

El mapa de caudales máximos de las cuencas intercomunitarias

Map of maximum flows of intercommunity basins

Antonio Jiménez Álvarez. Ingeniero de Caminos, Canales y Puertos.

Jefe del Área de Recursos Hídricos. Centro de Estudios Hidrográficos (CEDEX). Madrid (España). antonio.jimenez@cedex.es

Celia García Montañés. Ingeniero de Caminos, Canales y Puertos.

Directora de Programa. Centro de Estudios Hidrográficos (CEDEX). Madrid (España). celia.garcía@cedex.es

Luis Mediero Orduña. Dr. Ingeniero de Caminos, Canales y Puertos.

Personal de Investigación. Centro de Estudios Hidrográficos (CEDEX).

Profesor en la ETS de Ingenieros de Caminos, C y P. Universidad Politécnica de Madrid. Madrid (España). luis.mediero@upm.es

Lara Incio Caballero. Licenciada en Ciencias Biológicas.

Personal de Investigación. Centro de Estudios Hidrográficos (CEDEX). Madrid (España). lara.incio@cedex.es

Julio Garrote Revilla. Licenciado en Ciencias Geológicas.

Personal de Investigación. Centro de Estudios Hidrográficos (CEDEX). Madrid (España). julio.garrote@cedex.es

Resumen: Dentro del marco de un Convenio, la Dirección General del Agua encargó al CEDEX la elaboración de un mapa de caudales máximos en régimen natural en las cuencas intercomunitarias, trabajo que ha sido distribuido recientemente. Se presentan en este artículo las bases teóricas del trabajo que ha consistido en el estudio detallado de las leyes de frecuencia de caudales máximos en una selección de estaciones de aforos y embalses, aplicando técnicas de regionalización e incorporando datos sobre eventos históricos, y la posterior extrapolación de los resultados a puntos no aforados mediante modelos hidrometeorológicos y estadísticos. Se ha desarrollado una herramienta informática (CauMax) que permite la visualización y consulta de las capas raster con los resultados en el entorno de un SIG.

Palabras Clave: Crecidas; Caudales máximos; Cuencas intercomunitarias; Método racional; Regionalización estadística

Abstract: In a collaboration agreement, the Spanish Water Directorate entrusted the Center for Research and Experimentation of Public Works, CEDEX, with the development of a map of maximum natural flows of intercommunity basins which has recently been published. This article presents the theoretical basis of the work, this consisting of a detailed study of the distribution of maximum flows at selected gauging stations and reservoirs, applying regionalization techniques and the incorporation of data of historical events. The results have been extrapolated to ungauged sites using hydrological and statistical models. A computer program (CauMax) has been developed to view and consult raster layers in a GIS environment.

Keywords: Floods; Maximum flows; Intercommunity basins; Rational method; Statistical regionalization

1. Introducción

El fenómeno natural de la llegada de avenidas a un río es enormemente variable en frecuencia y magnitud, estando en esencia determinado por el azar. La caracterización y cuantificación de las crecidas de un río es indispensable para garantizar el correcto funcionamiento y la seguridad de determinadas infraestructuras hidráulicas y de transporte. Asimismo, el estudio hidráulico de los caudales de crecida permite conocer sus efectos sobre el territorio y, en consecuencia, caracterizar su peligrosidad

1. Introduction

The natural phenomenon of river flooding varies enormously in terms of frequency and scale and is essentially determined by chance. The characterization and quantification of river flooding is essential to guarantee the correct operation and safety of specific water infrastructure and transport. The hydraulic study of flood flows makes it possible to ascertain their effect on the territory and, in turn, characterise the ensuing hazard and risk of flooding, this being essential information for the

y su riesgo por inundación, información imprescindible para gestionar correctamente dicho riesgo y para la ordenación de los usos del suelo. Este último aspecto cobra un especial interés hoy en día dentro del proceso de implantación de la Directiva para la Evaluación y Gestión del Riesgo de Inundación y del desarrollo del Sistema Nacional de Cartografía de Zonas Inundables (SNCZI).

De esta forma, la Dirección General del Agua (DGA) encargó al Centro de Estudios y Experimentación de Obras Públicas (CEDEX) la caracterización de los caudales de crecida, en el ámbito territorial de las demarcaciones hidrográficas con cuencas intercomunitarias, mediante la elaboración de un mapa de caudales máximos. Para el establecimiento de la metodología a aplicar se seleccionó la cuenca del Tajo como cuenca piloto, cuyos resultados fueron presentados y distribuidos en el año 2009. Recientemente se ha finalizado el trabajo en el resto de demarcaciones hidrográficas con cuencas intercomunitarias (a excepción de la demarcación del Guadalquivir, que se encuentra en elaboración, y para la que en la actualidad únicamente se dispone de resultados en el territorio de las Comunidades Autónomas de Extremadura y Castilla-La Mancha), habiéndose puesto sus resultados a disposición de las distintas administraciones y del público en general. El objeto de este artículo es presentar las bases teóricas del trabajo realizado por el CEDEX para generar el mapa de caudales máximos.

La metodología aplicada en el estudio ha partido de la selección de estaciones de aforos y embalses con suficiente información sobre caudales máximos como para realizar una estimación suficientemente fiable de los valores correspondientes a distintas probabilidades o periodos de retorno (cuantiles). A las series de datos seleccionadas se les han aplicado diversas técnicas estadísticas para verificar la calidad de la información y se les ha ajustado una función de distribución con la finalidad de determinar su ley de frecuencia, aplicando técnicas de regionalización estadística e incorporando datos sobre eventos de avenida históricos para mejorar la estimación de los cuantiles de alto periodo de retorno. La información sobre cuantiles obtenida en los puntos aforados se ha extrapolado al resto de puntos de la red de drenaje mediante la aplicación de métodos hidrometeorológicos y modelos estadísticos en función de la superficie de cuenca vertiente.

correct management of this risk and for the organization of land use management. This latter aspect takes on particular importance today in view of the introduction of the EU Directive on the assessment and management of flood risks and the development of the National System of Flood Zone Cartography (SNCZI).

The Spanish Water Directorate (DGA) entrusted the Center for Research and Experimentation of Public Works (CEDEX) with the characterization of flood flows within the territorial area of intercommunity river basin districts through the preparation of maximum flow maps. In order to establish the methodology to be applied, the Tajo basin was selected as a pilot basin and the results were duly presented and published in 2009. The work on all other intercommunity river basin districts was recently completed (with the exception of the Guadalquivir basin which is still underway and where only the results for the Autonomous Communities (Devolved Regions) of Extremadura and Castilla-La Mancha are currently available), and the results have since been made available to the various authorities and the public in general. The object of this article is to present the theoretical grounding of the work conducted by CEDEX to establish the map of maximum flows.

The methodology applied in the study was based on the selection of gauging stations and reservoirs with sufficient information on maximum flows to allow a suitably reliable estimation of the values of the different probabilities or return periods (quantiles). Diverse statistical techniques were applied to the selected series of data to verify the quality of information and these were given a distribution function in order to establish their frequency law, applying statistical regionalization techniques and incorporating data on historic flood events to improve the estimation of the high return period flood quantiles. The information on the quantiles obtained at the gauging points was extrapolated to the rest of the points of the drainage area through the application of hydrometeorological methods and statistical models in accordance with the surface area of the watershed.

As a result, and with the assistance of geographic information system (GIS) tools, six raster layers were obtained with a spatial resolution of 500

Como resultado, y con el apoyo de las herramientas de un sistema de información geográfica (SIG), se han obtenido seis capas raster, con resolución espacial de 500mx500m, que ofrecen información sobre caudales máximos en régimen natural (no se tiene en cuenta la alteración del régimen hidrológico provocada por la presencia de presas en la cuenca), en cuencas con una superficie mayor de 50 km², para los periodos de retorno de 2, 5, 10, 25, 100 y 500 años. Con el objetivo de facilitar la consulta de estas capas se ha desarrollado una herramienta informática (CauMax) que permite su visualización y consulta en el entorno de un SIG.

2. Selección y tratamiento de las series temporales de caudales máximos anuales

2.1. Criterios de selección de los puntos aforados

La selección de los puntos aforados se ha realizado mediante el establecimiento de un equilibrio entre los dos aspectos contrapuestos siguientes: por una parte, se debe disponer de registros temporales con el mayor número de datos posible que permita obtener una precisión suficiente en la estimación de los cuantiles; por otra, es necesario disponer de un número suficiente de estaciones que suponga una buena cobertura espacial y que permita extrapolar los resultados al resto de la red de drenaje. Además, se han descartado aquellas estaciones (o parte de sus datos) cuyos registros se encuentran alterados significativamente por la presencia de presas de cierta importancia, con la finalidad de obtener series temporales representativas del régimen natural. En este sentido, se ha considerado que un registro no está alterado de manera significativa cuando el porcentaje de cuenca controlada por embalses es menor del 10% o cuando el volumen total de embalse es menor del 10% del volumen medio de las crecidas.

De esta forma, se han seleccionado aquellas estaciones de aforo o embalses cuyo registro no alterado tenga una longitud mínima de 20 datos anuales, aunque en algunas zonas con poca densidad espacial de estaciones se ha rebajado el número de datos exigido hasta un mínimo de 15 datos anuales. Los datos recopilados se refieren tanto a caudales máximos medios diarios como a máximos

m x 500 m and providing information on the natural maximum flows (without considering the alteration in the hydrological regime caused by the presence of dams in the basin), in those basins with a surface area of over 50 km², for return periods of 2, 5, 10, 25, 100 and 500 years. A computer application (CauMax) has been developed to aid the interpretation of these layers and allow visualization and consultation in a GIS environment.

2. Selection and processing of the time series of annual maximum flows

2.1. Gauge point selection criteria

The selection of the gauging points was made by establishing an equilibrium between the following opposing aspects: on the one hand, sufficient records with the greatest possible amount of data has to be available in order to make a sufficiently precise estimation of the quantiles; while, on the other, a sufficiently large number of stations are required in order to ensure correct spatial coverage and to allow the extrapolation of the results to the rest of the drainage area. We have similarly discarded those stations (or parts of the data) whose recordings are significantly altered by the presence of dams of certain importance, in order to obtain representative time series of the natural regime. In this respect, it has been considered that a recording is not significantly altered when the proportion of the basin controlled by reservoirs is less than 10% or when the total volume of the reservoir is less than 10% of the mean volume of the floods.

As such, a selection has been made of those gauging stations and reservoirs with an unaltered record of a least 20 annual gauge data, though in some areas with a more reduced spatial distribution of stations, the required number of data has been brought down to a minimum of 15 annual data. The data collected refers to both average daily maximum flows and instantaneous maximum flows, with the maximum annual average daily flows subsequently being transformed to instantaneous flows by the procedure indicated further on. In the case of the coastal area of the Júcar basin and on account of the difficulty in transforming the

instantáneos, habiéndose procedido posteriormente a transformar los caudales medios diarios en instantáneos mediante el procedimiento que se expondrá más adelante. En el caso de la zona costera de la cuenca del Júcar, debido a la dificultad para transformar los caudales medios diarios en instantáneos en una zona con un régimen hidrológico tan extremo, se han empleado adicionalmente datos procedentes de la red del Sistema Automático de Información Hidrológica (SAIH).

