aunque anteriormente se habían publicado algunos análisis extremos, puede decirse que la aparición de este libro marcó, al menos en los países tecnológicamente avanzados, el comienzo del uso generalizado del análisis extremal en la práctica ingenieril. El texto de Gumbel sigue siendo hoy la referencia básica del análisis extremal aplicado.

Sin embargo, no puede decirse que estas técnicas hayan sido aún desarrolladas de forma satisfactoria en todos sus aspectos. A pesar del tiempo transcurrido y de la enorme importancia práctica de este tema, una revisión de los resultados a que han llegado los numerosos autores que han realizado aplicaciones de la teoría extremal descubren un extenso desacuerdo. *En lugar de haberse dado un proceso de homogeneización progresiva de los criterios, como suponía el propio Gumbel en su libro, en realidad el tiempo ha ido acentuando las divergencias.* Cada vez se proponen más funciones distintas para la distribución extremal de una misma variable y hasta dentro de un mismo clima, funciones que al ser extrapoladas para su uso práctico divergen ampliamente implicando impactos sumamente diferentes en los proyectos.

Al propio tiempo son varios los autores que, al tratar este tema, explícitamente aseguran no

(*) Se admiten comentarios sobre el presente artículo, que pueden remitirse a la Redacción de esta Revista, hasta el 31 de julio de 1979.

existir un criterio con fundamento que permita discriminar la conveniencia del empleo de unos u otros métodos en el proceso de cálculo. Es decir, *podemos hablar de una situación de ambigüedad reconocida y grave.* La gravedad de esta situación la experimentan fundamentalmente los ingenieros que deben tomar la decisión de elegir por una de las posibilidades existentes, sabiendo que esas posibilidades pueden representar serias variaciones en el precio y la seguridad de las obras que proyectan, y sabiendo (o lo que es peor, sin saber) que no disponen de un criterio reconocido de elección.

En un artículo anterior (14) se trató de las fuentes de ambigüedad existentes en los criterios habituales de ajuste de función de distribución a partir de muestreos aleatorios. En el presente artículo se van a aplicar los resultados de aquel análisis a la elección de función de distribución extremal, con objeto de llegar a criterios de fiabilidad práctica desde un punto de vista ingenieril.

2. ESTADO ACTUAL DE LA METODOLOGIA

El sistema habitual de determinar qué tipo de función de distribución extremal corresponde a una determinada variable, se basa en obtener una o más muestras extremas de la variable y comprobar las bondades de los ajustes conseguidos con algunas funciones de distribución. *Respecto a qué funciones conviene probar, son numerosos los autores que han utilizado con exclusividad las denominadas "Distribuciones Extremales Asintóticas" (o Asintotas I, II y III).* Estas tres distribuciones son el producto final de la teoría de distribuciones asintóticas de valores extremos, popularizada en el tratado de Gumbel, y los autores aludidos han venido suponiendo, explícita o implícitamente, que su fundamento teórico como distribuciones extremas hace de las Asintotas los candidatos óptimos para la solución de cada caso (ver, por ejemplo, 21). Sin embargo, como se in-

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

dicó anteriormente, se ha llegado a una gran heterogeneidad de resultados en su uso práctico. A título de ejemplo se citan algunos casos significativos:

Precipitación caída en un intervalo de tiempo.

S. Nag y N. Dutta (1951) usaron la Asíntota-II en un estudio de precipitaciones máximas diarias, en la estación del monzón, dentro de la cuenca del río Barakar. Sin embargo, E. Gumbel (1958) probó que la Asíntota-I ajustaba mejor la misma serie de datos. A. Jenkinson (1955), por otra parte, consiguió los mejores ajustes con la Asíntota-II para la máxima precipitación en períodos de una hora, un día y cuatro días en Marsella, Nápoles y Trípoli. Puede señalarse que las Asíntotas-I y II producen extrapolaciones que divergen muy ampliamente, la segunda de ellas alcanzando valores considerablemente más elevados. En la figura 2.2.2 de la referencia (12) se puede ver cómo la Asíntota-I ajusta muy satisfactoriamente a las precipitaciones diarias máximas de Madrid, y en la figura 5.3.9 de la misma referencia se observa que una Asíntota-I también puede ajustar bien a las máximas precipitaciones mensuales en Los Llanos (Cuenca).

N. Barricelli (1943) utilizó la Asíntota-I para el ajuste de las *máximas presiones atmosféricas* anuales en Bergen, Noruega. Pero para la misma variable medida en Marsella, A. Jenkinson (1955) encontró más ajustada la Asíntota-II.

A. Court (1953) utilizó con éxito la Asíntota-I al estudiar las *velocidades máximas del viento* (medias en cinco minutos) en 25 estaciones de Estados Unidos. J. Bell (1961) usó también la Asíntota-I para la velocidad máxima del viento en Hong-Kong, y H. Wittingham (1964) para máximas rachas del viento en Australia. A. Davenport (1960) propuso emplear esa función de distribución como método estándar de análisis extremal del viento en las normas canadienses. Pero H. Thom, en una serie de estudios de velocidad máxima del viento en distintas partes del globo (1954, 1960, 1963, 1968, 1969, 1973) ha empleado siempre la Asíntota-II. A. Jenkinson (1955) empleó la Asíntota-III para la máxima velocidad media horaria del viento en Trieste. Por otra parte, en la figura 5.3.9 de la referencia (12) puede verse que una Asíntota-I ajustaría bien a las máximas velocidades medias diarias en Valladolid.

E. Gumbel (1940) utilizó por primera vez la Asíntota-I para el análisis extremal de *avenidas fluviales*. Después, la misma función de distribución ha sido empleada en una multitud de casos referentes a ríos americanos por W. Potter (1949), S. Rantz y H. Riggs (1954), R. Carter (1951), y E. Gumbel (1958) entre otros. A. Benham (1950) también la aplicó con resultados satisfactorios a ríos de Nueva Zelanda, y Shuh (1952) a China. Sin em-

bargo, J. Bernier (1956) encontró que la Asíntota-II proporcionaba mejores ajustes para las máximas avenidas diarias del Rhin en Rheinfelden, del Colorado en Black Canyon y del Duranse en Archidiacre. A. Jenkinson (1955) empleó también la Asíntota-II para Little River, pero prefirió una Asíntota-III para el río Connecticut.

M. St. Denis (1969) propuso utilizar la Asíntota-I al análisis extremal de la *altura de ola significativa*. P. Suárez Bores (1974) empleó esta distribución en el ajuste de varias muestras extremas de altura de ola significativa estimadas para puntos de la costa española por medio de estudios de previsión a partir de cartas barométricas y regímenes de viento. H. Thom (1970) prefirió, en cambio, la Asíntota-II para ajustar las observaciones máximas anuales de barcos meteorológicos situados en el Atlántico y Pacífico. El autor del presente trabajo (E. Copeiro, 1976) obtuvo, en algunos barcos meteorológico del Atlántico Norte, muestras extremas más largas que las utilizadas por Thom y consiguió ajustes satisfactorios con la Asíntota-I. Por otra parte, en un artículo posterior al reseñado arriba M. St. Denis (1973) se declaró partidario de la Asíntota-III por estimar que, según su experiencia, la Asíntota-I da predicciones excesivamente altas.

Como puede verse, se ha llegado a todo tipo de resultados para una misma variable. Es verdad que, entre los usuarios de las Asíntotas, han sido sólo una exigua minoría los que han intentado justificar que las condiciones en que fue deducida la Asíntota escogida como distribución de extremos se satisfacen en el caso concreto en que se la está empleando. Estos intentos de justificación, por lo demás, no han sido acompañados por el éxito debido (como se verá más adelante) a una interpretación errónea de la aplicabilidad de la ecuación extremal a las variables geofísicas.

La generalidad de autores que han empleado las Asíntotas se han limitado a elegir entre ellas de acuerdo con el criterio estrictamente empírico de su mejor o peor ajuste a la muestra extremal disponible. *Esto ha terminado por colocar a las tres Asíntotas en el campo del puro empirismo, donde se las alinea en igualdad de condiciones con otras funciones de distribución a la hora de probar cuál da mejor ajuste a las muestras.* Puede afirmarse que este punto de vista es el dominante en la actualidad. Como ejemplo puede citarse un comentario de J. Bernier, que en 1956 (6) había publicado un artículo sobre la aplicación de la Asíntota-II a avenidas fluviales máximas y en el año 1963 (7) pasa a ser de la opinión que es preferible hacer cada elección en función del ajuste conseguido, para lo cual lo más conveniente es disponer de un amplio arsenal de funciones de distribución que probar. Concluye que no debe con-

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

cederse un papel especial a las Asíntotas, ya que otras funciones, como por ejemplo la Log-Normal, dan en ocasiones mejores ajustes.

En esta línea, puede mencionarse (5) el estudio comparativo realizado por el Grupo de Trabajo Federal sobre métodos de frecuencias de avenidas (Water Resources Council, U.S.A.), en que se compararon las distribuciones Asíntota-I, Asíntota-II, Log-Normal, Log-Pearson-III, Gamma-II y Hazen, con ayuda de 10 muestras extremales. En el campo de estudios extremales de oleaje, dos ejemplos recientes de la misma tendencia pueden verse en las referencias (24) (Asíntota-I y Weibull) y (9) (Normal y Weibull).

En estas selecciones de distribución extremal en base a los ajustes conseguidos con muestras de valores extremos, los distintos autores han optado por uno de los dos criterios siguientes:

1. Selección a partir de varias muestras extremales de la variable en cuestión.

Se escoge aquella función de distribución que mejores ajustes dé al conjunto de muestras, y esa función se emplea en todos los casos correspondientes a aquella variable. Obviamente, este punto de vista parte implícitamente de suponer que cada variable específica posee un mismo tipo de distribución extremal en cualquier localización.

2. Selección para cada muestra extremal.

Este criterio, estrictamente casuístico, parte de suponer que una misma variable puede en principio adoptar tipos totalmente diferentes de distribución extremal en localizaciones distintas. Se escoge en cada caso aquella función que da un mejor ajuste a la muestra extremal disponible.

A continuación se discute la conveniencia de estos criterios.

3. DISCUSION DE LOS CRITERIOS EN USO

3.1. *El último de los dos criterios vistos (la elección casuística)* es indudablemente cómodo, puesto que se presta muy bien a reducir el análisis extremal de cada caso a una rutina de tanteos relativamente fácil de mecanizar utilizando ordenadores y sin necesidad de acudir a consideraciones generales. Sin embargo, un somero análisis de los resultados a que puede llegarse con esta técnica muestra que su uso es a la vez insatisfactorio desde un punto de vista teórico y altamente peligroso desde la perspectiva ingenieril.

Aun si fueran aceptados como válidos los criterios habituales de bondad de ajuste que tratan uniformemente la totalidad del recorrido muestral (lo que es incorrecto, según se discute en (12 y 14)),