2.2. Revisión de la calidad y representatividad de los datos de las series temporales

Un aspecto que condiciona de forma crítica, junto con la longitud de la serie temporal, la precisión de los cuantiles obtenidos mediante el análisis estadístico es la calidad de los datos. El tipo de proceso de medida llevado a cabo en las estaciones de aforo (basado en la utilización de una curva de gasto obtenida a partir de aforos directos, realizados habitualmente cuando circulan por el río caudales bajos o moderados) conlleva que normalmente los mayores errores en la medida de los caudales se cometan para los caudales más grandes, durante las avenidas, afectando de forma crítica a la calidad de los datos de las series de caudales máximos anuales. Por este motivo, se han llevado a cabo diversos procesos de revisión y contraste de la información para detectar aquellas estaciones con mayores problemas (en la totalidad o en parte de su registro) y proceder a su corrección o, de no ser posible, a eliminar la estación del análisis. Se han llevado a cabo los siguientes procesos de revisión:

- Identificación de tendencias temporales en las series: Las series temporales de máximos anuales corresponden a datos independientes entre sí, por lo que la existencia de tendencias temporales en las mismas puede suponer la presencia de algún problema en los datos. Para identificar estas posibles tendencias temporales se ha aplicado el test estadístico de Mann-Kendall, y se ha dibujado el gráfico del caudal acumulado de la serie temporal respecto al tiempo para identificar cambios en el comportamiento de los datos (figura 1).
- Identificación de inconsistencias entre los datos de caudal máximo anual de cada estación de

average daily flows into instantaneous flows in an area with such an extreme hydrological regime, recourse has also been made to data provided by the Automatic Hydrological Information System (SAIH).

2.2. Review of the quality and representation of time series data

Data quality is an aspect that, together with the length of the time series, seriously compromises the precision of the quantiles obtained by statistical analysis. The type of measuring process conducted at the gauging stations (based on the use of a rating curve obtained from direct gauging and as commonly carried out in rivers with low or moderate flows) implies that the greatest errors in the measurement of the flows tends to occur with the large flows, during flooding, and critically affects the data quality of the series of annual maximum flows. For this reason different review processes and information comparison have been conducted to detect those stations with the greatest problems (in respect to all or part of their recordings) to allow the correction of the same or, where this is not possible, the elimination of the station from the analysis. The following review processes were carried out:

- *Identification of time series trends: The time series of annual maximums correspond to data that is independent of each other and the presence of time trends within the same may imply the presence of some problem with the data. The Mann-Kendall test has been applied in order to identify these possible time trends with the ensuing plotting of a graph of the accumulated flow of the time series in order to identify changes in the behaviour of the data (Fig. 1).*
- *Identification of inconsistencies between the maximum annual flow of each gauging station with that of other stations set upstream from the same over the same recording period. In those cases where greater flows are detected upstream from those recorded downstream, a detailed study has subsequently been conducted to establish the possible causes for the inconsistencies.*

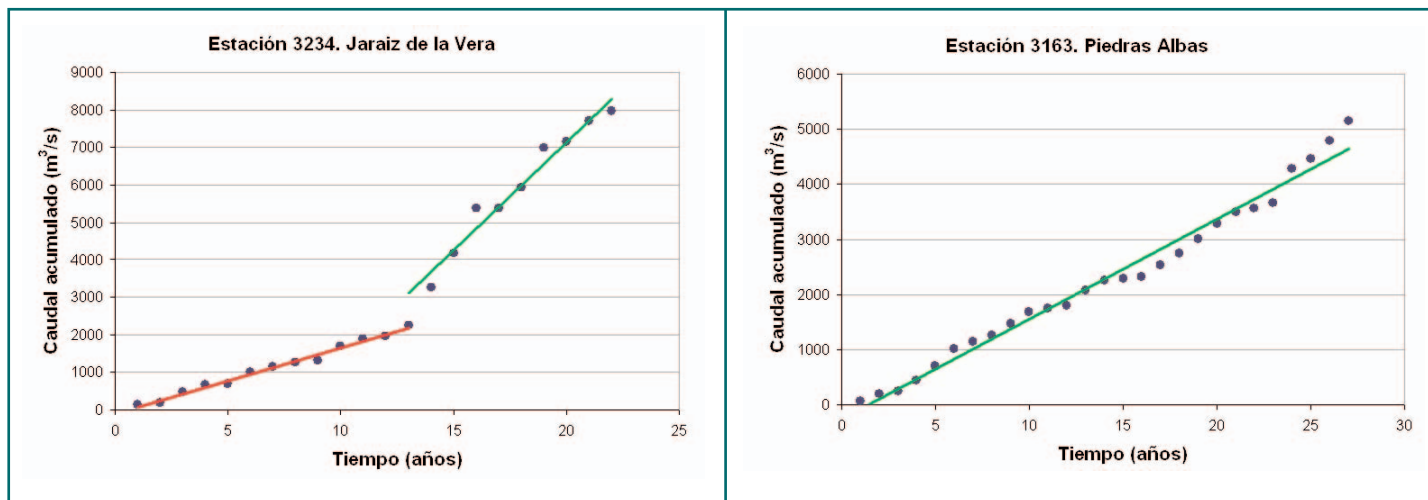


Fig. 1. Ejemplos de estaciones con (izquierda) y sin (derecha) tendencia temporal/ Examples of station with (left) and without (right) time trends.

aforos con los de aquellas otras estaciones situadas aguas arriba de la misma para los años de registro común. Aquellos casos en los que se han detectado medidas de caudal aguas arriba superiores a las registradas aguas abajo han sido objeto de un estudio detallado posterior para determinar las posibles causas de las inconsistencias.

- Identificación de datos anómalamente altos (*outliers*): Este proceso no pretende detectar deficiencias en la calidad de la información sino localizar aquellos datos singulares dentro de las series temporales poco representativos desde un punto de vista estadístico a la hora de realizar el ajuste de la ley de frecuencia. Se ha aplicado el procedimiento recomendado por el Water Resources Council de los Estados Unidos (1).
- Identificación de estaciones discordantes: Permite identificar aquellas series temporales que tienen unas características estadísticas significativamente diferentes al resto de las estaciones, mediante la medida de discordancia en base al análisis de los valores de los L-momentos (2).

2.3. Transformación de datos de caudal máximo diario en instantáneo

Como se comentó anteriormente, y con el objetivo de aumentar el número de estaciones disponible para conseguir una mayor cobertura espacial de la zona en estudio y de disponer de series temporales más largas, que permitan mejorar la estimación de los cuantiles, se decidió seleccionar tam-

- *Identification of outliers: This process does not aim to detect deficiencies in the quality of information but, instead, locate outlying data within the time series that show a marked deviation from a statistical point of view when making an adjustment of the law of frequency. The guidelines recommended by the U.S. Water Resources Council have been adhered to in this respect (1).*
- *Identification of discordant stations: Allowing the identification of those time series that have markedly different statistical characteristics from the other stations through the measurement of the discordance by analysis of the L-moment values (2).*

2.3. Transformation of maximum daily flow data to instantaneous flow

As indicated, and in order to increase the number of stations available to ensure the largest possible spatial cover of the study area and make use of longer data time series to improve the estimation of the quantiles, it was decided to incorporate those stations which only held recordings of the average daily maximum flows (Q_{md}) on all or part of their recordings. In these cases it was necessary to convert the average daily flows into their corresponding instantaneous flows (Q_i) by obtaining a (k) coefficient in relation to the same:

$$k = \frac{Q_i}{Q_{md}}$$

bién estaciones en las que en todo o en parte de su registro sólo se disponía del caudal máximo medio diario (Q_{md}). En estos casos fue necesario transformar los caudales medios diarios en sus correspondientes caudales instantáneos (Q_i) mediante la obtención de un coeficiente (k) que relaciona ambos:

$$k = \frac{Q_i}{Q_{md}}$$

En el caso de estaciones en las que se disponía de un registro temporal de caudales máximos instantáneos con escasas lagunas, el coeficiente k se obtuvo ajustando una regresión entre los valores de Q_i y Q_{md} en aquellos años en que se disponía de ambos tipos de información, aplicando el resultado para estimar el caudal instantáneo en aquellos años en que sólo se disponía del caudal medio diario (figura 2 izquierda). En las estaciones de aforo en las que los caudales instantáneos eran desconocidos en la mayor parte, o incluso en todo, el registro, así como en los embalses, el coeficiente se obtuvo mediante la aplicación de la fórmula de Fuller, que relaciona el coeficiente k con el área de la cuenca vertiente (A) y dos parámetros a y b que han sido calibrados regionalmente (figura 2 derecha).

$$k = \frac{Q_i}{Q_{md}} = 1 + \frac{a}{A^b}$$

Se ha llevado a cabo la calibración de los parámetros en cada una de las siguientes zonas: Miño-Sil más Galicia Costa, Cantábrico más Cuencas Internas del País Vasco, Duero, Tajo, cabecera del Guadiana hasta el río Jabalón, afluentes del río Guadiana por su margen izquierda aguas abajo del río Olivenza más ríos Tinto y Odiel, resto de la cuenca del río Guadiana más cuenca del río Guadalquivir, Segura, cabecera del Ebro hasta los ríos Arga en la margen izquierda y Queiles en la derecha, resto de la cuenca del Ebro por la margen izquierda y resto de la cuenca del Ebro por la margen derecha. En la demarcación hidrográfica del Júcar sólo se ha calibrado la ecuación de Fuller en la zona correspondiente a las cabeceras del río Júcar (hasta la confluencia con el río Cabriel) y Turia (hasta el embalse de Benageber), debido a la gran variabilidad existente en la relación entre caudales medios diarios e instantáneos en las cuencas más próximas a la costa.

In the case of those stations that had time recordings of maximum instantaneous flows at regular intervals, the k coefficient was obtained by adjusting for regression between the Q_i and Q_{md} values in those years where both types of information were available, and applying the result to estimate the instantaneous flow over those years where only the average daily flow was available (Fig. 2 left). In those gauging stations where the instantaneous flow was unknown throughout the majority or even all of the recordings, as well as in the reservoirs, the coefficient was obtained by applying the Fuller formula relating the k coefficient to the area of the drainage basin (A) and two regionally scaled parameters a and b (Fig. 2 right).

$$k = \frac{Q_i}{Q_{md}} = 1 + \frac{a}{A^b}$$

Parameters were calibrated in each of the following zones: Miño-Sil plus the Galician Coast, Cantabrian Basin and the Interior Basins of the Basque Country, Duero, Tajo, the headwaters of the Guadiana to the River Jabalón, tributaries of the River Guadiana on the left bank downstream of the River Olivenza together with the Tinto and Odiel rivers, the remainder of the Guadiana river basin plus the Guadalquivir river basin, Segura, the headwaters of the Ebro up to the Arga river on the left bank and the Queiles river on the right bank, the remainder of the Ebro basin on the left bank and the remainder of the Ebro basin on the right bank. In the Júcar river basin district the Fuller equation was only calibrated for the area corresponding to the headwaters of the River Júcar (up to where it meets the River Cabriel) and the River Turia (up to the Benageber reservoir), due to the large variation in terms of the relation between average daily flows and instantaneous flows in the basins closer to the coast.

2.4. Selection of gauging points

Following on from the preceding process, a series of gauging points were selected formed by both gauging stations, pertaining to both the Official Gauge Station Network (ROEA) and Automatic Hydrological Information System (SAIH), and reservoirs. Figure 3 shows the 400-plus measuring points selected.

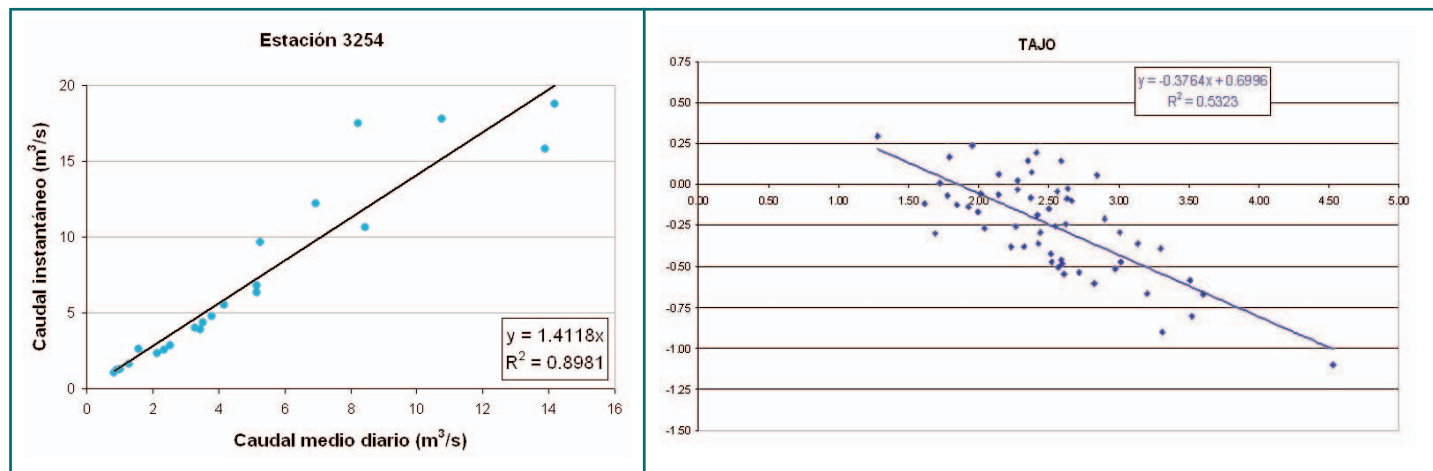


Fig. 2. Ejemplos de obtención del coeficiente que relaciona los caudales máximos medios diarios e instantáneos. Izquierda, regresión entre los caudales máximos medios diarios e instantáneos en la estación 3254. Derecha, ajuste de la ecuación de Fuller en la cuenca del Tajo/Examples of the procurement of coefficients relating the maximum average daily flows and instantaneous flows. Left, the regression between the maximum average daily flows and instantaneous flows at station 3254. Right, adjustment of the Fuller equation at the Tajo basin.

2.4. Selección de puntos aforados

Como consecuencia del proceso anterior, se ha seleccionado un conjunto de puntos aforados formado por estaciones de aforo, tanto de la Red Oficial de Estaciones de Aforo (ROEA) como del SAIH, y por embalses. En la figura 3 aparecen representados los más de 400 puntos de medida seleccionados.

3. Cálculo de los cuantiles en los puntos aforados

3.1. Selección del tipo de función de distribución y del procedimiento de ajuste

La estimación de los cuantiles de caudal máximo en los puntos aforados ocupa, como se ha dicho, una parte central en el desarrollo del trabajo. Con objeto de mejorar la precisión con que se realiza esa estimación, se ha llevado a cabo un análisis de las funciones de distribución y procedimientos de ajuste que presentan un mejor comportamiento para obtener dichos cuantiles, y que representan mejor las características estadísticas de las series en las distintas zonas geográficas. En ese análisis se ha tenido en cuenta tanto la capacidad descriptiva del modelo, es decir su capacidad para ajustarse con precisión a la distribución de caudales observados, como su capacidad de predicción, es decir, su robustez para estimar los cuantiles asociados a los mayores periodos de retorno (3). El error en la estimación aumenta al aumentar el periodo de retorno

3. Calculation of the quantiles at the gauge points

3.1. Selection of the type of distribution function and adjustment procedure

The estimation of the quantiles of the maximum flow at the gauging points is paramount to the development of the work. In order to make this



Fig. 3. Puntos de medida seleccionados para el trabajo: estaciones de la ROEA, estaciones de la red SAIH y embalses/Measuring points selected for the work: ROEA stations, stations on the SAIH network and reservoirs.

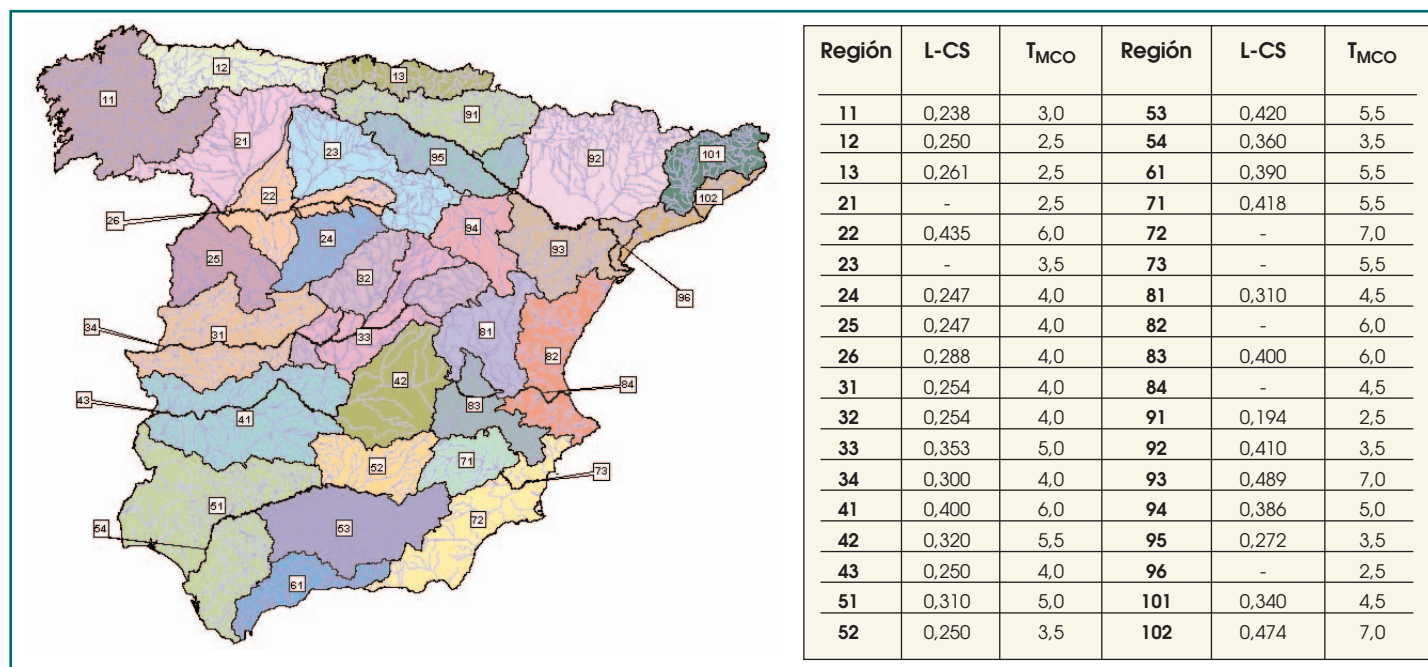


Fig. 4. Regiones identificadas con homogeneidad estadística y valor del L-coeficiente de sesgo y del periodo de retorno de la máxima crecida ordinaria en cada una de ellas/
 Regions identified with statistical homogeneity and value of the L-coefficient of skewness and the return period of the maximum ordinary flood in each of the same.

del cuantil estimado, siendo mayor el error cuanto menor es la longitud de la serie temporal disponible. Dicho error es especialmente importante en aquellos casos en los que es necesario emplear funciones de distribución de tres o más parámetros, que vienen determinadas por el valor de los momentos estadísticos de tercer o cuarto orden (relacionados con los coeficientes de sesgo y de curtosis), cuya estimación a partir de la muestra presenta un mayor error. Para reducir la incertidumbre en la estimación de dichos momentos se han empleado técnicas estadísticas de regionalización, las cuales tratan de mejorar el ajuste de las funciones de distribución obteniendo una estimación más precisa de los momentos estadísticos de mayor orden al emplear conjuntamente la información contenida en todas las estaciones existentes en una determinada región con comportamiento estadístico homogéneo.

De esta forma, en primer lugar, se ha realizado un estudio para la identificación de regiones estadísticamente homogéneas en el territorio de la España peninsular (figura 4). Esta identificación se ha realizado siguiendo principalmente criterios geográficos, a partir de la orografía de la cuenca, la altitud, el régimen de lluvias y el valor de los cuantiles de precipitación, etc., comprobando posteriormente mediante la aplicación de los tests estadísticos de homogeneidad de

estimate as precise as possible, an analysis was made of the distribution functions and adjustment procedures best suited to the establishment of these quantiles and that best represented the statistical characteristics of the series in the different geographical areas. This analysis took into account both the descriptive capacity of the model, that is to say its capacity for precise adjustment to the observed distribution flows, and its forecasting capacity or robustness to estimate quantiles associated with longer return periods (3). The error in estimation increases with the length of the return period of the estimated quantile, and the smaller the length of available time series the greater the error. This error is particularly important in those cases where it is necessary to employ distribution functions of three or more parameters, and determined by the value of the third or fourth-order statistical moments (related to the skewness and kurtosis), whose estimation from the sample reveals the greatest error. Statistical regionalization techniques have been employed to reduce the uncertainty in the estimation of these moments, these techniques serving to improve the adjustment of the distribution functions in order to obtain a more precise estimate of the higher order statistical moments through the combined use of information contained at all the

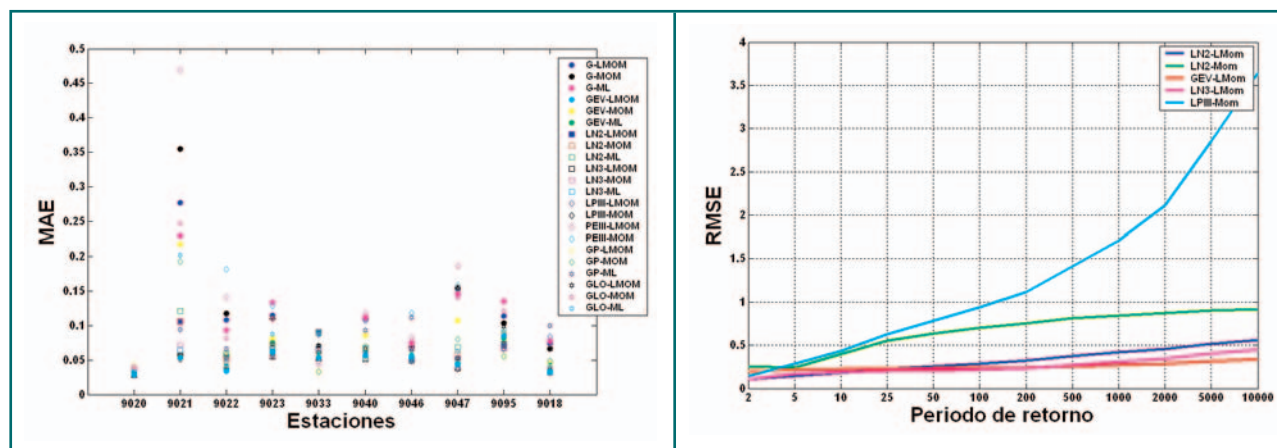


Fig. 5. Ejemplos de error en el ajuste a los datos observados (izquierda) y de análisis de la capacidad predictiva de las distintas funciones y procedimientos de ajuste (derecha) correspondientes a la región 92 (cuenca del Ebro)/Examples of the error in the adjustment of observed data (left) and analysis of the forecasting capacity of the different adjustment procedures and functions (right) corresponding to region 92 (Ebro basin).

Wiltshire (4) y de Hosking y Wallis (2) lo adecuado de dicha identificación. Este método tiene la ventaja de facilitar la asignación de una región a un punto determinado, ya que las regiones están delimitadas perfectamente por los accidentes geográficos de las cuencas. Como resultado de los análisis llevados a cabo, se ha decidido adoptar un valor regional únicamente para el coeficiente de sesgo, determinando el coeficiente de variación a partir de la información local.