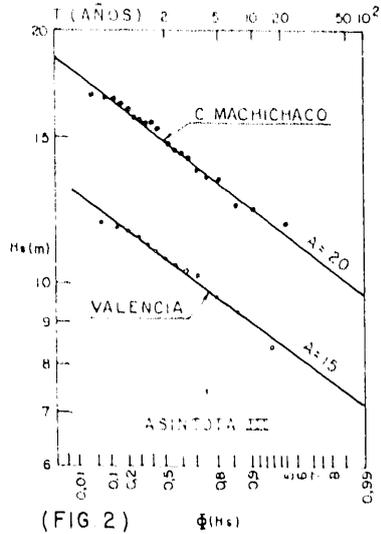
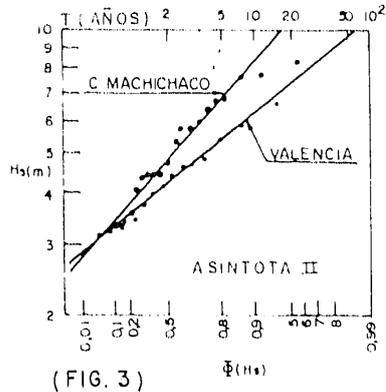
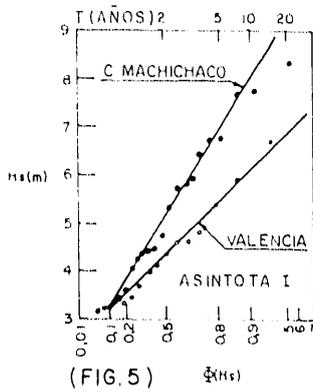
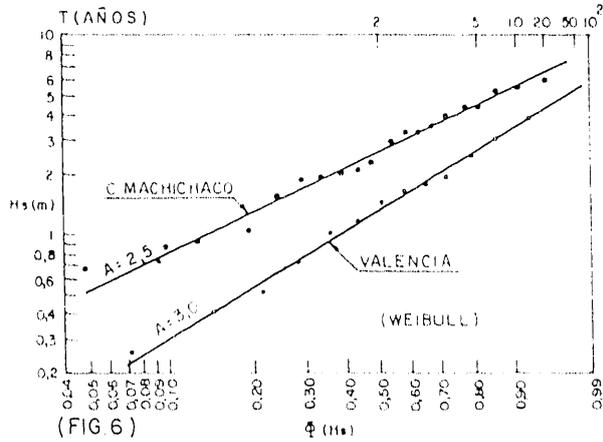
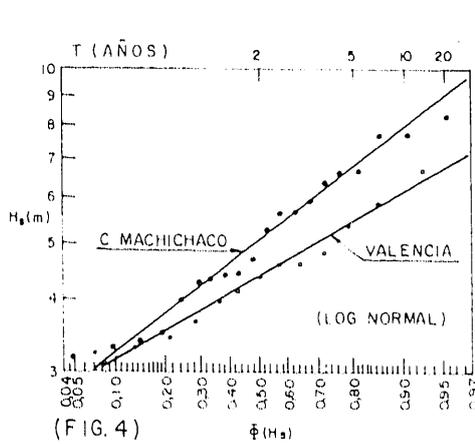
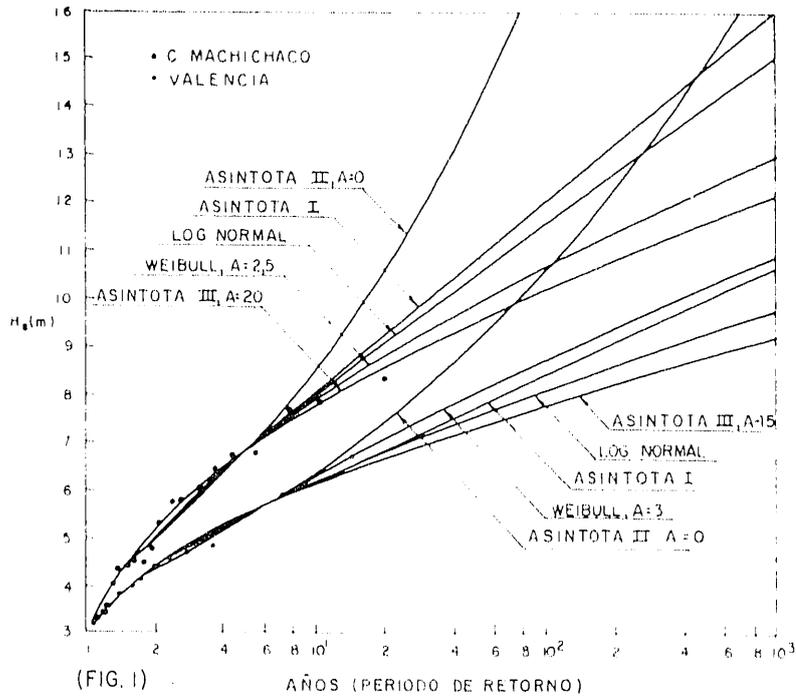
estos criterios no son capaces de proporcionar, a partir de una sola muestra, una elección de FDD satisfactoria en la práctica cuando se pretende extrapolar aquella FDD. La razón está en que, para los tamaños muestrales que son corrientes en la realidad, hay normalmente varios tipos distintos de FDD que son capaces de proporcionar ajustes de bondades muy similares. Esas funciones toman valores muy próximos a lo largo del recorrido de los valores muestrales, pero al ser extrapoladas pueden divergir ampliamente produciendo resultados muy diferentes al nivel de los períodos de retorno requeridos usualmente para el uso práctico. En estos casos (que son la norma más bien que la excepción) no es ingenierilmente permisible el realizar la elección entre funciones con el solo criterio de una pequeña diferencia entre las bondades calculadas de los ajustes: Esas diferencias menores son perfectamente atribuidas a la inevitable variabilidad entre muestras distintas de una misma población.

Para mostrar un caso real se han tomado dos muestras extremales de altura de ola significativa, obtenidas mediante métodos de previsión para Cabo Machichaco (Vizcaya) y Valencia (tomadas de P. Suárez Bores, 1974). En las figuras 2 a 4 se muestra el resultado de ajustar visualmente ambas muestras, representadas según metodología de E. Gumbel (1958), con funciones de distribución Asíntota-I, Asíntota-II, Asíntota-III, Weibull y Log-Normal. Todos los ajustes son aceptables según los contrastes estadísticos habituales, si bien existen pequeñas diferencias entre ellos. En la figura 1 se comparan las extrapolaciones correspondientes a las distintas distribuciones. Las diferencias, al nivel de los períodos de retorno usuales (por ejemplo, $T = 500$ años: Riesgo de superación en 50 años = 10 por 100) son muy grandes, totalmente fuera de proporción con las diferencias entre los ajustes. Es tan importante la repercusión que supone, en proyectos de obras marítimas, la diferencia entre estas estimas, que un ingeniero que tuviese que hacer uso práctico de estos resultados no podría permitirse el hacer la elección entre las distribuciones con el solo criterio de las pequeñas diferencias en la bondad de aquellos ajustes. Para realizar en este caso una elección razonable, sería necesario acudir a otro tipo de criterio.

Cuando se utiliza un criterio racional para los ajustes, que tenga en cuenta la fiabilidad variable de la estima muestral a lo largo de su rango de valores (12 y 14), se desechan dos tramos en los extremos del recorrido muestral. El recorrido muestral restante (*útil o efectivo*) es más corto y resulta, por tanto, aún más susceptible de ser estrechamente ajustado por un amplio número de FDD diferentes.

Puede concluirse que, al menos en el campo

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO



de los tamaños muestrales habituales y a efectos de extrapolar resultados a períodos de retorno también habituales (superiores al menos en un orden de magnitud al tamaño muestral, según es corriente), *la mejor bondad de ajuste no proporciona un criterio fiable de elección de FDD cuando se analiza una sola muestra aisladamente.*

3.2. *El otro de los criterios mencionados en el apartado anterior (la selección a partir de varias muestras extremas de la variable en cuestión), tampoco ha producido hasta el momento elecciones satisfactorias a pesar de que, como se verá a continuación, está en una posición teórica favorable.*

Un caso significativo que es interesante comentar, a este respecto, es el estudio comparativo citado anteriormente (5) en que se comparó el comportamiento de seis funciones de distribución con 10 muestras extremas correspondientes a avenidas fluviales. El Grupo de Trabajo Federal escogió la distribución Log-Pearson-III por ser la que mejores bondades de ajuste dio en el conjunto de muestras. Sin embargo, algunos años más tarde (8) otros autores compararon aquella función con la Pearson-III, con ayuda de otras 27 muestras extremas pertenecientes a 13 países distintos de cuatro continentes, y concluyeron que la distribución Pearson-III proporciona mejores ajustes y debe, por tanto, ser preferida a la otra. Si ambas distribuciones tomaran valores muy próximos en las extrapolaciones, el problema planteado sería sólo aparente puesto que desde un punto de vista práctico el uso de una u otra función daría resultados muy similares y, desde el punto de vista teórico, no debe olvidarse que este tipo de búsqueda de FDD sólo pretende llegar a una aproximación empírica de la distribución real. Sin embargo, de hecho las extrapolaciones producidas por las distribuciones Pearson-III y Log-Pearson-III divergen ampliamente, originando repercusiones importantes en su uso práctico dentro de los períodos de retorno usuales. Por tanto, el problema de selección entre ellas resulta crucial.

Se podrían citar otros casos análogos a este último. No parece que la metodología en uso actualmente esté en camino de proporcionar selecciones fiables. Por el contrario, a medida que se incorporan nuevas funciones y nuevas muestras a los estudios comparativos, la corona de favorita va pasando de una distribución a otra de una forma realmente caótica, desgraciadamente alejada del deseable proceso convergente que fuera centrando los resultados en un acercamiento progresivo a la realidad.

La raíz del problema está en la aplicación de los criterios de bondad de ajuste. Los criterios usuales de ajuste, y en particular los que utilizan las ventajas del papel probabilístico, emplean para

el ajuste todos los puntos muestrales y en igualdad de condiciones. Este tipo de criterio está sometido a errores potencialmente importantes. Puede demostrarse fácilmente (12 y 14) que la fiabilidad de las estimas muestrales de función de distribución no es uniforme, sino que varía a lo largo del recorrido de valores muestrales. En el caso de muestras aleatorias (de los que las muestras extremas son en general casos típicos), la fiabilidad (aproximación esperada de los valores estimados mediante el muestreo, a los valores reales de la población) de la probabilidad muestral es máxima en el centro de la distribución de la población ($F(x) = 0,5$) y disminuye hacia ambas colas de ella. Por lo tanto, *suponiendo que la muestra obtenida tenga un tamaño suficiente, en la serie de puntos correspondientes a las probabilidades muestrales pueden distinguirse dos regiones características:*

- Una zona "central", donde las probabilidades muestrales constituyen una buena (según criterio del usuario) aproximación esperada de las probabilidades de la población.
- Dos zonas, situadas en ambas colas de la serie de puntos muestrales, donde la aproximación esperada es pobre y, por tanto, existe una elevada probabilidad de que las probabilidades muestrales presenten amplias desviaciones aleatorias respecto a las de la población. La aleatoriedad de esas desviaciones no excluye la existencia de un sesgo en ellas, dado por la distribución binomial, empleada en (12 y 14) para mostrar la existencia de la dispersión).

Si se conoce previamente el tipo de FDD adecuada, un criterio eficaz de ajuste consiste en eliminar las dos zonas extremas de baja fiabilidad, ajustando sólo la zona central. Cuando (como es el caso que nos ocupa) el problema estriba precisamente en averiguar cuál es el tipo de FDD adecuada, a efectos de extrapolar en uno de los dos sentidos, el estudio del comportamiento de las colas de los puntos muestrales proporciona un criterio de gran utilidad práctica. Para probar la conveniencia de determinado tipo de FDD, cuando se dispone de un cierto número de muestras aleatorias, se ajusta con aquella función el tramo "central" de cada una de las muestras y se observa la tendencia de las colas muestrales situadas en el extremo correspondiente al sentido en que desea extrapolarse la FDD. *Si el conjunto de colas presenta una dispersión aleatoria (contando con el sesgo indicado antes) en torno a la FDD ajustada, no hay razón para rechazar esa FDD. Por el contrario, la FDD se rechaza si se observa sistematicidad en las desviaciones de las colas. A medida que se vaya disponiendo de mayor número y ta-*

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

maño de muestras y se prueben nuevos tipos de FDD, el uso de este criterio irá centrando progresivamente los resultados hacia la distribución real de la población, para cada variable estudiada.

Una ventaja adicional del criterio propuesto está en que permite analizar aisladamente cada cola de la FDD. La conveniencia práctica de centrar el análisis en la cola de interés para las extrapolaciones está en que, en ocasiones, el tipo de FDD que se comporta más adecuadamente en las extrapolaciones hacia un sentido, es distinto del más adecuado para extrapolar hacia el sentido opuesto. Este fue el resultado obtenido en la referencia (12), donde el criterio de elección que se está tratando fue empleado satisfactoriamente en la selección de FDD para la variable altura de ola significativa (con muestras casi continuas, que incluyen algunas peculiaridades que se indicarán más adelante). Cuando se da un caso así, los criterios "convencionales" de selección, basados en la bondad del ajuste bruto a la totalidad del recorrido muestral, son doblemente inadecuados.

En cualquier caso, el criterio de elección propuesto precisa del uso de un cierto número de muestras distintas de la misma variable. Para conseguir estas muestras puede operarse de dos formas:

1. Efectuar mediciones en un cierto punto durante un tiempo muy largo para conseguir suficientes muestras de tamaño aceptable. Esto es, en general, inviable, a nivel práctico, cuando se trata de seleccionar FDD extremal. Resultaría, en cambio, abordable para la FDD no extremal en el caso de bastantes variables de interés.

2. Tomar muestras individuales en un cierto número de puntos. Esta opción resulta más asequible que la anterior, y es la única de las dos abordable para distribuciones extremales. Sin embargo, se precisa formular la hipótesis previa de que la variable en cuestión sigue el mismo tipo de FDD en todos los lugares muestreados. Si bien la hipótesis de que cada variable obedece a la misma FDD en cualquier lugar es una generalización excesiva, resulta en cambio razonable el suponer que existen zonas de comportamiento homogéneo, dentro de las cuales la variable sí sigue un mismo tipo de FDD. Esto supone establecer una discriminación entre tipos "climáticos" diferenciales de cada variable. En tal discriminación, es lógico suponer que el comportamiento estadístico diferencial entre aquellos tipos responde a cambios sustanciales en la constitución física de la variable; de ahí la denominación alusiva al clima. Esta hipótesis fue empleada con buen resultado en (12) al realizar la elección (aludida antes) de FDD para la altura de ola significativa: aparecieron dos tipos de FDD distintos según que dominase el "Sea" o

el "Swell" en el lugar de observación y en los rangos de altura de ola considerados.

En resumen, se considera inviable la elección de FDD extremal a partir de una sola muestra extremal y sin otra información adicional, pero *resulta abordable la elección cuando se dispone de un cierto número de muestras extremales de la misma variable* y se utiliza el criterio de elección que ha sido propuesto, junto con la hipótesis de diferenciación climática.