El siguiente paso ha consistido en analizar, en cada una de las regiones identificadas, la capacidad descriptiva de los distintos modelos. Se ha estudiado el ajuste de las diferentes funciones de distribución a los caudales observados en las estaciones de cada región, a los que se ha asignado probabilidades muestrales. En cada una de las estaciones se han ajustado las diferentes funciones consideradas (Gumbel, valores extremos generalizada (GEV), log-normal de 2 y 3 parámetros (LN2 y LN3), Pareto generalizada (GP), logística generalizada (GLO), log-Pearson tipo III (LPIII) y Pearson tipo III (PEIII)) mediante diferentes métodos de ajuste (momentos (Mom), L-momentos (L-Mom) (5) y máxima verosimilitud (ML)) y se ha cuantificado el error existente entre los datos observados y dichas funciones. Los ajustes de las funciones se han realizado utilizando únicamente la información local sin la aplicación de ningún método de regionalización. Se ha obtenido el error como el valor medio de las distancias adimensionales entre cada pareja de resultados. Para cada modelo o función de distribución se ha calculado la media de los errores absolutos de todas las estaciones existentes en la región (MAE) (figura 5 izquierda).

existing stations with statistically homogeneous behaviour in a given region.

To this end a study was first made to identify the statistically homogeneous regions within the Spanish peninsular (Fig. 4). This identification was primarily made on geographical criteria, on the basis of the orography of the basin, the height, the rainfall regime and the value of the precipitation quantiles, etc., and the suitability of this identification was then verified by the application of the Wiltshire (2) and Hosking and Wallis (2) statistical test for homogeneity. This method has the advantage of allowing the allocation of a region to a specific point, as the boundary lines of the regions are perfectly established by the geographical features of the basins. As a result of the analysis carried out, it was decided that a regional value would only be adopted for the skewness coefficient and that the coefficient of variation would be established on the basis of local information.

The next stage consisted of an analysis of the descriptive capacity of the different models in each of the identified regions. A study was made of the adjustment of the different distribution functions to the flows observed at the stations of each region which were assigned sample probabilities. In each of the stations adjustment was made in accordance with the different functions considered (Gumbel, generalized extreme values (GEV), log-normal of 2 and 3 parameters (LN2 and LN3), generalized Pareto (GP) distribution, generalized logistic distribution (GLO), Log-Pearson III (LPIII) and Pearson Type III (PEIII) by different adjustment methods (moments

Previo al análisis de la capacidad predictiva, se ha comprobado que el esquema de regionalización adoptado es capaz de reproducir en cada región unas características estadísticas similares a las que tienen los datos observados. Para ello se ha utilizado un método basado en simulaciones de Monte Carlo. Se han generado regiones sintéticas similares a la región observada, es decir, con un número igual de estaciones y con series de datos sintéticas de igual longitud. Se han comparado las características estadísticas (coeficiente de variación (Cv), coeficiente de sesgo (Cs), máximo valor estandarizado, L-coeficiente de variación (L-Cv) y L-coeficiente de sesgo (L-Cs)) de las regiones sintéticas con las de la región real. Los resultados muestran que el patrón de regionalización adoptado es capaz de reproducir con bastante precisión la distribución en la región de los estadísticos considerados. No obstante, en cada región hay una serie de funciones de distribución y procedimientos de ajuste que son los que mejor reproducen el comportamiento de los distintos estadísticos, y que han sido los utilizados al analizar el comportamiento predictivo de las funciones de distribución en cada región.

El análisis de la capacidad predictiva se ha realizado para periodos de retorno superiores a 100 años y su objetivo es valorar la eficiencia y robustez de los modelos en cuanto a la extrapolación de los cuantiles. Como base de comparación, a los datos reales de las estaciones de cada región se les ha ajustado una función kappa de cuatro parámetros, más general que todas las utilizadas, usando para el ajuste los valores regionales de L-coeficiente de sesgo y L-curtosis, y los valores locales de la media y el L-coeficiente de variación. A partir de estas funciones se han generado regiones sintéticas similares a la observada, es decir, con un número igual de estaciones y con series de datos sintéticas de igual longitud, mediante simulaciones de Monte Carlo. Por otra parte, se han ajustado los modelos identificados previamente como los mejores reproduciendo las características estadísticas de la región, a cada una de las series sintéticas generadas, tomando el valor regional del L-Cs de cada región sintética y valores locales de la media y el L-Cv.

Para cuantificar la capacidad predictiva de los distintos modelos seleccionados se ha calculado, en cada estación de la región y para cada periodo de retorno, el error cuadrático medio como suma del

(Mom), L-moments (L-Mom) (5) and maximum likelihood (ML)) and the error between observed data and these functions was then quantified. The adjustment of functions was made purely on the basis of local information without applying any regionalization method. The error has then been obtained as the average value of the dimensionless distances between each pair of results. The mean absolute error (MAE) of all the stations in the region has been calculated for each distribution function or model (Fig. 5 left).

Prior to the analysis of the forecasting capacity, it was necessary to ensure that the adopted regionalization arrangement was capable of reproducing similar statistical characteristics in each region to those established from observed data. To this end, a method was employed based on Monte Carlo simulations. Synthetic regions similar to the observed region were generated, using the same number of stations and series of synthetic data of equal length. The statistical characteristics (coefficient of variation (Cv), skewness coefficient (Cs), maximum standardized value, L-coefficient of variation (L-Cv) and L-coefficient of skewness (L-Cs)) of the synthetic regions were then compared with those of the real region. The results show that the adopted regionalization arrangement was capable of reproducing the distribution of the statistical considerations in the region with a sufficient degree of precision. However, there are a series of distribution functions and adjustment procedures in each region that best reflect the behaviour of the different statistics and which have been employed to analyze the forecasting behaviour of the distribution functions in each region.

The analysis of forecasting capacity has been made for return periods of over 100 years with the objective of appraising the efficiency and robustness of the models in terms of the extrapolation of the quantiles. By way of comparison, the real data of the stations in each region was adjusted by a four parameter kappa distribution function, more general than all those employed, using the regional values of the L-coefficient of skewness and the L-kurtosis and the local values of the mean and the L-coefficient of variation for the adjustment. Synthetic regions have been generated from these functions by way of Monte Carlo simulations, these regions being similar to those observed, that is, with the same number of

cuadrado del sesgo, valor medio de los errores o diferencias entre los cuantiles base y los cuantiles obtenidos con el modelo, y la varianza de los mismos. Finalmente, se ha obtenido la media en la región del valor estandarizado del error cuadrático medio para cada periodo de retorno (RMSE), lo que ha permitido valorar el comportamiento de cada modelo (figura 5 derecha).

Como resultado de los análisis anteriores se ha concluido que el modelo que mejor describe el comportamiento de los caudales máximos anuales en cada región es la función de valores extremos generalizada (GEV) ajustada por el método de los L-momentos (L-Mom) (5) asumiendo, como se ha dicho, un valor regional para el L-coeficiente de sesgo (L-CS) en cada región (figura 4) (3). Como casos especiales, se ha decidido utilizar la función Gumbel ajustada por el método de los L-momentos en las regiones 21 y 23 de la cuenca del Duero y en el eje principal del Ebro hasta su confluencia con el río Segre, y la función de valores extremos de dos componentes (TCEV) en las regiones 72, 82 y 84 del levante y sureste peninsular, de la que se hablará en el próximo apartado.

3.2. Información histórica empleada en el ajuste de las leyes de frecuencia

Como se ha expuesto en el apartado anterior, una forma de aumentar la precisión en la estimación de los cuantiles consiste en reducir la incertidumbre con que se estiman los momentos estadísticos de mayor orden mediante el empleo de información regional. Otra posible forma es emplear en el análisis estadístico información sobre grandes avenidas ocurridas en el pasado mucho antes de que se iniciase la medida sistemática de los caudales y, por tanto, no registrada en las estaciones de aforo. Esta información tiene el inconveniente de ser incompleta y requerir de técnicas especiales para ser introducida en el análisis, pero tiene un gran valor para mejorar la estimación de la cola derecha de la distribución al aportar datos directos sobre el comportamiento de la población en ese rango de probabilidades. De esta forma, y con objeto de mejorar la estimación de los cuantiles de alto periodo de retorno, se ha introducido, siempre que ha sido posible, información sobre avenidas históricas en el ajuste de las leyes de frecuencia.

stations and with series of synthetic data of equal length. In addition to which, those models that were previously identified as providing the best representation of the statistical characteristics of the region were adjusted to each of the synthetic series generated, taking the regional value of the L-Cs of each synthetic region, the local mean values and the L-Cv.

In order to quantify the forecasting capacity of the different models selected, a calculation was made for each station in the region and for each return period of the mean square error as a sum of the skew squared, the mean value of errors or differences between the base quantiles and the quantiles obtained from the model and the variance of the same. The mean standardized value of the root mean square error (RMSE) for each return period was then obtained for the region, which allowed an appraisal of the behaviour of each model (Fig. 5 right).

As a result of this analysis it was concluded that the model that best described the behaviour of the maximum annual flows in each region was the generalized extreme value (GEV) function adjusted by the method of L-moments (L-Mom) (5) when assuming, as indicated, a regional value for the L-coefficient of skewness (L-Cs) in each region (figure 4) (3). As a special case, the Gumbel function adjusted by the method of L-moments was employed in regions 21 and 23 of the Duero basin and the main axis of the Ebro up to its confluence with the River Segre, and the two-component extreme value function (TCEV) was employed in eastern and south-eastern regions 72, 82 and 84, as referred to in the following section.

3.2 Historical information employed in the adjustment of the frequency laws

As indicated above, one of the ways of improving the quantile estimation is by reducing the uncertainty related to the estimation of higher order statistical moments by using regional information. Another possibility is to employ statistical information in the analysis regarding large flood events that occurred in the past, long before the systematic measurement of flows, and that have subsequently not been recorded at the gauging stations. This information, while having the inconvenience of



Fig. 6. Localización de las avenidas históricas consideradas en el ajuste de las leyes de frecuencia/ Location of historical floods considered in the adjustment of the laws of frequency.

Se han seleccionado para su utilización en el estudio (figura 6) sólo aquellos eventos de los que se conoce el caudal punta en algún sitio próximo a alguno de los puntos aforados, así como su orden de importancia en un determinado periodo de tiempo (periodo en el cual se dispone de información histórica), lo cual permite asignar probabilidad al caudal histórico. Para introducir esta información en el ajuste (figura 7) se ha empleado el método de ponderación de las series de datos propuesto por el Water Resources Council de los Estados Unidos (1) por estar especialmente indicado cuando se utiliza el método de los momentos (o una de sus variantes) como procedimiento de ajuste.

La introducción de información histórica en el ajuste de las leyes de frecuencia en la zona costera del levante y sureste peninsular (regiones 72, 73, 82 y 84 de la figura 4) pone de manifiesto la existencia de un quiebro en dicha ley (figura 7 derecha), que es el reflejo de un comportamiento hidrológico muy extremo

being incomplete and requiring special techniques to allow introduction in the analysis, serves as great assistance for the improved estimation of the right tail of the distribution by providing direct data in the behaviour of the population in this range of probabilities. In this way, and in order to improve the estimation of high return period quantiles, information of historic floods have been introduced, whenever possible, in the adjustment of the frequency laws.

For the purposes of the study (Fig. 6) recourse has only been made to those events in which the point flow somewhere near to the gauging points and their level of importance within a specific period of time (the period for which historical information is available) are both known and which make it possible to assign probability to the historical flow. The data series weighting method proposed by the U.S. Water Resources Council (1)

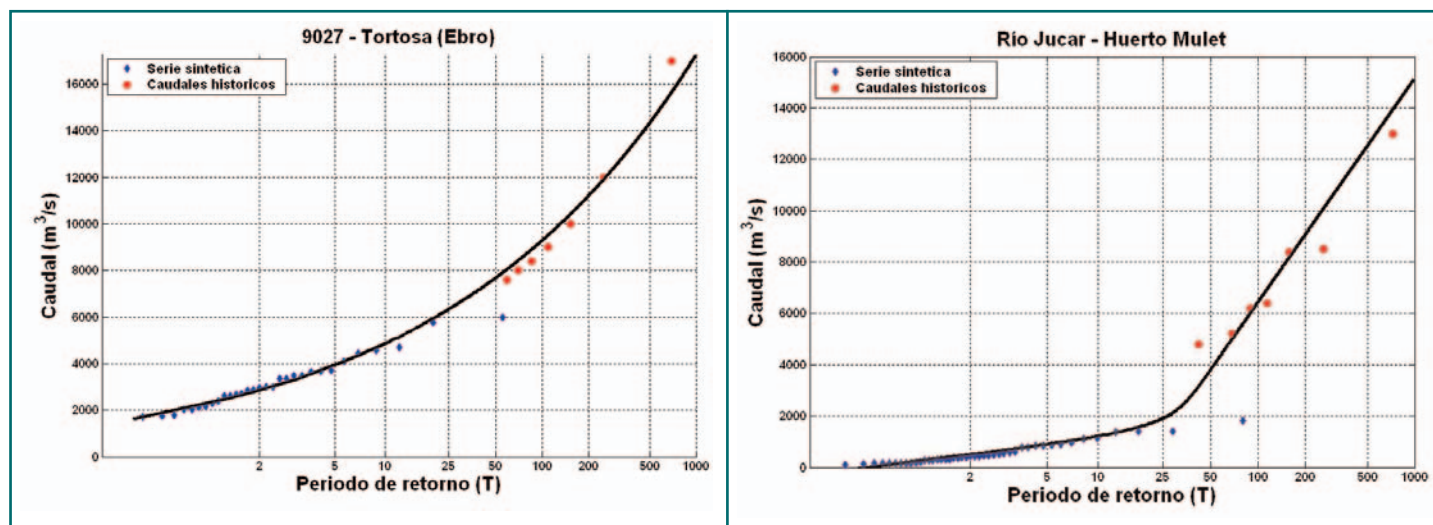


Fig. 7. Ejemplos de funciones de distribución ajustadas empleando datos de avenidas históricas. Izquierda, función GEV en Tortosa (río Ebro). Derecha, función TCEV en Huerto Mulet (río Júcar)/Examples of the adjusted distribution functions employing data of historical floods. Left, GEV function in Tortosa (River Ebro). Right TCEV function in Huerto Mulet (River Júcar).

y de la existencia de dos mecanismos muy diferenciados de generación de las avenidas. Este tipo de comportamiento hidrológico se puede representar adecuadamente, desde un punto de vista estadístico, mediante la utilización de la función de distribución de valores extremos de dos componentes (composición de dos funciones Gumbel). El ajuste de este tipo de función presenta una gran dificultad ya que la segunda rama de la función corresponde a eventos de baja probabilidad de ocurrencia, por lo que no suelen estar suficientemente representados en muestras de escasa longitud. De esta forma, la disponibilidad de información sobre avenidas históricas cobra una especial relevancia a la hora de ajustar la segunda rama de esta función. En aquellos casos en los que no se dispone de ese tipo de información, el ajuste de la función presenta una elevada incertidumbre. Con objeto de reducir esa incertidumbre se ha desarrollado un procedimiento de ajuste consistente en ajustar cada rama de la función, correspondiente a una función Gumbel, de manera independiente, obteniendo posteriormente la función conjunta mediante el producto de las dos funciones Gumbel (3).

La primera rama de la función se puede estimar mediante el ajuste local de una función Gumbel a los datos de la muestra mediante el método de los L-momentos. La estimación de los parámetros correspondientes a la segunda rama se realiza asumiendo un valor regional del L-coeficiente de variación $((t_2)_2)$ y estimando el valor de la media $((\lambda_1)_2)$ mediante una

has been used to introduce this information in the adjustment (Fig. 7) as this is particularly recommended when using the method of moments (or one of its variations) as the adjustment procedure.

The introduction of historical information in the adjustment of the laws of frequency in the eastern and south-eastern coastal zone (regions 72, 73, 82 and 84 of figure 4) reveals the breakdown in this law (Fig. 7 right), as reflected by very extreme hydrological behaviour and the presence of two very different means of generating floods. This type of hydrological behaviour may be suitably represented, from a statistical point of view, by using the extreme value distribution function of two components (two-parameter Gumbel distribution). The adjustment of this type of function is very difficult as the second branch of the function corresponds to events that are very unlikely to occur and which, in the main, are not sufficiently represented in long scale samples. To this extent, the availability of information on historical floods takes on particular importance when adjusting the second branch of this function. In those cases where no such information is available, the adjustment of the function presents a high level of uncertainty. In order to reduce this uncertainty an adjustment procedure has been developed consisting of the independent adjustment of each branch of the function, corresponding to a Gumbel function, and then

Tabla 1. Parámetros para ajuste de la función TCEV.
Table 1. Parameters for adjustment of the TCEV function.

Región	L-Cv	a	b	c
72 y 73	-0.26	1.5846	1.2280	0.8554
82 y 84	-0.24	2.6039	0.5659	0.6861

regresión con los estadísticos correspondientes a la primera rama $((\lambda_1)_1$ y $(t_2)_1$) (tabla 1):

$$(\lambda_1)_2 = -10^a (\lambda_1)_1^b (t_2)_1^c$$

4. Estimación de los cuantiles en los puntos de la red fluvial no aforados

4.1. Estimación en cuencas pequeñas: método hidrometeorológico

Una vez determinados los cuantiles en los puntos aforados, la estimación de los mismos en los puntos de la red fluvial no aforados y con pequeña superficie de cuenca vertiente se ha basado en la aplicación del método racional modificado (6), que se ha calibrado y validado a partir de los datos registrados en las estaciones de aforo. El modelo se basa en la siguiente formulación:

$$Q = \frac{CIA}{3.6} K$$

Donde Q es el caudal correspondiente a un determinado periodo de retorno en m^3/s , C es el coeficiente de escorrentía cuyo valor se obtiene en función de la precipitación máxima diaria y el umbral de escorrentía (P_0), I es la intensidad de precipitación media para un determinado periodo de retorno en mm/h , A es la superficie de la cuenca en km^2 y K es el coeficiente de uniformidad que se calcula en función del tiempo de concentración de la cuenca.

Para la calibración y validación del método de cálculo se han seleccionado, en una primera fase, un conjunto de estaciones de aforos (185 estaciones para calibración y 60 estaciones para validación de los resultados) en base a criterios de calidad de sus registros y de localización espacial, teniendo en cuenta la aplicabilidad del método y la cobertura del territorio. Mediante un SIG se ha obtenido, en cada una de las estaciones consideradas, el valor de las variables que intervie-

obtaining a joint function as the product of the two Gumbel functions (3).

The first branch of the function may be estimated by the local adjustment of a Gumbel function to the sample data by the L-moments method. The estimation of the corresponding parameters of the second branch is made by assuming a regional value of the L-coefficient of variation $((t_2)_2)$ and estimating the value of the mean $((\lambda_1)_2)$ by a regression with the statistics corresponding to the first branch $((\lambda_1)_1$ and $(t_2)_1$) (table 1):

$$(\lambda_1)_2 = -10^a (\lambda_1)_1^b (t_2)_1^c$$

4. Estimation of the quantiles at ungauged points on the river network

4.1. Estimation in small basins: hydrometeorological method

Once the quantiles at the gauging points have been established, the estimation of these at ungauged points on the river networks and in locations with small basin drainage area was then based on the application of the modified rational method (6) and the calibration and validation of the same in accordance with the data recorded at the gauging stations. This model is based on the following formula:

$$Q = \frac{CIA}{3.6} K$$

Where Q is the flow in m^3/s corresponding to a specific return period, C the Runoff Coefficient established as the ratio between the maximum daily rainfall and the runoff threshold (P_0), I the average rainfall intensity in mm/h for a specific return period, A the surface area of the basin in km^2 and K the uniformity coefficient calculated in accordance with the basin time of concentration.

For the calibration and validation of the calculation method a group of gauging stations were initially selected (185 stations for calibration and 60 stations for the validation of the results) on the basis of the quality of their data time series and location, in view of the applicability of the method

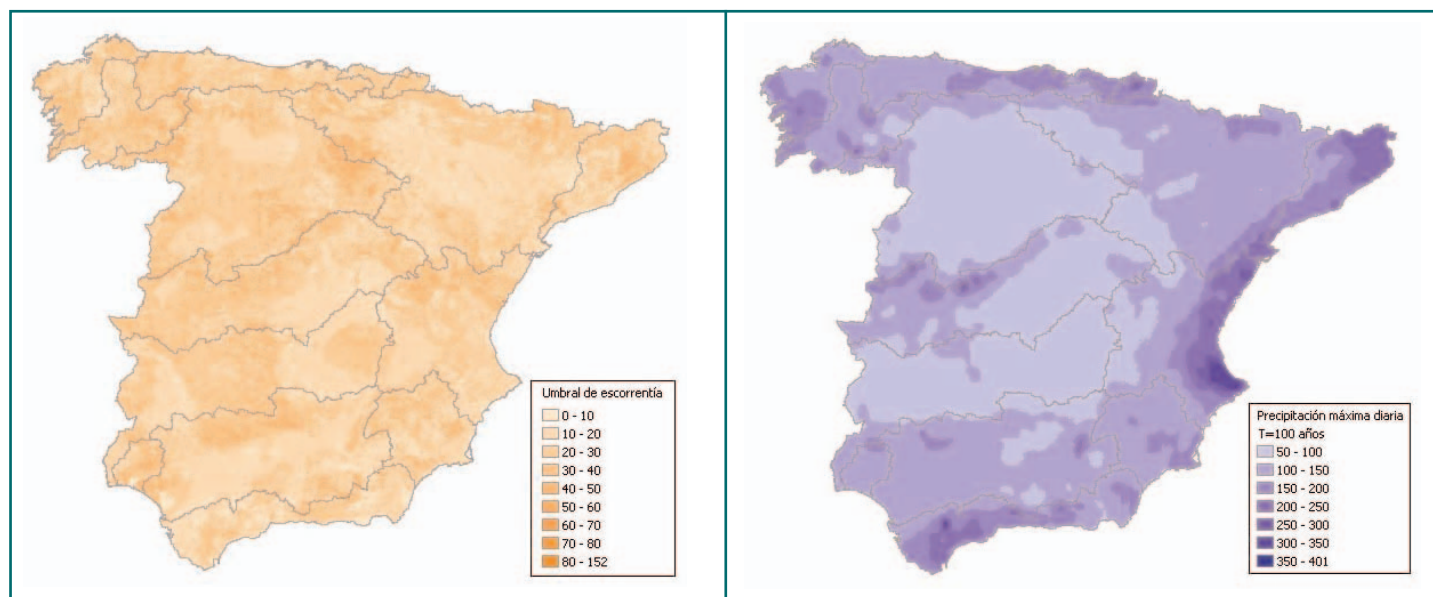


Fig. 8. Mapa de umbral de escorrentía en condiciones medias de humedad antecedente (izquierda) y mapa de precipitaciones máximas diarias para 100 años de periodo de retorno (derecha)/Map of the runoff threshold in average antecedent moisture conditions (left) and maximum daily rainfall map for 100-year return period (right).

nen en el método de cálculo (superficie de cuenca, tiempo de concentración, umbral de escorrentía, precipitación máxima diaria y factor de torrencialidad) (figura 8), considerando el umbral de escorrentía en condiciones medias de humedad antecedente, y se han calculado los cuantiles de caudal máximo para distintos periodos de retorno. El umbral de escorrentía se ha obtenido siguiendo la metodología utilizada en (7), a partir del mapa de usos de suelo Corine Land Cover 2000 y de datos de infiltración actualizados. Los mapas de precipitación se han elaborado a partir de los resultados obtenidos en (8).