Existen, sin embargo, variables para las que aún aquella segunda alternativa es inviable por carecerse de muestras extremales suficientes en número, extensión y/o precisión. Un caso significativo a este respecto es la altura de ola. Para abordar estos casos se propone utilizar directamente la ecuación extremal, que no precisa de muestras de valores extremos. Como se verá en los apartados siguientes, *la estima de los parámetros relevantes en la ecuación extremal precisa de tiempos de observación muy inferiores (al menos en un orden de magnitud) a los que serían necesarios para obtener resultados de fiabilidad comparable utilizando muestras extremales con los criterios expuestos hasta aquí.*

4. ECUACION EXTREMAL

La función de distribución extremal de una variable expresa la probabilidad de que un valor cualquiera de esa variable no sea superado (o no sea minorado) en un intervalo de tiempo dado, es decir, la probabilidad de que aquel valor sea el máximo (o mínimo) que se presente en el intervalo.

En la hipótesis de que una variable X es aleatoria, su función de distribución $F(x)$ en un intervalo es conocida, y a lo largo de ese intervalo la variable toma un número n de valores estadísticamente independientes, la función de distribución extremal $\Phi(x)$ en el intervalo es:

$$\Phi(x) = [F(x)]^n \quad [1]$$

Esta sencilla ecuación es la piedra angular de la estadística de extremos, y en su acepción más elemental expresa la probabilidad de que x sea el mayor (o menor) valor obtenido en n experimentos estadísticos independientes con la variable aleatoria X .

El interés de los técnicos que han venido haciendo uso de la estadística de extremos no se ha centrado en la utilización directa de la ecuación extremal exacta, sino que se ha dirigido hacia el empleo de algunas funciones de distribución que, en ciertas condiciones, aproximan a aquella ecuación de forma asintótica. Estas distribuciones asin-

tóticas han venido siendo estudiadas y desarrolladas desde hace unos setenta años por una serie de autores, entre los cuales R. Fisher y L. Tippett (1928) y M. Fréchet (1927) sentaron las bases de partida y dieron forma a las funciones fundamentales.

La utilidad de las formas asintóticas de $\Phi(x)$ está en que a veces, no se conoce la función de distribución $F(x)$ o/y el parámetro n , pero, en cambio, se dispone de un muestra de valores extremos de la variable. Si se conoce cuál es la forma analítica que toma asintóticamente la distribución extremal, los parámetros de ésta pueden ser estimados mediante ajuste de aquella muestra. Una vez realizada esta estima, se pueden realizar predicciones extremas fuera del rango ocupado por la muestra. Para la determinación del tipo de distribución extremal asintótica que corresponde a la variable en cuestión, es preciso conocer ciertas propiedades de la cola de la función de distribución $F(x)$ y asegurar la conveniencia de las hipótesis formuladas sobre el parámetro n .

En la referencia (20) pueden consultarse los fundamentos y demostraciones de la teoría asintótica; en (12) se da un resumen de las características de las asíntotas y del uso práctico que se viene haciendo de ellas.

El uso directo de la ecuación extremal precisa obviamente de la caracterización de los dos parámetros $F(x)$ y n . Sobre la elección de función de distribución se han indicado ya algunos conceptos básicos. En cuanto al parámetro n , una hipótesis básica implícita en las deducciones de las tres asíntotas es $n = \text{constante}$. Por otra parte, la buena convergencia de las asíntotas con sus correspondientes distribuciones extremas depende de que n tome un valor suficientemente alto; además, naturalmente, de que la cola de la distribución $F(x)$ converja a su vez con suficiente rapidez hacia la forma característica indicada para cada asíntota.

A continuación se va a considerar con algún detenimiento este parámetro n , que apenas ha merecido la atención de la literatura especializada y que, sin embargo, presenta características peculiares que llevan a conclusiones de primera importancia. Una de ellas es que, como se verá en el apartado siguiente, para un amplio grupo de variables geofísicas importantes en ingeniería civil, una de las hipótesis básicas que permitieron la formulación de las tres distribuciones extremas asintóticas no se cumple. Este hecho ha sido ignorado hasta el momento. La consideración del parámetro n lleva, como se expone a continuación, a establecer una *diferenciación de las variables climatológicas en dos tipos básicos con comportamientos característicos distintos*.

El elemental modelo estadístico expuesto anteriormente permite el planteamiento de la ecuación

extremal sobre la base de que la variable se presenta un cierto número n de veces, por término medio en el intervalo de tiempo (generalmente un año) considerado. Se parte, por tanto, de suponer que la variable en cuestión caracteriza un fenómeno que aparece de forma discreta en el tiempo. Pueden citarse variables que de forma natural se adaptan a ese modelo: Una de ellas podría ser, por ejemplo, la cantidad total de nieve caída en una localidad en el paso de cada frente frío sobre esa localidad. El valor correspondiente del parámetro n perteneciente a la distribución extremal en el año medio, sería el número medio de frentes fríos que pasan por el lugar en un año.

Sin embargo, *muchas de las variables geofísicas no son directamente asimilables al modelo descrito*. Como ejemplos característicos pueden mencionarse la humedad, la temperatura o la velocidad instantánea del viento o de una corriente. *No tiene sentido hablar del "número de veces" que una de estas variables se presenta en un año*. La variable presenta una evolución continua a lo largo del tiempo, se define en cada instante y, por tanto, se "presenta" tantas veces como uno quiera considerar, en cualquier intervalo de tiempo. *Puede afirmarse que la gran mayoría de variables geofísicas significativas en ingeniería pertenecen a esta clase, cuya evolución es continua*. Dentro de estas últimas, un grupo particularmente importante es el constituido por magnitudes totalizadas o promediadas en un cierto intervalo de tiempo. Como ejemplos pueden citarse el recorrido del viento (o su velocidad media) en un intervalo estándar como una, seis o veinticuatro horas; el caudal medio (o el volumen total desaguado) de un río en veinticuatro horas; la precipitación total en veinticuatro horas o treinta días; la temperatura media en una hora, etc. *En publicaciones estadísticas que tratan de las distribuciones extremas asintóticas, es habitual considerar que la ecuación extremal [1] es directamente aplicable a estas variables con sólo igualar n al cociente entre el intervalo de definición de la distribución extremal y el intervalo de definición de la variable*. Como ejemplo característico puede citarse la referencia fundamental de E. Gumbel (1958), que al tratar de la distribución extremal anual del volumen desaguado por ríos en veinticuatro horas ("Statistics of Extremes", pág. 237), hace $n = 356$ con la observación (matizada) de que no todas las 365 observaciones anuales son estadísticamente independientes. Sin embargo, *esta forma de tratar a las variables citadas es inadecuada y conduce a resultados erróneos*. Un ejemplo en que se toman dos variables-promedio con intervalos de definición de pequeña duración, ayuda a mostrar cómo se llega a conclusiones absurdas: Consideremos el caudal medio de un río en cinco minutos. Si su función de distribución en el año medio es $F(x)$,

la distribución extremal sería $\Phi(x) = [F(x)]^{105 \cdot 120}$ donde la cifra en el exponente es el número de intervalos de cinco minutos que se yuxtaponen en un año. Si el río es de grandes dimensiones, la función de distribución del caudal medio en diez minutos es prácticamente idéntica a la del caudal medio en cinco minutos ($F(x)$), ya que la curva de variación de estas variables en el tiempo, cuando éste se toma a escala de minutos, es frecuentemente tendida. Por tanto, las distribuciones extremales de ambas variables deben ser también iguales. Sin embargo, la distribución extremal para el caudal del medio en diez minutos, obtenida a partir de la misma hipótesis anterior, sería $\Phi(x) = [F(x)]^{52.5(120)}$, que es sustancialmente diferente a la expresión obtenida para la otra variable. Los períodos de retorno correspondientes a ambos casos difieren en un factor ≈ 2 dada la siguiente relación aproximada, válida para $T(x) > 5-10$:

$$T(x) = \frac{1}{1 - [F(x)]^n} \approx \frac{1}{n [1 - F(x)]}$$

En este caso la hipótesis de independencia estadística entre las observaciones es obviamente inadecuada, pero aun así ambas ecuaciones conservarían asintóticamente su aplicabilidad en virtud de un teorema de G. Watson, citado por E. Gumbel ("Statistics of Extremes", pág. 164), según el cual si la variable es ilimitada la distribución asintótica del valor mayor es la misma en el caso de observaciones interdependientes que para las independientes.

El resultado absurdo al que se ha llegado, indica que no es esta la forma correcta de aplicar la ecuación extremal a las variables totalizadas o promediadas en un intervalo. Una mirada más atenta descubre que este tipo de variables es, en realidad, en todo análogo a las variables de definición instantánea, a que se aludió anteriormente, en cuanto que todas ellas toman un número ilimitado de valores en cualquier intervalo (mayor que el intervalo de definición de la variable), que se desee. Desde este punto de vista, todas ellas son variables de evolución continua en el tiempo. *El caudal medio (o el volumen total desaguado) de un río en veinticuatro horas posee, en efecto, una evolución continua a lo largo del año y el número de valores que toma en un año no es 365, sino infinitud de ellos.* Es cierto que sólo 365 de ellos se yuxtaponen exactamente, pero ello no tiene una relevancia particular: Ni en la propia constitución de la variable, cuyo ritmo de evolución natural no tiene conexión alguna con su intervalo de definición, ni en las aplicaciones ingenieriles, puesto que *quien utilice la distribución extremal está interesado en conocer el "máximo volumen desa-*

guado en veinticuatro horas, en un año" y no solamente "el máximo volumen desaguado en aquellos intervalos de veinticuatro horas que comienzan a las cero horas, en un año". Ambos conceptos pueden diferir grandemente, dando el segundo de ellos valores significativamente inferiores al otro en la gran mayoría de los casos. El segundo de aquellos conceptos, pues, alude a una interpretación de la variable que no solamente está desconectada de la constitución física de ésta, sino que también se aleja peligrosamente de los requerimientos prácticos de la predicción extremal.

Una variable de gran interés en ingeniería marítima, que pertenece a este mismo tipo cuya evolución es continua, es la altura de ola significativa (o cualquier otro promedio de la altura de las olas) que en un cierto instante se halla en una cierta área del mar. Fue precisamente tratando de esta variable cuando por primera vez (según los datos conocidos por el autor del presente trabajo) se mostró que una hipótesis $n = \text{constante}$ produce resultados incoherentes en el análisis extremal. El autor en cuestión es J. Battjes (1970), quien empleó una demostración por reducción al absurdo parecida a la anterior para refutar un análisis extremal de H_s efectuado por N. Nordestrom (1969).

El análisis extremal criticado partió de registros de doce minutos de duración, con los cuales se estimó la probabilidad de excedencia en el año medio, $q(H_s)$. El período de retorno que fue asignado, en años (núm. de minutos al año = 525.600), es:

$$T(H_s) = \frac{12}{525.600 q(H_s)}$$

Battjes observó que podría haberse obtenido la misma función $q(H_s)$ con observaciones mucho más frecuentes, por ejemplo cada minuto, utilizando fotografía aérea o un procedimiento similar. En ese caso, el período de retorno a que se llega con el mismo planteamiento es:

$$T(H_s) = \frac{1}{525.600 q(H_s)}$$

Resultado incompatible con el anterior. Battjes no intentó dar solución al problema en su artículo, cuyo tema central no era precisamente el análisis extremal.

En el apartado siguiente va a tratarse de la forma en que la ecuación extremal [1] puede ser aplicada a las variables de evolución continua en el tiempo.

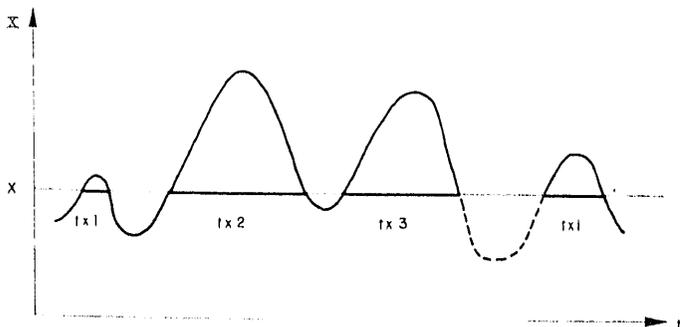
5. LA ECUACION EXTREMAL PARA VARIABLES DE EVOLUCION CONTINUA

La búsqueda de una forma de aplicar la ecuación extremal a las variables de evolución continua, debe dirigirse a encontrar una variable relacionada con ella cuya evolución en el tiempo sea discreta. Sólo con una variable de este tipo el parámetro n tiene sentido, y, por lo tanto, en esas condiciones la ecuación puede ser utilizada. Esta discretización de la variable es sólo útil a condición de que, una vez utilizada la ecuación extremal, se pueda volver de nuevo a la variable continua. A continuación va a utilizarse la curva que describe la variable en el tiempo para realizar ese proceso.

La figura siguiente muestra la evolución de los valores de la variable. Se ha visto cómo el considerar un cierto número de valores correspondientes a un valor fijo del espaciamiento entre ellos conduce a un resultado absurdo. Como solución alternativa, vamos a considerar las ondulaciones que describe la curva de evolución. Cada ondulación constituye un cierto fenómeno físicamente independiente del resto y con una forma propia. Sin embargo, no podemos asignar a cada ondulación entera una determinada duración. En lugar de ello, efectuamos un corte de la curva a un nivel x cualquiera de la variable y consideramos los tramos ondulados que quedan encima de él.

Cada tramo señala un intervalo de tiempo durante el cual la variable toma valores que exceden a x . Podemos denominar a estos tramos "curvas de excedencia del nivel x ", o abreviadamente "excedencias de x ".

Supongamos que las excedencias de x se presentan de forma independiente. La referencia (12) se detiene en esta afirmación para precisar algo más su alcance; por ahora la consideraremos válida, en función de la evidente individualidad física de cada curva de excedencia. Vamos ahora e *establecer, en cada nivel x , una dicotomía constituida por dos posibilidades: la aparición de una excedencia de x , o su no aparición.* Para ello necesi-



tamos hacer abstracción del tiempo y considerar tan solo las "veces" que aparece al año una excedencia de x , y las "veces" que "podría" aparecer, o número de pruebas estadísticas aleatorias.

Para reducir el tiempo durante el cual el nivel x es excedido a las "veces" que esto ocurre, tomamos como duración de cada "vez" o prueba estadística, la duración media $t(x)$ de las excedencias de x . El número de pruebas estadísticas en un año medio es:

$$n(x) = \frac{T_y}{t(x)} \quad [2]$$

donde T_y es el tiempo total del año, expresado en las mismas unidades que $t(x)$.

La probabilidad $q(x)$ de que en una de las $n(x)$ pruebas aparezca una excedencia de x , es:

$$q(x) = \frac{n_x}{n(x)} \quad [3]$$

Donde n_x es el número medio de apariciones de la excedencia de x en un año.

$$n_x = \frac{\sum t_{xi}}{t(x)} \quad [4]$$

Combinando las dos expresiones anteriores:

$$q(x) = \frac{\sum t_{xi}}{T_y} \quad [5]$$

Esta es precisamente la expresión de la probabilidad absoluta de presentación de valores de la variable superiores a x , es decir, el valor complementario de la función de distribución $F(x)$. Con ello hemos vuelto a la variable continua original:

$$q(x) = P(X > x) = 1 - F(x) \quad [6]$$

Ahora podemos llegar a la función de distribución extremal de la variable. La probabilidad de que en el año medio no se presente ninguna curva de excedencia de x , es:

$$[1 - q(x)]^{n(x)} = [F(x)]^{n(x)} \quad [7]$$

Esta es la probabilidad de que el valor x de la variable no sea superado en el año medio, es decir, la función de distribución extremal de la variable:

$$\Phi(x) = [F(x)]^{n(x)} \quad [8]$$

La expresión [8], distribución extremal de una variable de evolución continua, es similar a la [1],

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

distribución extremal de una variable de evolución discreta, excepto que el exponente n es una constante en el caso discreto pero es función de x en el caso continuo. Esta diferencia origina, naturalmente, que las propiedades características de las distribuciones extremales difieran para ambos casos. En particular, las propiedades límites que dieron lugar a las tres distribuciones asintóticas fueron deducidas para $n =$ constante, y, por tanto, la extensiva aplicación que se ha hecho de ellas para variables continuas carece totalmente de base teórica, contra lo que se ha venido suponiendo.

El modelo descrito hasta aquí se presentó ya en (11), donde fue deducido para su uso en análisis extremal de altura de ola significativa.

Para definir la distribución extremal en la ecuación [8] es suficiente disponer de la función de distribución de la variable $F(x)$, y la función de duraciones medias de las excedencias de la variable $n(x)$. En los apartados siguientes se trata de las formas que poseen ambas funciones en

ciertos casos característicos, a partir del análisis de un número de datos referentes a algunas variables geofísicas significativas en ingeniería civil.

6. FUNCION $n(x)$.

Para abordar la caracterización de la forma funcional de $n(x)$ debe primeramente investigarse cuál es el tipo de relación existente entre los valores correspondientes a la población y las estimas obtenidas a partir de muestras limitadas. Dado un determinado período de observación, donde el valor muestral de $n(x)$ para cada nivel de la variable se obtiene a partir de la duración media de las excedencias contenidas en la muestra a ese nivel, las relaciones entre estima muestral y población son:

- La fiabilidad de las estimas muestrales de la duración media es tanto mayor cuanto mayor sea el número de excedencias con-

DATOS DE CAUDAL MEDIO EN VEINTICUATRO HORAS, EMPLEADOS EN LA DETERMINACION DEL PARAMETRO $n(x)$

Nombre estación	Situación	Período observado	Duración de cada registro	Intervalo entre registros	Fuente
Trillo.	Río Tajo.	1 octubre 1963 1 octubre 1973	24 horas	24 horas	Centro de Estudios Hidrográficos (Madrid).
Orusco.	Río Tajuña.	1 octubre 1952 1 octubre 1963			

DATOS DE VELOCIDAD MEDIA DEL VIENTO EN VEINTICUATRO HORAS, EMPLEADOS EN LA DETERMINACION DEL PARAMETRO $n(x)$

Nombre estación	Situación	Período observado	Duración de cada registro	Intervalo entre registros	Fuente
Valladolid.	Valladolid.	1 octubre 1970 1 octubre 1975	24 horas	24 horas y 12 horas	Instituto Meteorológico Nacional (Madrid).
Peinador.	Vigo.	1 febrero 1970 30 septiembre 1972			

DATOS DE PRECIPITACION EN TREINTA DIAS, EMPLEADOS EN LA DETERMINACION DEL PARAMETRO $n(x)$

Nombre estación	Situación	Período observado	Duración de cada registro	Intervalo entre registros	Fuente
Los Llanos.	Cuenca del Guadiana. (Prov. de Cuenca).	1 octubre 1940 1 octubre 1970 (excepto 1966-67)	1 mes	1 mes	Intecsa (Madrid).

CUADRO 1-A

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

DATOS DE ALTURA DE OLA EMPLEADOS EN LA DETERMINACION DEL PARAMETRO $n(x)$

Nombre estación	Situación	Area marítima	Profundidad (m)	Período observado	T. útil (h)	% T. útil	Tipo de observac.	Duración de cada registro	Intervalo entre registros	Fuente
Osborne Head	Nova Scotia (Canadá)	Atl. Norte (W)	30,3	15 diciembre 1970 1 septiembre 1976	37.138	70,1	Instrument. (boya)	20 min.	3 horas	
Western Head	Nova Scotia (Canadá)	Atl. Norte (W)	40,2-43,0	15 abril 1970 5 mayo 1973	19.494	74,2	Instrument. (boya)	20 min.	3 horas	
Chebucto Bay	Nova Scotia (Canadá)	Atl. Norte (W)	26,7	24 octubre 1974 4 febrero 1976	9.091	81,3	Instrument. (boya)	20 min.	3 horas	Don Birell (Marine Information Directorate, Canadá)
Roberts Bank	British Columbia (Canadá)	Estrecho de Georgia	138,9	7 febrero 1974 3 abril 1976	12.282	70,0	Instrument. (boya)	20 min.	3 horas	
Toronto	Toronto (Canadá)	Lago Ontario	107,8	15 abril 1972 18 junio 1973	7.163	81,2	Instrument. (boya)	20 min.	3 horas	
Owers Lightvessel	Owers (Inglaterra)	Canal de la Mancha	13-15	1 octubre 1968 30 septiembre 1969	8.704	99,4	Instrument. (boya)	12 min.	3 horas	L. Draper (I.O.S., Gran Bretaña)
Plat. "Penrod 36"	(53° 24,6' Norte-4° 12,9' Este) (Holanda)	Mar del Norte	26,0	1 marzo 1973 28 febrero 1974	7.150	81,9	Instrument. (boya)		3 horas	E. Bouws (K.N.M.I., Holanda)
Dique P. Lucero	Bilbao	Cantábrico	40,0	2 abril 1976 2 abril 1978	17.335	98,6	Instrument. (boya)	12 min.	3 h - 4 h	L. Tejedor (Thalassa, para el Puerto de Bilbao)
Barco meteor. "B"		Atlántico Norte		1949-1972		91,0	Visual		1 h - 3 h	
Barco meteor. "C"		Atlántico Norte		1952-1972		92,0	Visual		1 h - 3 h	
Barco meteor. "E"		Atlántico Norte		1952-1971		89,0	Visual		1 h - 3 h	
Barco meteor. "I"		Atlántico Norte		1949-1971		88,0	Visual		1 h - 3 h	Environmental Data Service (U.S.A.)
Barco meteor. "J"		Atlántico Norte		1949-1971		88,0	Visual		1 h - 3 h	
Barco meteor. "K"		Atlántico Norte		1949-1969		93,0	Visual		1 h - 3 h	

CUADRO 1-B

tenido de la muestra. Este número (ver figuras 7 y 8, referentes al cuadro 1) es máximo en un cierto nivel de la variable y disminuye hacia niveles superiores e inferiores. Del mismo modo varía la fiabilidad de las estimas de $n(x)$.

- Aceptando que existe una cierta zona "central" de valores de la variable donde las estimas muestrales de $n(x)$ son correctas, los tramos inmediatos por encima y por debajo dan estimas de baja fiabilidad debido al pequeño número de excedencias contenidas en la muestra. Estas estimas se *desvían de forma aleatoria* en torno a los valores de la población.
- Supongamos que la población de la variable en cuestión es ilimitada en ambas direcciones. Cada muestra específica está limitada por un valor máximo de la variable y otro mínimo. En estos niveles, los valores muestrales de $n(x)$ son respectivamente ∞ y 0. Los tramos anteriores de desviación aleatoria terminan en dos tramos que se *desvían sistemáticamente* de los valores correspondientes a la población, tendiendo asintóticamente a ∞ y 0. En el caso de variables cuyo límite natural inferior es cero (como ocurre con H_s), esta desviación sistemática sólo ocurre en el extremo superior muestral.

Por tanto, para cualquiera de estas últimas variables, toda muestra de $n(x)$ con tamaño suficiente posee un tramo central que se ajusta bien a la función correspondiente a la población; dos tramos adyacentes con desviación aleatoria acusada (de los cuales el superior es el más largo, debido a la más lenta disminución del número de excedencias hacia los niveles altos, ver figuras 27 y 28 referentes al cuadro 1), y, por último, un tramo límite superior que se desvía asintóticamente tendiendo hacia $n(x) = \infty$ en el mayor valor muestral.

En la referencia (11), donde se tuvo que abordar la caracterización de $n(H_s)$ sin disponerse de muestra alguna de la curva de evolución de la variable, se admitió una ley exponencial para aquel parámetro con el fin de llegar a una función doble exponencial para la distribución extremal de H_s (función que pareció comportarse adecuadamente en varias muestras extremas obtenidas con la siempre dudosa exactitud de las estimas visuales individuales en temporales). Posteriormente han sido analizados varios años de registro casi-continuo de algunas variables significativas en ingeniería civil (altura de ola significativa, velocidad media del viento en veinticuatro horas, caudal fluvial en veinticuatro horas, precipitación en treinta días) con el fin de estimar directamente una forma funcional adecuada para $n(x)$ en estos casos. En

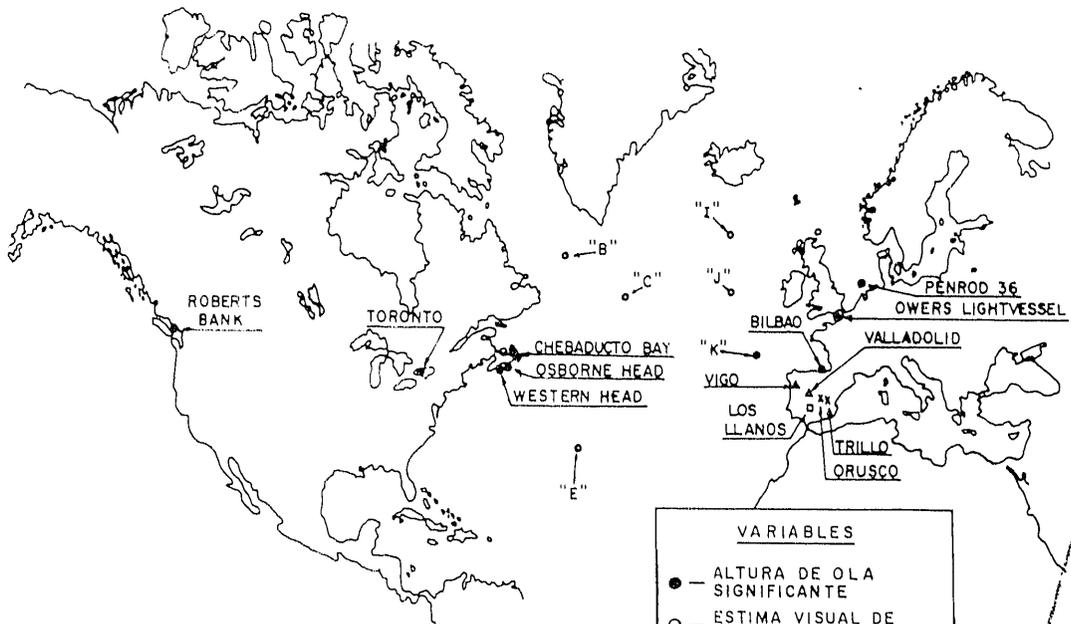
el cuadro 1 y figura 9 se muestran las características de las muestras utilizadas. El resultado obtenido del análisis es que la sencilla función lineal, $n(x) = Ax + B$, se comporta correctamente en todo el conjunto de muestras analizado. En las figuras 10 a 19 se muestran los puntos muestrales con las funciones lineales ajustadas a sus tramos centrales. En estas figuras no se han representado los puntos para cuya estima se contaba con un número de excedencias inferior a 10, con lo cual se han limitado considerablemente las desviaciones. En los tramos superiores de desviación aleatoria (los más largos), esa desviación es hacia la derecha de la recta en un mayor número de casos que hacia la izquierda. Esto es coherente con el hecho de que las funciones de distribución de la duración de las excedencias para cada nivel de la variable (de la que se han analizado un cierto número de casos), son apreciablemente sesgadas en el sentido de que las duraciones inferiores a la media resultan más frecuentes que las superiores a ella (a cambio, naturalmente, de tener un recorrido inferior).