La calibración del método se ha realizado mediante un coeficiente corrector (β) que multiplica al umbral de escorrentía correspondiente a condiciones medias de humedad antecedente. En cada estación de aforos considerada se ha determinado un valor de β que permite ajustar los cuantiles derivados del análisis estadístico de los datos de las estaciones de aforos con los estimados mediante el método racional modificado (figura 9). Este ajuste se ha realizado atendiendo principalmente al periodo de retorno de 10 años, ya que para periodos de retorno inferiores el cálculo agregado de la escorrentía puede no ser correcto, y para periodos de retorno superiores existe mayor incertidumbre en la determinación de los cuantiles debida a la escasa longitud de las series de datos.

Con objeto de disponer de un mapa del coeficiente β a escala nacional se ha realizado una interpola-

and the territorial coverage. By means of a GIS the value of the variables intervening in the calculation method (basin surface, time of concentration, runoff threshold, maximum daily rainfall and torrentiality factor) (Fig. 8) was obtained for each of the stations in question, considering the runoff threshold for average antecedent moisture conditions and the maximum flow quantiles were calculated for different return periods. The runoff threshold has been obtained by the method used in (7) on the basis of the Corine Land Cover 2000 map of land uses and updated infiltration data. The precipitation maps were prepared on the basis of the results obtained in (8).

The calibration of the method has been made by a correction factor (β) that multiplies the runoff threshold corresponding to the average antecedent moisture conditions. A β value was established for each gauging station to adjust the quantiles derived from the statistical analysis of the gauging station data to that estimated by the modified rational method (Fig. 9). This adjustment has been made considering a 10-year return period at the outset, as the aggregate calculation of runoff over shorter return periods may well be incorrect and over longer return periods there is greater uncertainty in the establishment of the quantiles on account of the reduced length of the data series.

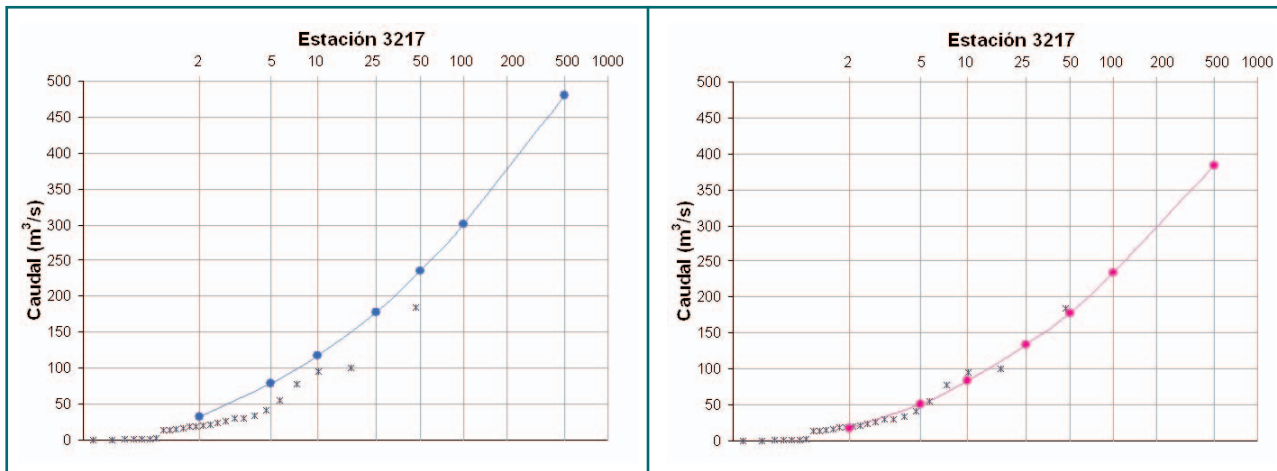


Fig. 9. Ajuste de la ley de frecuencia obtenida mediante el método racional a los datos registrados en la estación de aforos 3217 (izquierda: resultado obtenido con el P_0 para condiciones medias de humedad antecedente, derecha: resultado obtenido con el P_0 calibrado)/Adjustment of the law of frequency obtained by the rational method to the data recorded at gauging station 3217 (left: result obtained with the P_0 for average antecedent moisture conditions, right: result obtained with calibrated P_0).

ción espacial mediante kriging ordinario de los valores puntuales de dicho coeficiente resultantes del proceso de calibración, obteniendo como resultado un mapa raster con resolución espacial de 500mx500m (figura 10 izquierda). La distribución de los resultados obtenidos es acorde con el comportamiento general de este tipo de coeficiente según distintos autores (9) (10). Las áreas más húmedas del norte son las que presentan valores más bajos mientras que en la zona del sudeste, donde el clima es más árido, los valores del coeficiente son altos.

Del contraste de resultados en las estaciones de validación se concluye que las estimaciones regionales

In order to prepare a map of the β factor on a national scale, a spatial interpolation was made by ordinary kriging of the point values of the said factor resulting from the calibration process, this then providing a raster map with a spatial resolution of 500 m x 500 m (Fig. 10 left). The distribution of the results obtained is in agreement with the general behaviour of this type of factor according to different authors (9) (10). The wetter areas to the north are seen to show the lowest values while the drier area to the southeast is shown to have the highest factor values.

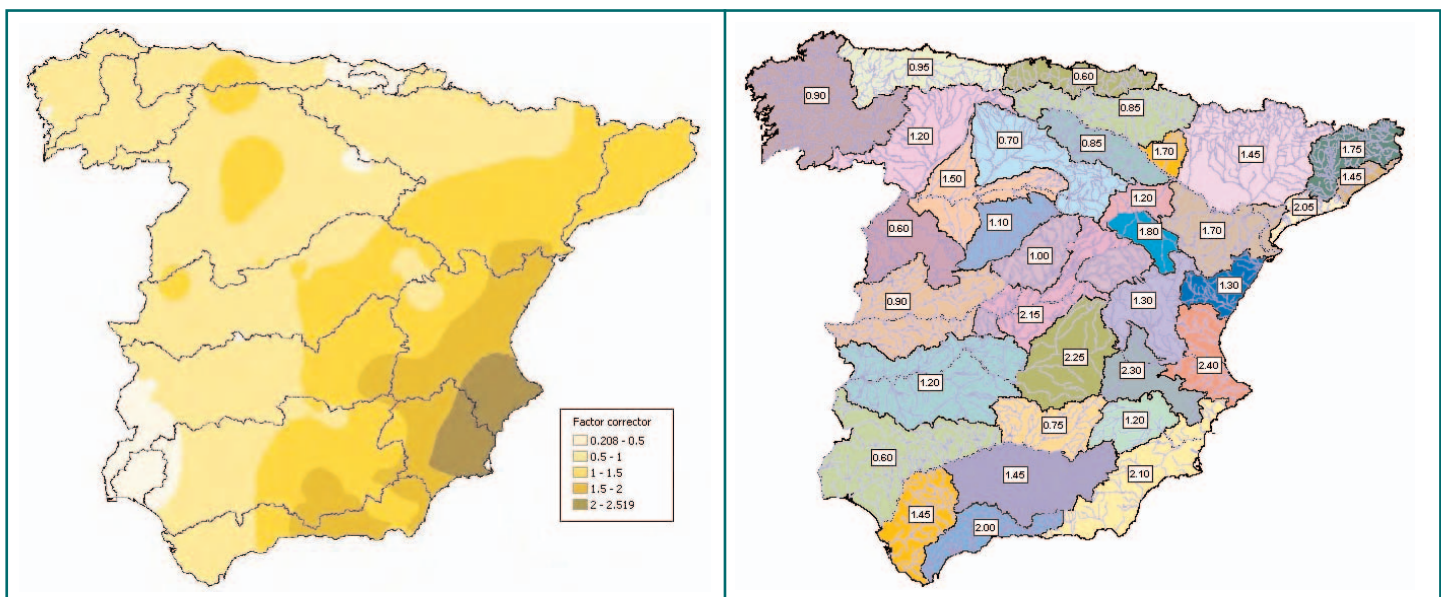


Fig. 10. Mapa de coeficientes correctores del umbral de escorrentía (izquierda) y mapa de valores medios por regiones (derecha)/Map of correction factors of the runoff threshold (left) and map of average values for regions (right).

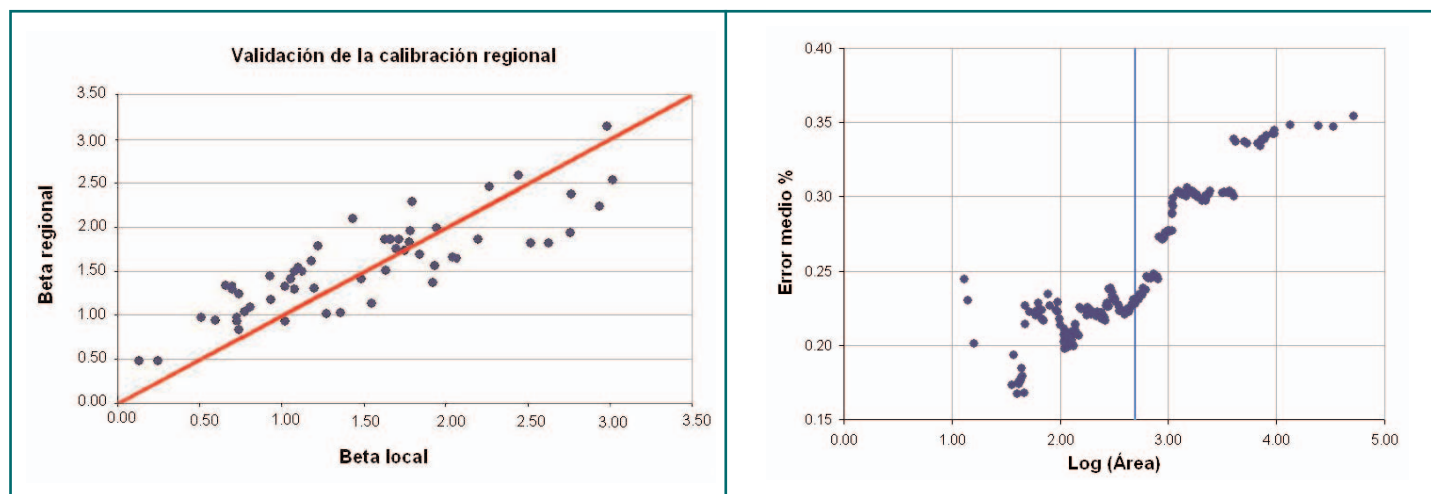


Fig. 11. Izquierda: Comparación entre los valores regionales y locales del coeficiente β para las estaciones de validación. Derecha: Errores en la estimación de los cuantiles mediante el método racional modificado en función del área de la cuenca vertiente/Left: Comparison between the regional and local values of the β coefficient for the validation stations. Right: Errors in estimation of the quantiles by the modified rational method in accordance with the area of the drainage basin.

del mapa no son sesgadas y proporcionan unos valores razonables a pesar de que existe cierta dispersión respecto a los valores locales (figura 11 izquierda). Por otra parte, y con la finalidad de determinar el tamaño máximo de cuenca vertiente para la aplicación del método racional modificado que permite obtener la ley de frecuencia de caudales con cierta precisión, se ha analizado la distribución del valor medio del error en función del área (figura 11 derecha). Se observa que hasta superficies de cuenca del orden de 500 km² el error medio es más o menos estable y del orden del 20-25%. Para superficies de cuenca mayores el valor del error va aumentando de forma constante. Teniendo en cuenta estos resultados, para la elaboración del mapa de caudales máximos, se ha decidido aplicar el método racional en aquellos puntos de la red fluvial con cuencas vertientes de hasta unos 500 km² de superficie.