El ajuste correcto de las funciones lineales para $n(x)$ requiere el uso de "intervalos de precisión", tal como se describe en la referencia (12). Sin embargo, el ajuste visual al tramo de valores centrales resulta aceptablemente aproximado para el uso de la ecuación extremal, siempre que la *longitud efectiva* de la muestra (longitud del tramo central de estima correcta) sea suficiente. Esto último sólo puede determinarse convenientemente tras análisis de muestras de cada variable objeto de interés. En la referencia (12) se han obtenido algunos resultados interesantes respecto a las variables H_s y V .

Como puede observarse, no se ha detectado dependencia de la tipología climática en la caracterización funcional de $n(x)$ en los casos estudiados. Sólo el análisis de un mayor número de muestras de cada variable permitirá evaluar hasta qué punto es esto correcto; mientras tanto, los resultados obtenidos en (12) relativos a la baja sensibilidad de $n(x)$ a variaciones apreciables de $n(x)$ llevan a estimar que la función lineal puede ser utilizada con confianza en la práctica.

Entre las figuras anteriores, las referentes a observaciones visuales de altura de ola no tienen un valor elevado en la selección de la relación lineal para $n(H_s)$, ya que la precisión de aquellas observaciones es menos clara de lo deseable en la región de valores altos de la altura de ola. Se las incluye, sin embargo, como comprobación "a posteriori" de que la función lineal produce ajustes razonables y de valor práctico para estas estimas de oleaje realizadas un tanto rudimentariamente, pero que constituyen una fuente de datos de primera importancia (a menudo la única disponible).

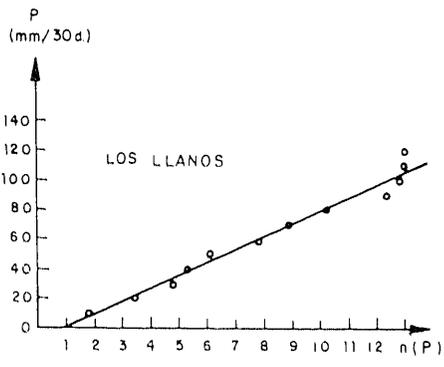
FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO



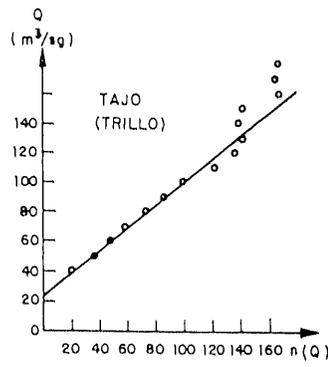
(FIG.7) DETERMINACION DEL PARAMETRO $n(x)$: SITUACION DE LAS ESTACIONES

- VARIABLES**
- - ALTURA DE OLA SIGNIFICANTE
 - - ESTIMA VISUAL DE ALTURA DE OLA
 - X - CAUDAL FLUVIAL
 - △ - VELOCIDAD DEL VIENTO
 - - PRECIPITACION

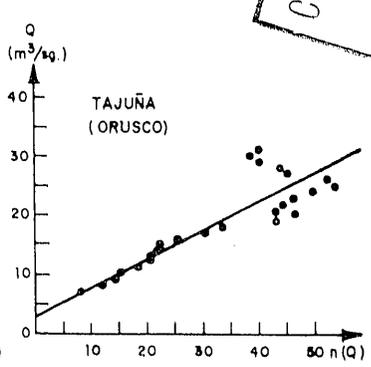
COLEGIUM INSTITUTUM ECOMINOS
BILALITECA



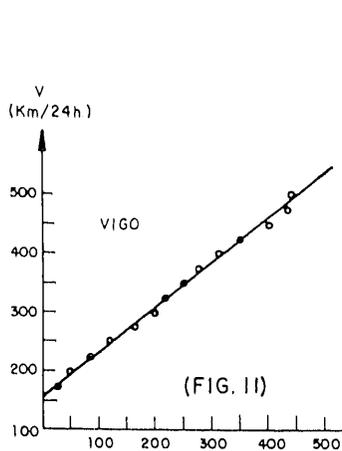
(FIG. 8)



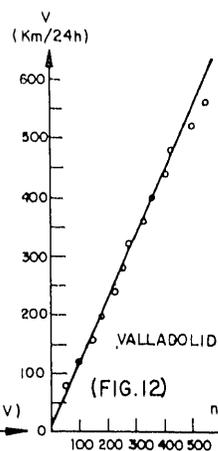
(FIG. 9)



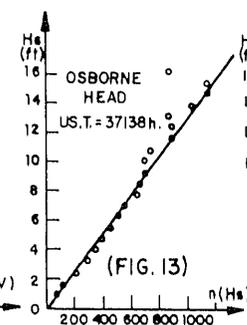
(FIG. 10)



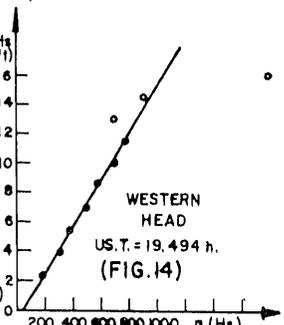
(FIG. 11)



(FIG. 12)

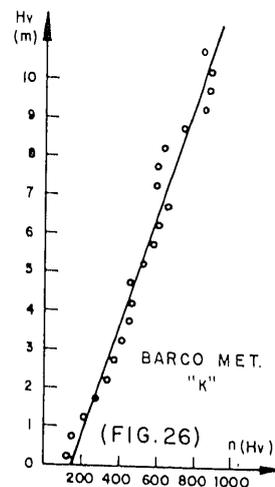
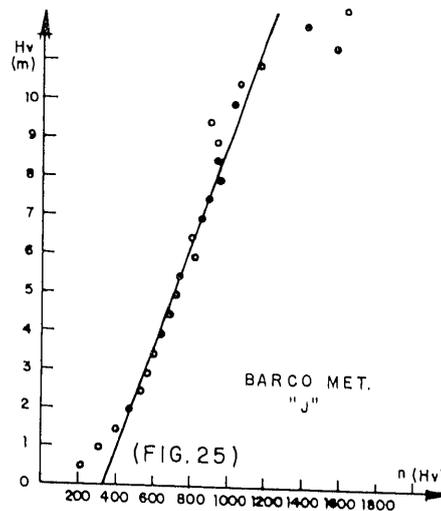
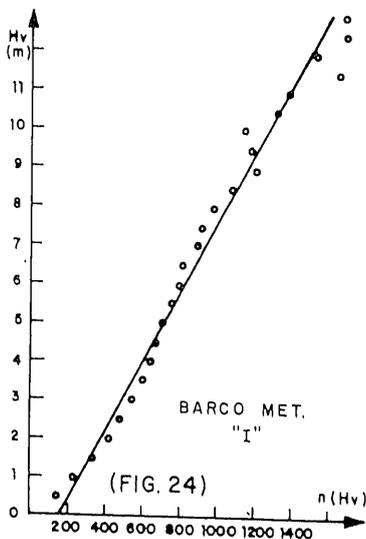
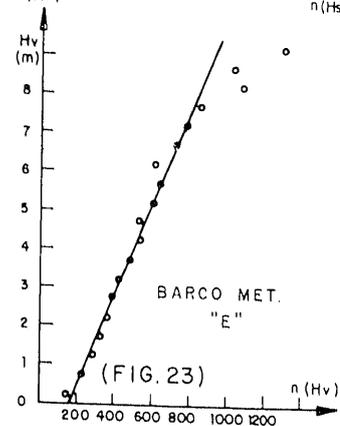
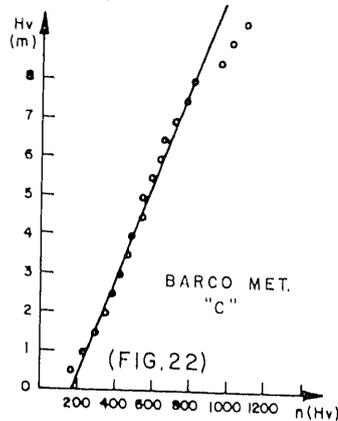
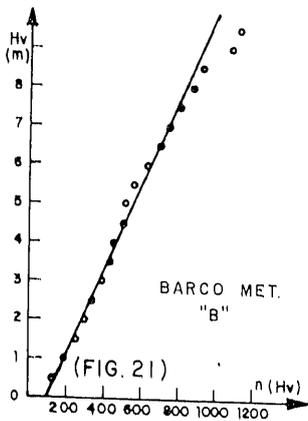
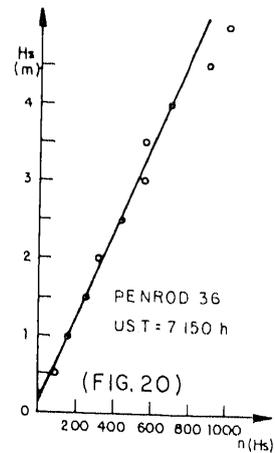
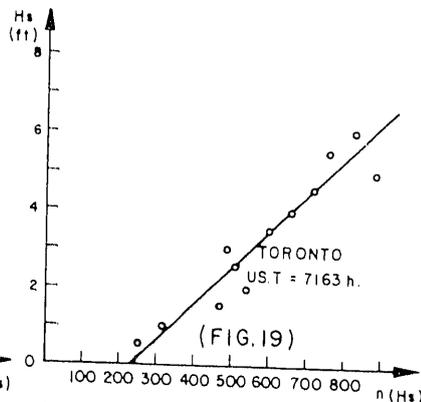
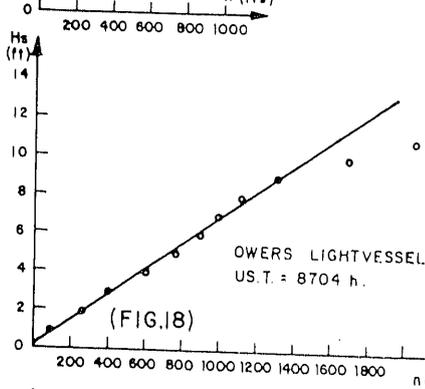
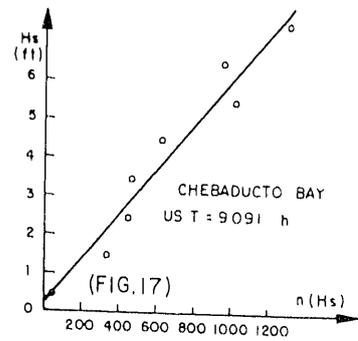
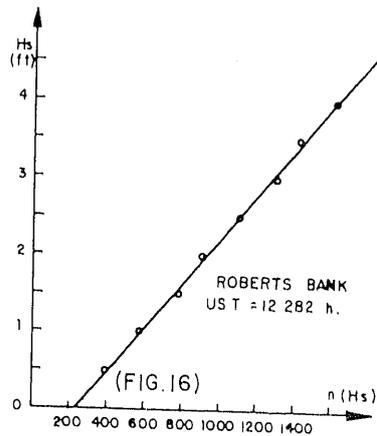
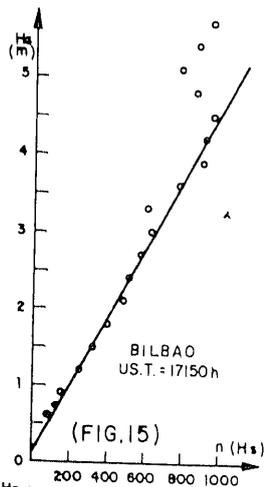


(FIG. 13)



(FIG. 14)