Posteriormente, el proceso de calibración y validación se amplió mediante la incorporación de las estaciones seleccionadas para elaborar el mapa de caudales máximos. Este análisis permitió constatar que, aunque el mapa regional de coeficientes β refleja correctamente las variaciones de dicho coeficiente a lo largo del territorio nacional en función de las características climáticas de cada zona, presenta en algunos casos desviaciones importantes respecto a los valores obtenidos en las estaciones de aforo. En consecuencia, los valores proporcionados por dicho mapa deben tomarse únicamente como orientativos y no como un valor exacto. De esta forma, se ha considerado preferible y más acorde con su variabilidad real, el caracterizar estadísticamente los valores del coeficiente corrector β

From a comparison of the result at the validation stations it was seen that the regional estimates of the map are not skewed and provide reasonable values in spite of a certain dispersion with respect to the local values (fig. 11 left). Additionally, and in order to establish the maximum drainage basin size for the application of the modified rational method in order to obtain the frequency law of flows with a certain degree of precision, an analysis has been made of the distribution of mean error value according to area (Fig. 11 right). It has been seen that for basin surfaces of up to 500 km² the mean error is more or less stable and is around 20-25%. For larger basins the error is seen to rise in a constant manner. These results have been taken into account in the preparation of the maximum flow map by applying the rational method to those points on the river system network with drainage basins of up to 500 km² in area.

The calibration and validation process was then extended by incorporating the selected stations for the preparation of the maximum flow map. From this analysis it was seen that while the regional map of β factors correctly reflected the variations of this factor throughout the national territory in accordance with the climatic characteristics of each zone, in some cases there were considerable deviations with respect to the values obtained at the gauging stations. As such, the values provided by this map should only be taken as guidelines and not as exact values. It was subsequently considered

en cada una de las regiones estadísticas definidas anteriormente, en lugar de proporcionar un valor concreto en cada punto del territorio a través de un mapa. La mencionada caracterización estadística se ha realizado determinando para cada región el valor medio de los valores del coeficiente β obtenidos al calibrar el método racional en cada una de las estaciones de la región (figura 10 derecha), así como una valoración de la dispersión del conjunto de valores en torno a su valor medio a través de los intervalos de confianza del 50, 67 (relacionado con el error estándar) y 90%.

Por último, como se ha dicho, la calibración se ha llevado a cabo atendiendo principalmente al periodo de retorno de 10 años. Al aplicar dicho coeficiente al cálculo de otros periodos de retorno, se observa que los resultados no son del todo satisfactorios, no existiendo coincidencia en la forma funcional de la ley de frecuencia obtenida a partir de los datos de aforos mediante métodos estadísticos con la calculada mediante el método racional. De esta forma, ha sido necesario determinar unos factores correctores del coeficiente β en función del periodo de retorno, de tal manera que la forma funcional de la ley de frecuencia calculada mediante la aplicación del método racional coincida con la obtenida mediante procedimientos estadísticos.

4.2. Estimación en cuencas grandes: método estadístico

Para la estimación de los cuantiles en los puntos no aforados de la red fluvial con cuencas superiores a los 500 km², se han aplicado modelos estadísticos de regionalización. Este tipo de modelos pueden ser la única opción viable en aquellos puntos con una importante área vertiente en los que no es adecuada la aplicación de métodos hidrometeorológicos, debido a las hipótesis y simplificaciones que asumen en su planteamiento. El método estadístico utilizado consiste en definir ecuaciones de regresión a nivel regional que permiten estimar los cuantiles en los puntos no aforados a partir de distintas características de las cuencas vertientes, método que ha sido utilizado en EE.UU. por el U.S. Geological Survey (USGS) para el cálculo de los caudales máximos. En la selección de este tipo de procedimiento se han tenido en cuenta las ventajas que ofrece para considerar tanto la información sobre avenidas históricas como la correlación entre estaciones, además de la amplia experiencia existente en su aplicación a escala nacional para el estudio de los caudales máximos.

preferable and more in accordance with their real variability, to statistically characterize the values of the β correction factor in each of the regions that had been statistically defined beforehand, instead of providing a specific value at each point of the territory by a map. This statistical characterization has been made by establishing for each region the average value of the β factor values obtained by calibrating the rational method in each of the stations within the region (Fig. 10 right), as well as an evaluation of the dispersion of the combined values with respect to the average value by means of confidence intervals of 50, 67 (related to the standard error) and 90%.

Finally, and as indicated above, the calibration was carried out primarily in view of a 10-year return period. On applying the said coefficient to the calculation of other return periods, the results are not seen to be totally satisfactory, there being no coincidence in the functional form of the law of frequency obtained from the gauging data by statistical methods with that calculated by the rational method. It has subsequently been necessary to establish adjustment coefficients of the β correction factor in accordance with the return period, so that the functional form of the law of frequency calculated by the application of the rational methods coincides with that obtained by statistical processes.

4.2. Estimation in large basins: statistical method

Regional statistical models have been employed to estimate the quantiles at ungauged points on the river system network with basins of over 500 km². These types of models may be the only feasible option on those points with a large drainage area where hydrometeorological methods cannot be applied due to the hypotheses and simplifications assumed in the same. The statistical method employed consisted of the definition of regional regression equations that would allow the estimation of the quantiles at ungauged points on the basis of the different characteristics of the drainage basins, this being a method employed in the United States by the U.S. Geological Survey (USGS) for the calculation of maximum flows.

Se ha comprobado que las regiones con homogeneidad estadística definidas con anterioridad presentan un buen comportamiento a la hora de ser empleadas para determinar los modelos regionales de regresión, permitiendo la definición de modelos con errores moderados y elevados coeficientes de determinación. No obstante, en algunos casos se ha decidido crear subregiones para obtener un mejor ajuste de las ecuaciones de regresión.

En el estudio de los caudales máximos es habitual emplear ecuaciones de regresión logarítmicas, que se pueden expresar mediante ecuaciones potenciales relacionando la variable dependiente a regionalizar (Y_T), que representa los cuantiles de caudal máximo, con las variables explicativas (X_i) o características físicas de la cuenca vertiente y los coeficientes de regresión (b_i), que han sido ajustados por el método de los mínimos cuadrados ordinarios.

$$Y_T = 10^{b_0} \cdot X_1^{b_1} \cdot X_2^{b_2} \cdot \dots \cdot X_n^{b_n}$$

Antes de definir el modelo es necesario determinar las variables explicativas que puedan tener mayor influencia en el valor de los cuantiles y, por tanto, que sean más adecuadas para explicar la variabilidad de los mismos entre los distintos puntos de medida, seleccionando aquellas variables que presentan un grado de correlación significativo con los cuantiles y que, al mismo tiempo, no están correlacionadas entre sí (figura 12). Se han considerado variables fisiográficas (área de la cuenca (A), área drenante con altitud mayor de 1500 m, perímetro de la cuenca, longitud del río principal, pendiente del río principal, altitud media de la cuenca (H) y pendiente media de la cuenca), climáticas (cuantiles de precipitación máxima diaria (P_T), precipitación media anual) y características del tipo y uso del suelo en la cuenca (umbral de escorrentía, tasa mínima de infiltración). De entre todos los modelos posibles se han seleccionado aquellos que presentan un mejor comportamiento, en función de los valores que adoptan distintos estadísticos (error estándar de regresión (e), coeficiente de Mallow, error cuadrático de predicción y coeficiente de determinación corregido (R^2_{corr})), y que además tienen sentido físico a la hora de representar el comportamiento hidrológico de la cuenca (tabla 2).

Para obtener los caudales en los puntos de las corrientes principales correspondientes a grandes cuencas vertientes no se han empleado las ecuaciones de

It has been show that regions with a previously defined statistical homogeneity reveal good behaviour when used to determine the regional regression models and allow the definition of models with moderate errors and high coefficients of determination. However, in some cases it was decided to create sub-regions in order to obtain a better adjustment of the regression equations.

In the study of maximum flows it is common to use logarithmic regression equations which may be expressed by potential equations that relate the dependent variable to be regionalized (Y_T), that represents the quantiles of maximum flow, to the explanatory variable (X_i) or the physical characteristics of the drainage basin and the regression coefficients (b_i), adjusted by the ordinary least squares method.

$$Y_T = 10^{b_0} \cdot X_1^{b_1} \cdot X_2^{b_2} \cdot \dots \cdot X_n^{b_n}$$

Prior to defining the model it is necessary to determine the explanatory variables that may have the greatest influence on the value of the quantiles and that, subsequently, provide a more suitable explanation of the variability of the same between the different measuring points, and selecting those variables that show a significant degree of correlation with the quantiles while, at the same time, not being correlated between themselves (Fig. 12). Consideration has been given to geographical variables (basin area (A), drainage area with heights of over 1500 m, basin perimeter, length of main river, gradient of main river, average height of basin (H) and average gradient of basin), climate variables (maximum daily rainfall quantiles (P_T), average annual rainfall) and characteristics of the type and use of soil in the basin (runoff threshold, minimum infiltration rate). Out of all the possible models, we have selected those that represent the best behaviour in terms of the statistical values adopted (standard regression error (e), Mallow coefficient, square forecasting error and the corrected coefficient of determination (R^2_{corr})), which also have a physical rationale when representing the hydrological behaviour of the basin (Table 2).

The multiple regression equations described earlier were not used to obtain the flows at the

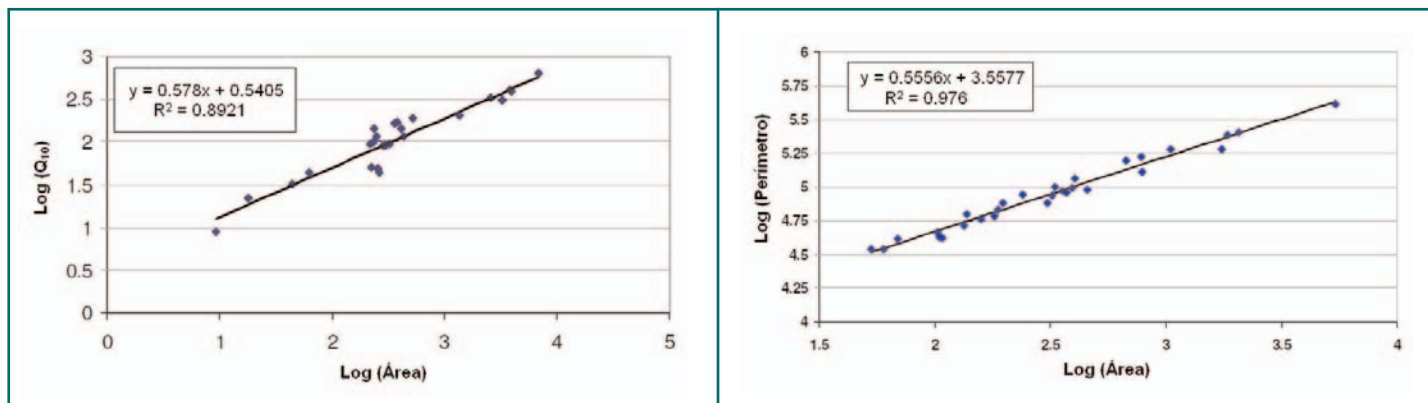


Fig. 12. Ejemplo de correlación de la variable área con el cuantil de caudal máximo anual de 10 años de periodo de retorno (izquierda), y entre las variables área y perímetro de la cuenca (derecha)/Example of the correlation of the variable basin area with the quantile of the maximum annual flow for a 10-year return period (left) and between the basin area and perimeter variables (right).

regresión múltiple descritas anteriormente, ya que las relaciones entre los caudales y las características medias de las cuencas podrían diferir de las correspondientes a cuencas de menor tamaño e, incluso, el valor medio en la cuenca de alguna de las variables podría no ser significativo a la hora de determinar el caudal cuando se trata de cuencas de gran tamaño. De esta manera, y puesto que en estos tramos de río se dispone de varias estaciones de aforos con series suficientemente largas y fiables a lo largo de los ejes principales, algunas de ellas con información histórica, se ha decidido establecer simplemente unas relaciones sencillas en función del área vertiente al punto, que permitan interpolar y dar continuidad a los cuan-

points of the main currents corresponding to large drainage basins, as the relations between the flows and the average characteristics of the basins may differ from those corresponding to smaller basins and the average value in the basin of some of these variables may be insignificant when establishing the flow in the case of the larger basins. As such, and when considering that various gauging stations have been placed in these sections of the river and that these have sufficiently long and reliable series throughout the main currents, some of these also including historical information, it has simply been decided to establish simple relations in accordance with the drainage area at the point, which then

Tabla 2. Ejemplos de modelos de regresión correspondientes a la región 32 de la cuenca del Tajo y al curso principal del río Duero (región 26).
Table 2. Regression models for region 32 in Tajo basin and for region 26 or main current of river Duero.