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO



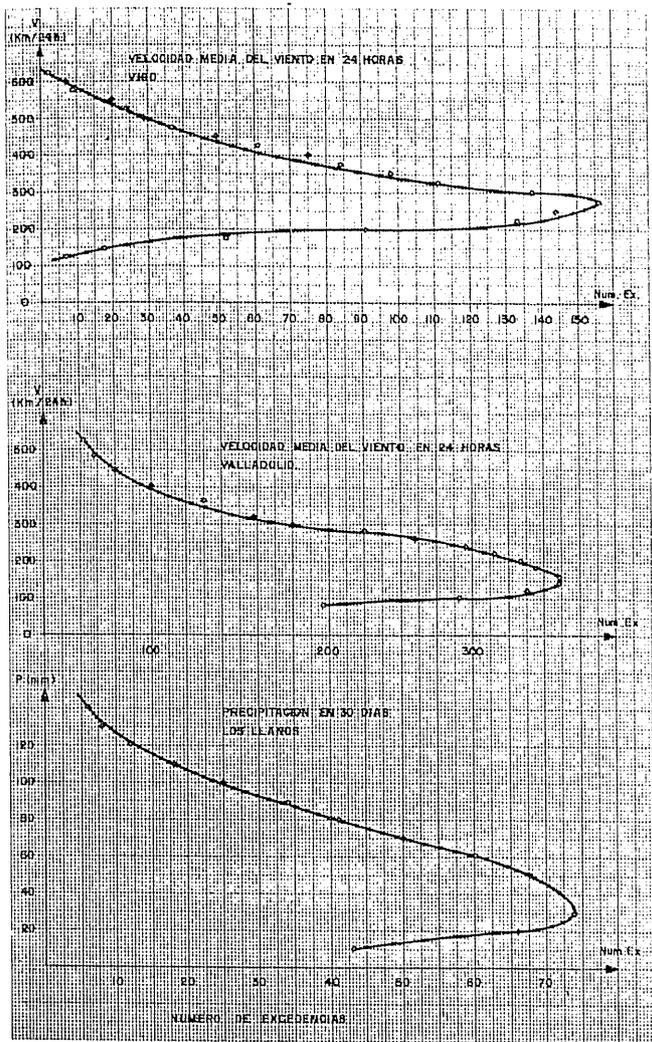


Figura 27.

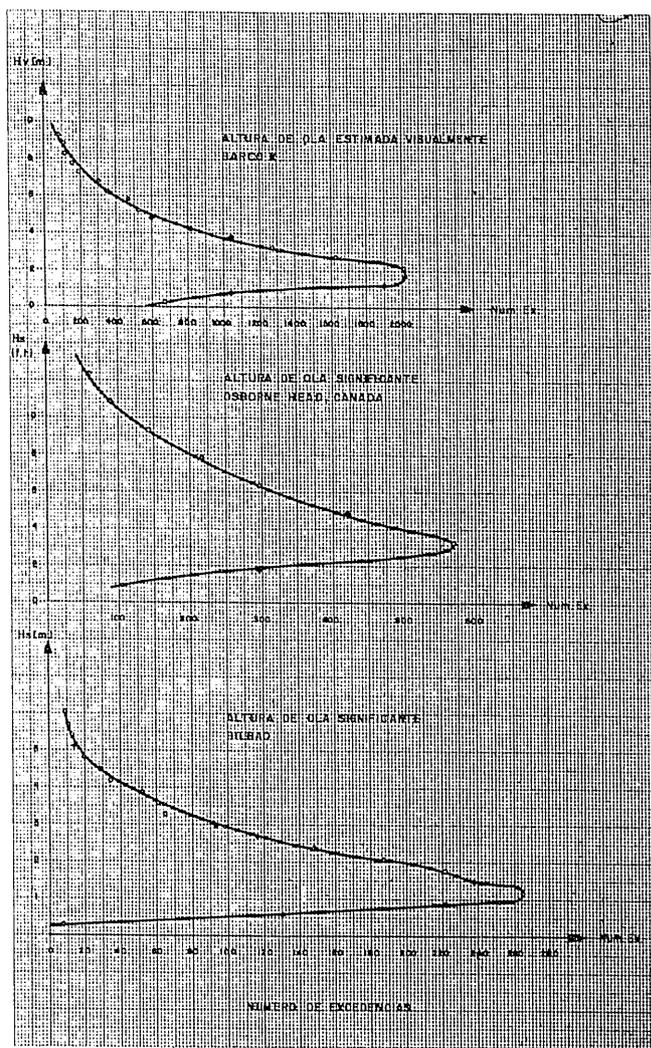


Figura 28

7. FUNCION $F(x)$.

Anteriormente se ha descrito la relación que existe entre los valores de $F(x)$ correspondientes a la población, y las estimas obtenidas a partir de muestras limitadas, en el caso de muestreo aleatorio. Para muestreo continuo o casi-continuo, la situación es similar incluyendo además tramos de desviación sistemática en los extremos muestrales (12 y 14). Según se indicó en aquel apartado anterior, la determinación de función de distribución para cada variable necesita del análisis comparado de un elevado número de muestras diferentes (en la referencia (12) se realiza esta determinación para la variable altura de ola significativa, con ayuda de 20 muestras distintas pertenecientes a diversas áreas marítimas).

Sin embargo, una sola muestra de buen tamaño puede servir de base a una determinación efectiva de $F(x)$ si se dispone al mismo tiempo de una muestra extremal cuyo recorrido útil tenga una

longitud apreciable, aunque esta sola no permita, como se indicó antes, la elección directa de distribución extremal. Para ello hay que partir de suponer que el tipo de función adecuado para $n(x)$ es conocido: En el apartado anterior se vio que $n(x)$ puede admitirse lineal, al menos para varias variables importantes. Como ejemplo van a tomarse dos casos: La velocidad media del viento en veinticuatro horas registrado en Valladolid (1970-1975) y la precipitación total en treinta días registrada en Los Llanos, Cuenca (1940-1969), que son dos de las estaciones que figuran en la tabla 1. En las figuras 29 a 31 se muestran ajustes de las distribuciones observadas para ambas estaciones, por funciones de distribución Weibull, Log-Normal de tres parámetros, y Doble Exponencial. Con estas tres hipótesis para $F(x)$, y las funciones $n(x)$ que fueron determinadas en las figuras 8 y 12, se han calculado las distribuciones extremales correspondientes, que se han representado en las figuras 32 a 34 junto con dos muestras extremales de ambas

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

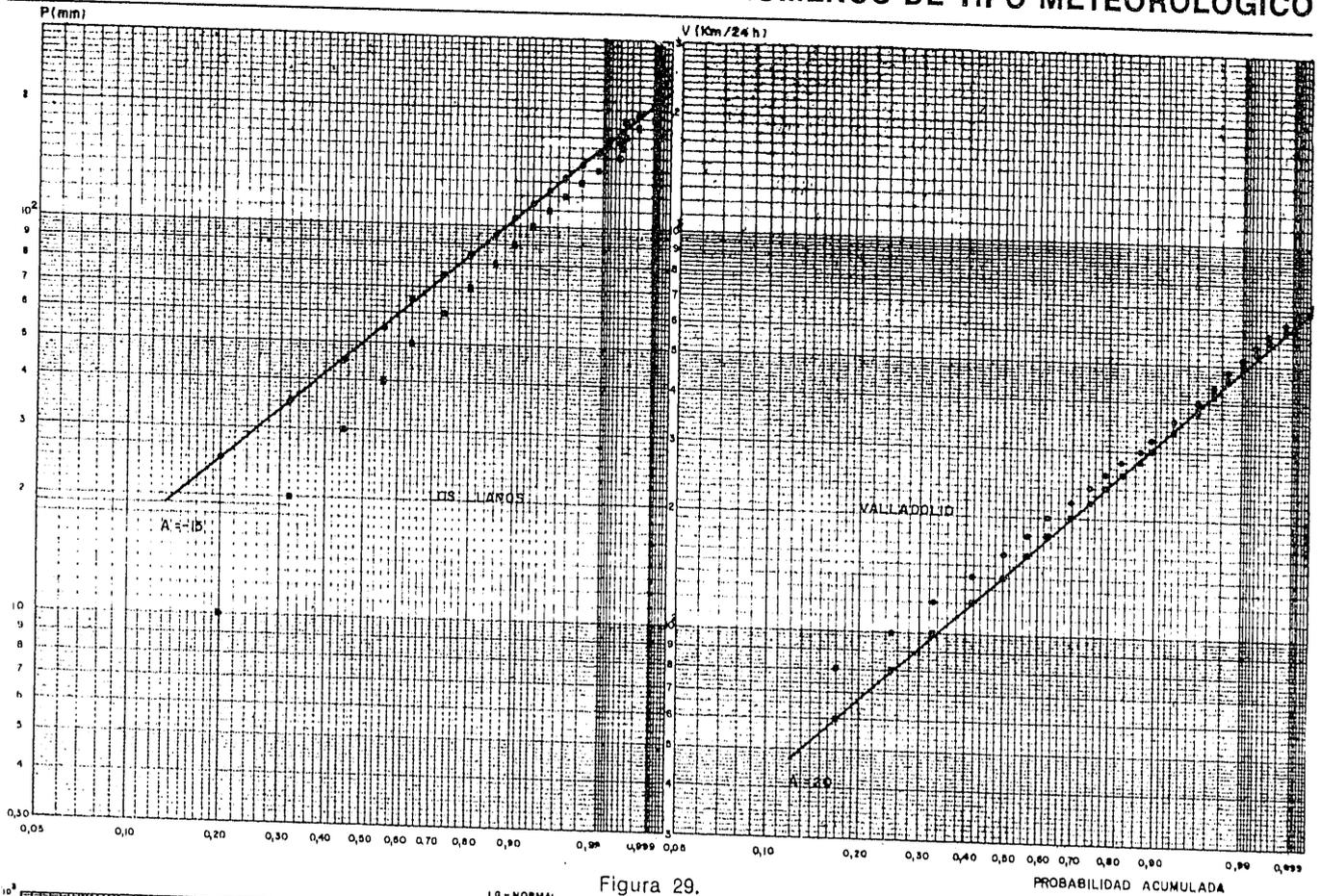


Figura 29.

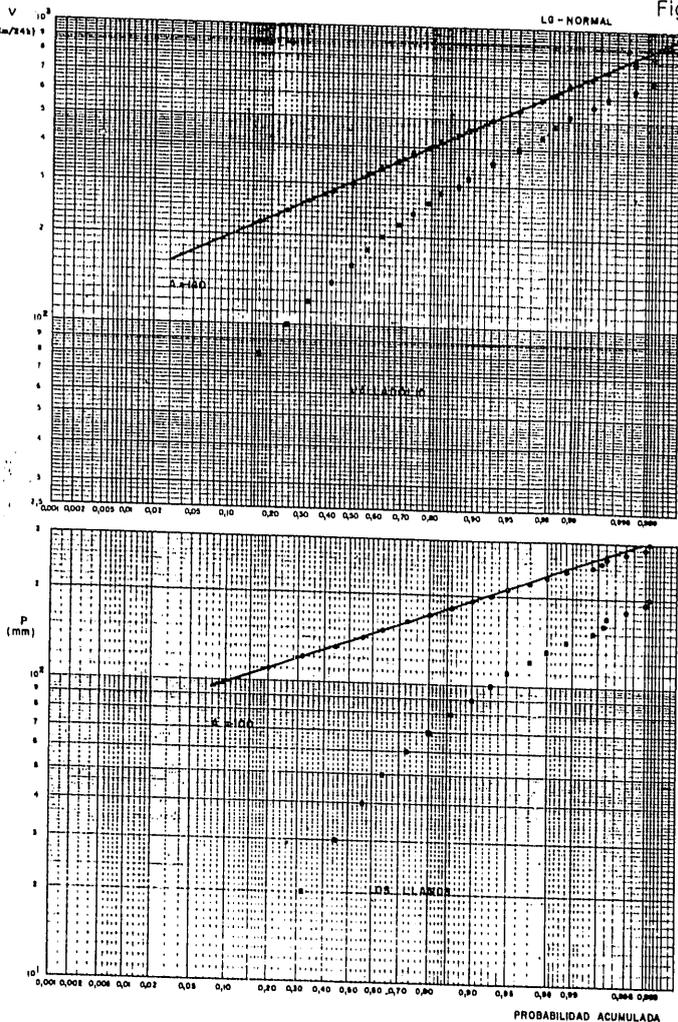


Figura 30.

estaciones. Estas últimas han sido representadas según tres criterios distintos: frecuencias observadas (12 y 14); fórmula de Weibull-Gumbel; fórmula de Gringorten; sin establecer intervalos de precisión y, por tanto, sin delimitar un recorrido útil.

De la inspección de las figuras se desprende que la función Doble Exponencial es, en ambos casos, adecuada para $F(x)$. Al mismo tiempo, se han determinado las distribuciones extremales $\Phi(x)$ correspondientes a las dos estaciones. Ninguna de estas funciones $F(x)$ y $\Phi(x)$ hubiera podido ser determinada directamente de forma individual, a partir de sus muestras respectivas, a no ser que por otro lado se conociese ya previamente cuáles eran los tipos de función adecuados para ambas distribuciones.

8. RESUMEN

El cuadro 2 esquematiza los sistemas de estimación de distribución extremal que han sido discutidos en este artículo.

Han quedado por tratar algunos aspectos complementarios de esta problemática, que son importantes para utilizar el análisis extremal con el ne-

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

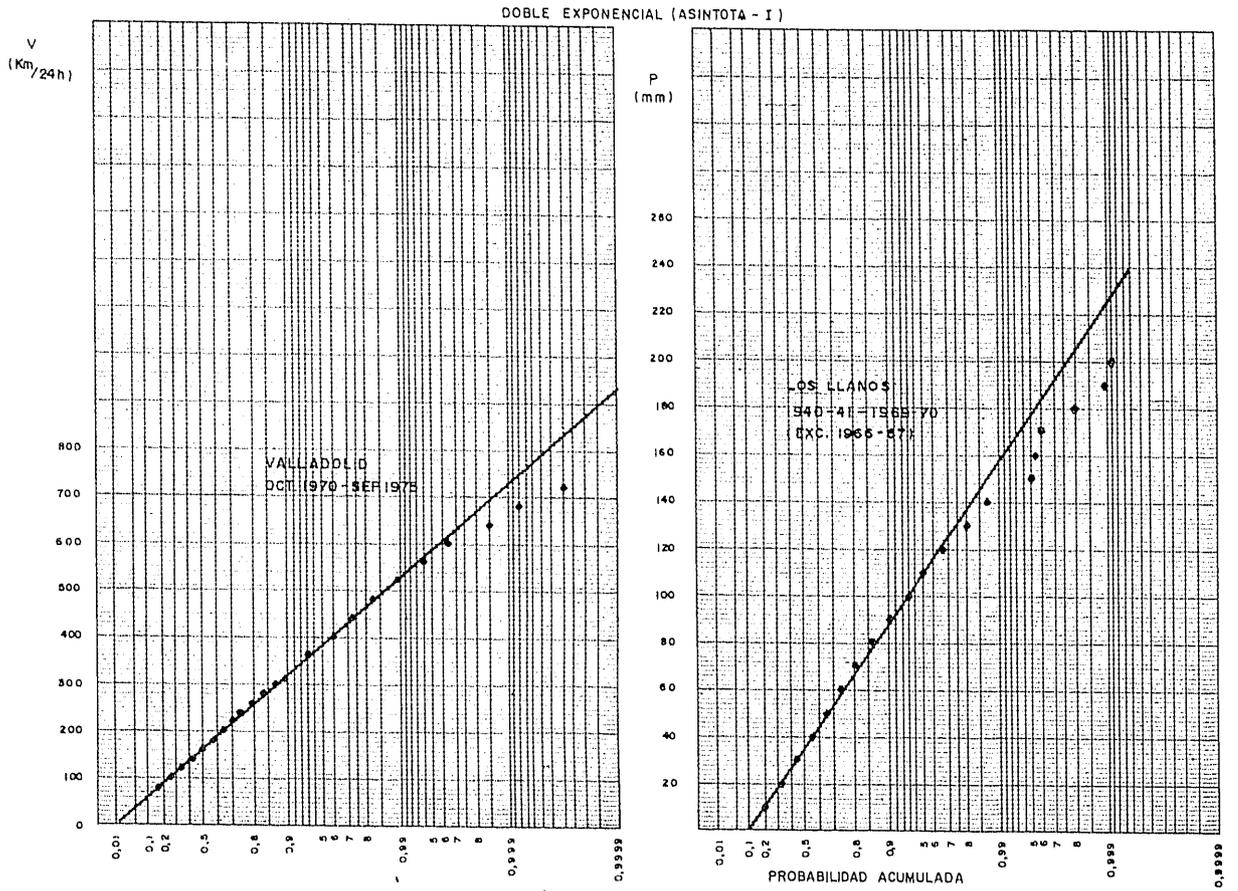


Figura 31.

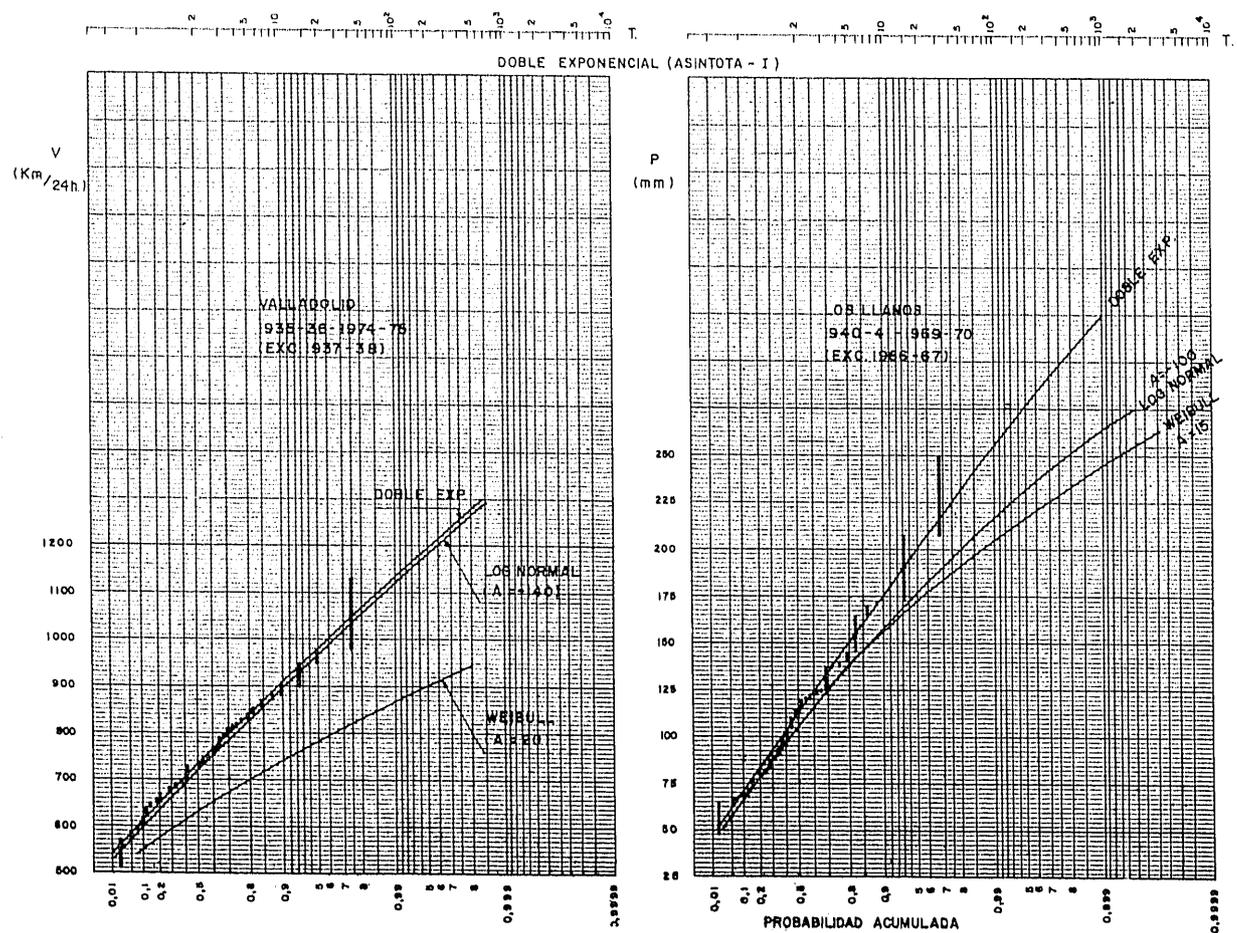


Figura 32.

MUESTRAS EXTREMALES (FRECUENCIAS OBSERVADAS) COMPARADAS CON CURVAS DE PREDICION

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

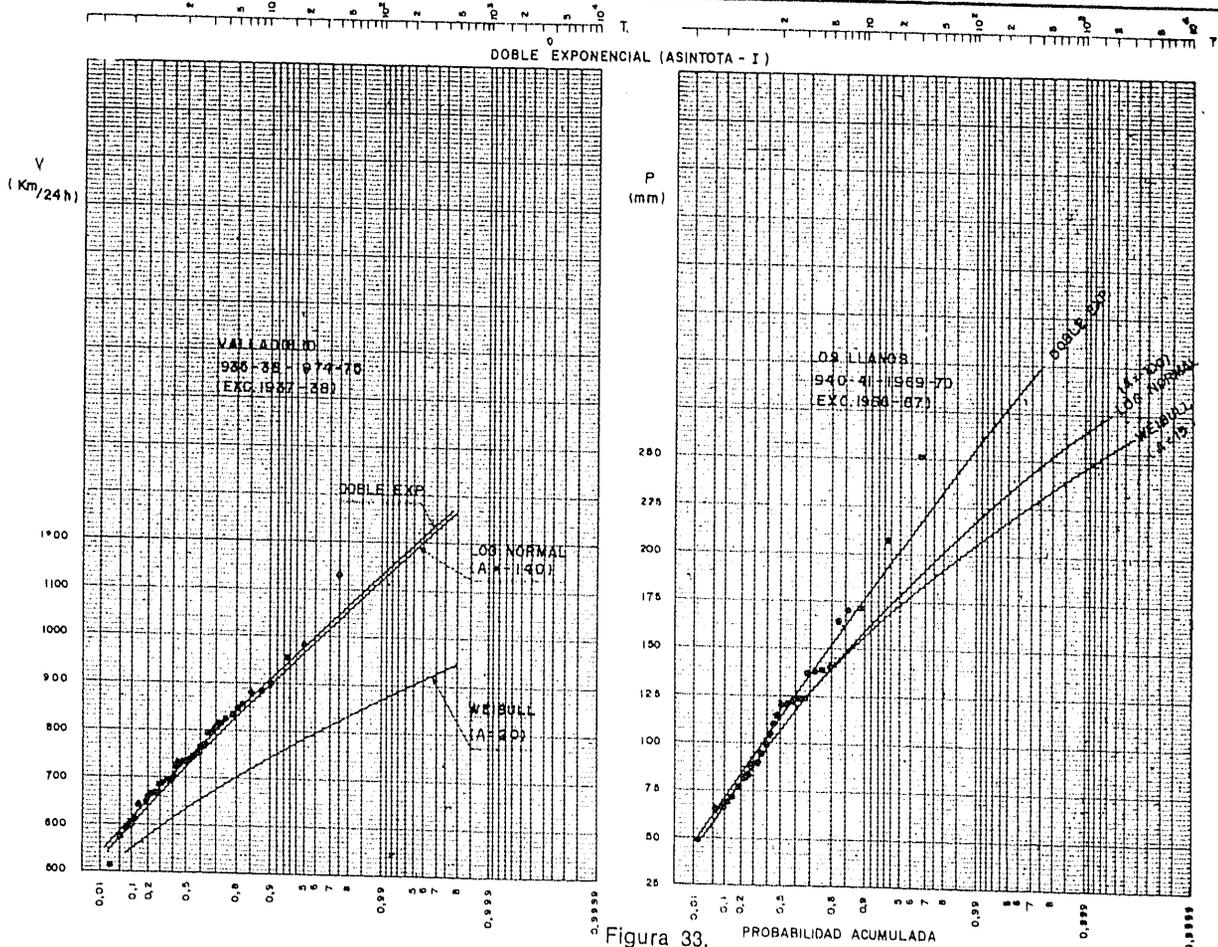


Figura 33.

MUESTRAS EXTREMALES (FRECUENCIA DE REPRESENTACION DE WEIBULL - QUMBE.)
COMPARADAS CON CURVAS DE PREDICION

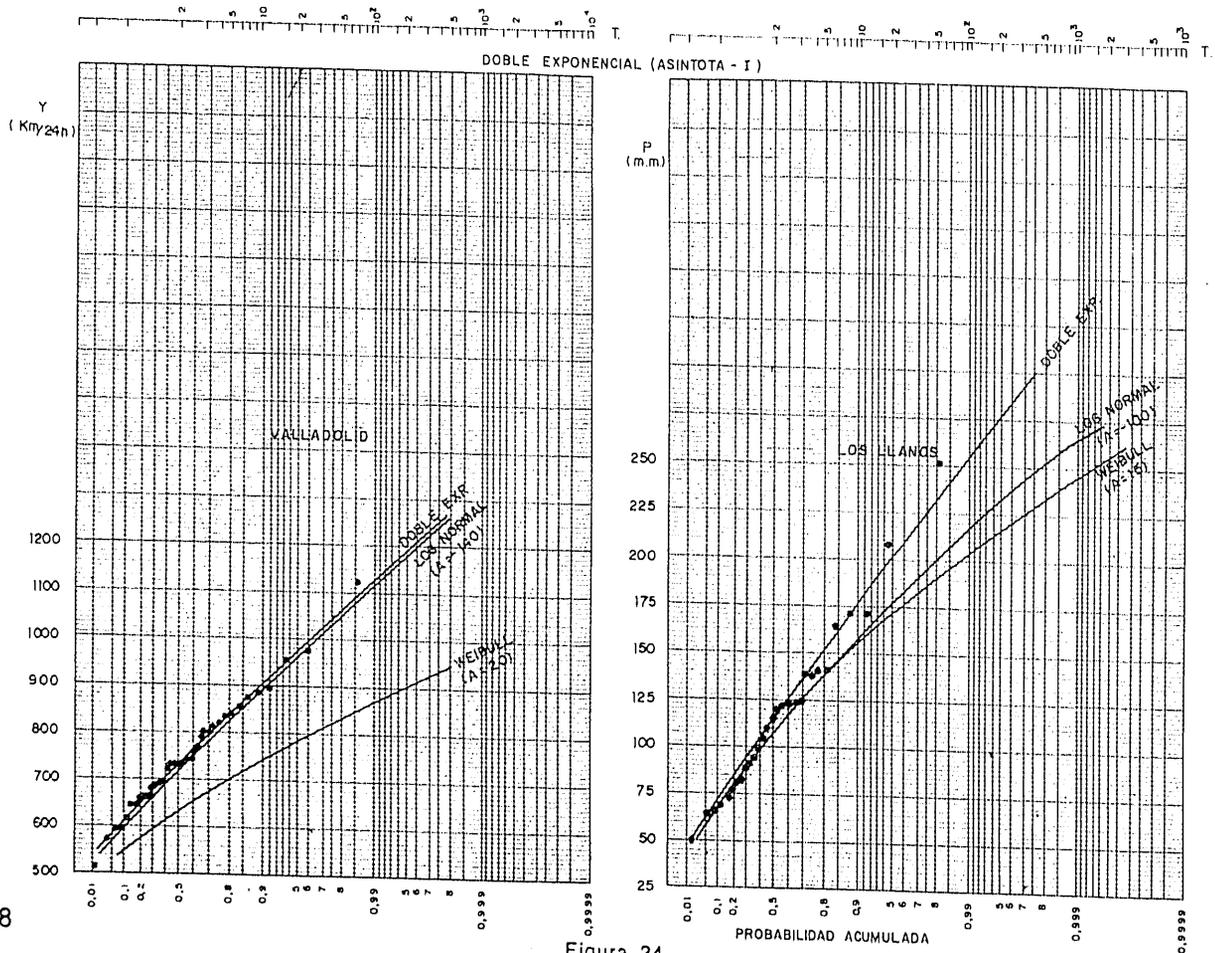


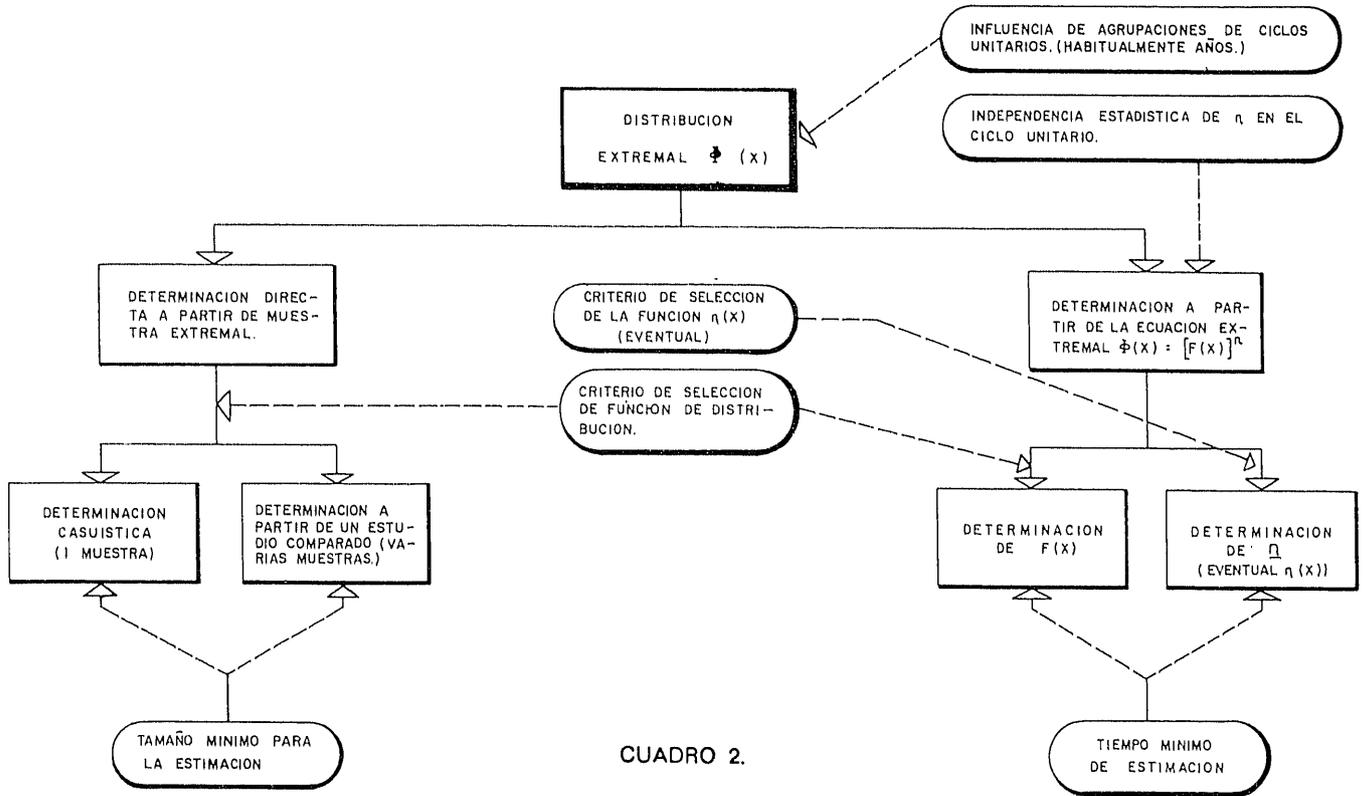
Figura 34.

MUESTRAS EXTREMALES (FRECUENCIA DE REPRESENTACION DE GRINGORTEN) COMPARADAS CON
CURVAS DE PREDICION

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

cesario conocimiento de las limitaciones y posibilidades de la metodología a nivel de su empleo en la práctica: Por un lado, los tiempos mínimos de observación o tamaños muestrales mínimos para una estima correcta de las funciones utilizadas.

Por otro, las repercusiones de la mayor o menor corrección de la hipótesis de independencia estadística para n dentro del ciclo unitario escogido y para los propios ciclos unitarios entre sí. Estos temas se discuten en (12).



CUADRO 2.

9. AGRADECIMIENTOS

En la obtención y tratamiento de los datos empleados para la caracterización de varias funciones se ha contado con la amable colaboración de un número de personas e instituciones enumeradas en (12). Entre ellas ha resultado particularmente importante la colaboración de Intecsa (Madrid), que por medio de Victoriano Fernández Dupuy proporcionó uso libre de su ordenador IBM-370-138 (VM) y la asistencia de sus especialistas en programación para elaborar los programas de ordenador necesarios.

El presente artículo constituye una parte de la tesis doctoral presentada por el autor a la E. T. S. de Ingenieros de Caminos, Canales y Puertos de Santander, y publicada en (12). Parte del mismo fue también parte de una comunicación presentada (13) a la XVI Conferencia Internacional de Ingeniería de Costas (I.C.C.E.; Hamburgo, 1978).

REFERENCIAS

- (1) N. BARRICELLI: "Les plus grands et les plus petits maxima ou minima annuels d'une variable climatique". Arkif for Matematik og Natur., 46-6. Oslo, 1943.
- (2) J. BATTJES: "Long term wave height distribution at seven stations around the British Isles". N.I.O. Internal Report No. A-44, 1970.
- (3) J. BELL: "Surface Winds in Hong Kong Typhoons". U.S. Asian Weather Symp. Feb. 1961.
- (4) A. BENHAM: "The estimation of extreme flood discharges by statistical methods". Proc. New Zealand Inst. Civ. Eng., 36, 1950.
- (5) M. BENSON: "Uniform Flood-Frequency Estimating Methods for Federal Agencies". Jour. Water Resources Research, Vol. 4, No. 5, 1968.
- (6) J. BERNIER: "Sur l'application des diverses lois limites des valeurs extrêmes au problème des débits de crue". La houille Banche, No. 5, 1956.
- (7) J. BERNIER: "Portée et limites des méthodes de prévision statistique des crues utilisées dans le calcul éco-

FUNCION DE DISTRIBUCION EXTREMAL EN FENOMENOS DE TIPO METEOROLOGICO

- nomique des ouvrages de protection contre les crues". Bull. Centre de Rech. et d'Essays de Chatou, 6 December 1963.
- (8) B. BOBEE and R. ROBITAILLE: "The Use of the Pearson Type 3 and Log-Pearson Type 3 Distributions Revisited". Jour. Water Research, 13-2, April 1977.
 - (9) L. BROWN, M. NEKHOM and M. PATCH: "A Monte-Carlo Determination of Extreme Ocean Waves". Offshore Techn. Conf., Houston, May 1975.
 - (10) R. CARTER: "Floods in Georgia, Frequency and Magnitude". U.S. Geol Survey, Circ. No. 100, 1951.
 - (11) E. COPEIRO: "Un método práctico de estimar oleaje extremal para el cálculo de estructuras marítimas". Revista de Obras Públicas. Junio 1976.
 - (12) E. COPEIRO: "Análisis extremal de variables geofísicas". Publicación 14, Laboratorio de Puertos, C.E.E.O.P. Madrid, 1978.
 - (13) E. COPEIRO: "Extremal prediction of significant wave height". 16th Int. Conf. Coastal Eng., Hamburgo, 1978.
 - (14) E. COPEIRO: "Estima de función de distribución a partir de una muestra aleatoria". Revista de Obras Públicas. Marzo 1979.
 - (15) A. COURT: "Wind Extremes as Design Factors". J. Franklin Institute, 256, 39, 1953.
 - (16) A. DAVENPORT: "Wind Loads on Structures". Tech. Paper No. 88, Division of Building Res., Ottawa, 1960.
 - (17) M. St. DENIS: "Determination of Extreme Waves". Topics in Ocean Engineering, Gulf Publ. Comp., 1969.
 - (18) M. St. DENIS: "Some Cautions on the Employment of the Spectral Technique to Describe the Waves of the Sea and the Response Thereto of Oceanic Systems". Preprints, Offsh. Techn. Conf., Houston, 1973.
 - (19) E. GUMBEL: "Les Crues du Rhône". Ann. de l'Université de Lyon (serie 3), No. 3, sect. A, 1940.
 - (20) E. GUMBEL: "Statistics of Extremes". New York, Columbia Un. Press, 1958.
 - (21) A. JENKINSON: "The frequency distribution of the annual maximum (or minimum) values of meteorological elements". Q. J. Royal Met. Soc., 87, 1955.
 - (22) S. NAG and N. DUTTA: "Determination of maximum likely one, two, three, four and five day rainfall in a river catchment". Bull. Internat. Inst., Calcutta, 33, 1951.
 - (23) N. NORDESTROM: "Long term distribution of wave heights and periods". Rep. No. 69-21-5, Det Norske Veritas, Oslo, 1969.
 - (24) C. PETRAUSKAS and P. AAGAARD: "Extrapolation of Historical Storm Data for Estimating Design Wave Heights". Offshore Techn. Conf. Houston, April 1970.
 - (25) S. RANTZ and H. RIGGS: "Magnitude and Frequency of Floods in the Columbia River Basin". U.S. Geol. Survey, Water Supply Paper, 1808, 1949.
 - (26) SHUH: "The Return Period and Variability of Floods in China". Essays and Papers in Memory of the late President Fu Ssu-nien, Taipan, 1952.
 - (27) P. SUAREZ BORES: "Sea observation in coastal areas: The Spanish offshore network". Int. Symp. Wave Measurement and Analysis, New Orleans, 1974.
 - (28) H. THOM: "Frequency of maximum wind speeds". Proc. A.S.C.E., 80, Sep. 1954.
 - (29) H. THOM: "The Distribution of Annual Tropical Cyclone Frequency". Journ. Geophys. Res., 65-1, Jan. 1960.
 - (30) H. THOM: "Tornado probabilities". Monthly Weather Review, 91-10, 12 Oct.-Dec. 1963.
 - (31) H. THOM: "Toward a Universal Climatological Extreme Wind Distribution". Nat. Res. Council. Proc., Int. Res. Sem., Wind Effects on Buildings and Structures, Ottawa, 1968.
 - (32) H. THOM: "New Distributions of Extreme Winds in the United States". Journ. Struct. Div., A.S.C.E., 94, July 1968.
 - (33) H. THOM: "Asymptotic Extreme-Value Distributions of Wave Heights in the Open Ocean". Proc. A.S.C.E., Volume 99, No. WW-3, Aug. 1973.
 - (34) H. THOM: "Distributions of Extreme Winds over Oceans" Journ. Waterways and Harbors Div., A.S.C.E., Feb. 1973.
 - (35) H. WITTINGHAM: "Extreme Wind Gusts in Australia". Commonw. of Australia, Bur. of Meteor., Bull. No. 46, 1964.