Región 32			Región 26	
$Q_2 = 10^{-4.3438} A^{0.6792} P_2^{0.8427} H^{0.9362}$	$R^2_{\text{corr}} = 0.87$	e = 34	$Q_2 = 0.0168 \cdot A^{1.0155}$	$R^2_{\text{corr}} = 0.99$
$Q_5 = 10^{-2.9825} A^{0.6569} P_5^{0.8573} H^{0.5532}$	$R^2_{\text{corr}} = 0.91$	e = 29	$Q_5 = 0.0246 \cdot A^{1.0388}$	$R^2_{\text{corr}} = 0.99$
$Q_{10} = 10^{-2.3975} A^{0.6419} P_{10}^{0.7751} H^{0.4446}$	$R^2_{\text{corr}} = 0.92$	e = 29	$Q_{10} = 0.0311 \cdot A^{1.0459}$	$R^2_{\text{corr}} = 0.99$
$Q_{25} = 10^{-1.7958} A^{0.6228} P_{25}^{0.6470} H^{0.3687}$	$R^2_{\text{corr}} = 0.92$	e = 28	$Q_{25} = 0.0409 \cdot A^{1.0514}$	$R^2_{\text{corr}} = 0.99$
$Q_{100} = 10^{-0.8303} A^{0.6009} P_{100}^{0.4582} H^{0.2288}$	$R^2_{\text{corr}} = 0.93$	e = 29	$Q_{100} = 0.0588 \cdot A^{1.0561}$	$R^2_{\text{corr}} = 0.99$
$Q_{500} = 10^{0.3026} A^{0.5748} P_{500}^{0.1928} H^{0.1032}$	$R^2_{\text{corr}} = 0.92$	e = 31	$Q_{500} = 0.0862 \cdot A^{1.0593}$	$R^2_{\text{corr}} = 0.99$

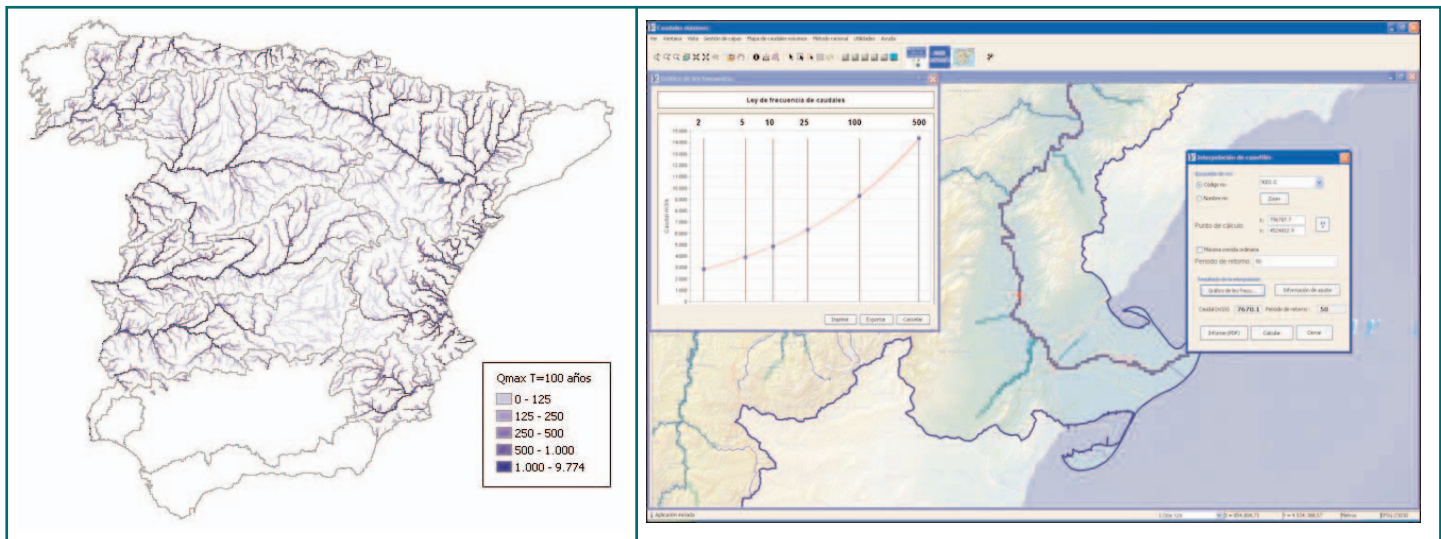


Fig. 13. Izquierda: Mapa de caudales máximos para el periodo de retorno de 100 años. Derecha: Aplicación informática CauMax para visualización y consulta del mapa de caudales máximos/Left: Map of maximum flows for a 100-year return period. Right: CauMax computer application for the visualization and consultation of the maximum flow map.

tiles calculados en las distintas estaciones de aforo presentes a lo largo del río (tabla 2).

5. Resultados

Los mapas obtenidos mediante la aplicación del método racional y las ecuaciones de regresión han sido objeto de ajustes posteriores con el objetivo de garantizar la adecuada coherencia de los resultados entre todos los puntos de la red fluvial. Como resultado del trabajo se ha obtenido una colección de mapas que proporcionan una estimación de los caudales máximos correspondientes a distintos periodos de retorno en régimen natural en cualquier punto de la red fluvial con una cuenca superior a 50 km² (figura 13 izquierda).

Con el objetivo principal de facilitar la consulta de los mapas de caudales máximos para los distintos periodos de retorno se ha desarrollado una herramienta informática (CauMax) que permite visualizar y consultar dichas capas en el entorno de un SIG (figura 13 derecha). Junto con la consulta de los mapas de caudales máximos, la aplicación informática permite realizar estimaciones de los cuantiles correspondientes a otros periodos de retorno mediante interpolación entre los valores facilitados por los mapas. Esta herramienta permite también obtener un valor aproximado para la máxima crecida ordinaria, a la que previamente se le ha asignado un periodo de retorno de forma regional siguien-

allow interpolation and the continuity of the quantiles calculated at the different gauging stations throughout the river (Table 2).

5. Results

The maps obtained by the application of the rational method and the regression equations have subsequently been subject to adjustment in order to guarantee a suitable coherence in results between all the points on the river system network. As a result of this work we have obtained a collection of maps that provide an estimation of the maximum natural flows for different return periods on any point of the river system network with basins of over 50 km² (Fig. 13 left).

A computer application (CauMax) has been developed to aid the consultation of maximum flow maps for different return periods which allows the visualization and consultation of these layers in a GIS environment (Fig. 13 right). In addition to aiding the consultation of maximum flow maps, the computer application allows the estimation of quantiles corresponding to other return periods by means of interpolation between the values provided by the maps. This application also provides an approximate value for the maximum ordinary flood which has previously been assigned a return period in a regional manner in accordance with the

do la metodología expuesta en (11) (figura 4 derecha). Asimismo, permite aplicar el método racional modificado en los puntos de la red fluvial con cuencas inferiores a 50 km², obteniendo las variables necesarias de forma automática a partir de las coberturas incluidas en la aplicación informática¹.

6. Agradecimientos

Este trabajo se ha llevado a cabo dentro del Convenio de colaboración "Asistencia técnica, investigación y desarrollo tecnológico en materia de gestión del dominio público hidráulico y explotación de obras", firmado entre la Dirección General del Agua y el CEDEX. Los autores desean expresar su agradecimiento a D. Jesús Yagüe Córdova y D. Javier Sánchez Martínez, de la Subdirección General de Gestión Integrada del Dominio Público Hidráulico, por sus aportaciones al trabajo y su labor de coordinación y dirección. Asimismo, se quiere agradecer la labor de revisión llevada a cabo por las distintas Confederaciones Hidrográficas, así como sus sugerencias y comentarios para mejorar la precisión de los resultados en determinadas zonas. ♦

methodology indicated in (11) (figure 4 right). This also allows the application of the modified rational method at those points on the river system network with basins of under 50 km² and automatically obtaining the necessary variables from the cover included in the computer application¹.

6. Acknowledgements

This work has been carried out under the collaboration agreement for "Technical assistance, research and technological development of public water management and exploitation of waterworks", signed between the Spanish Water Directorate and CEDEX. The authors wish to express their acknowledgement to Jesus Yagüe Córdova and Javier Sánchez Martínez, of the Sub-directorate for Integral Management of Hydraulic Public Domain for their contributions and their work of direction and coordination, and to acknowledge the revision work conducted by the different River Basin Authorities, as well as their suggestions and comments to improve the accuracy of the results in certain areas. ♦

(1) La aplicación informática CauMax puede descargarse gratuitamente de la página web del Ministerio de Agricultura, Alimentación y Medio Ambiente (<http://www.magrama.es/es/agua/temas/gestion-de-los-riesgos-de-inundacion/snczi/Mapa-de-caudales-maximos/>), o desde el CEDEX mediante el siguiente enlace: http://hercules.cedex.es/caumax/caumax_v2.rar.

(1) *The CauMax computer application may be downloaded free of charge from the web page of the Ministerio de Agricultura, Alimentación y Medio Ambiente (<http://www.magrama.es/es/agua/temas/gestion-de-los-riesgos-de-inundacion/snczi/Mapa-de-caudales-maximos/>) or from CEDEX at the following link: http://hercules.cedex.es/caumax/caumax_v2.rar.*

Referencias/References:

- (1) U.S. Water Resources Council. "Guidelines for determining flood flow frequency". Bulletin 17B. 1981.
- (2) Hosking, J.R.M.; Wallis, J.R. "Regional frequency analysis". Cambridge University Press. 1997. 224 pp. ISBN 0-521-43045-3.
- (3) Jiménez, A.; Mediero, L. "Caracterización del comportamiento estadístico de los caudales máximos anuales y estacionales de los ríos de la España peninsular. Análisis y propuesta de procedimientos para su determinación". Proc. Jornadas de Ingeniería del Agua 2009. Madrid. ISBN 978-90-78046-09-7.
- (4) Wiltshire, S.E. "Identification of homogeneous

- regions for flood frequency analysis". Journal of Hydrology. 84 (1986), pp. 287-302. 1986.
- (5) Hosking, J.R.M. "L-moments: Analysis and estimation of distributions using linear combinations of order statistics". J. Royal Statistical Soc., B. Vol. 52. nº2. pp. 105-124. 1990.
- (6) Témez, J.R. "Extended and improved rational method. Version of the highways administration of Spain". Proc. XXIV Congress. Madrid. Vol. A. pp 33-40. 1991.
- (7) Ferrer-Juliá, M.; Estrela, T.; Sánchez del Corral Jiménez, A.; García-Meléndez, E. "Generación de un mapa de valores continuos de número de curva basado en la conductividad hidráulica saturada". Revista de Ingeniería Civil. nº 131. 3er trimestre. pp. 181-187. 2003.

- (8) Dirección General de Carreteras. "Máximas lluvias diarias en la España peninsular". Serie monografías. Ministerio de Fomento. 1999.
- (9) Hawkins, R.H.; Hjelmfelt, A.T.; Zevenbergen, A.W. "Runoff probability storm depth and curve numbers". Journal of the Irrigation and Drainage Division, 111(4). pp. 330-340. 1985.
- (10) Ponce, V.M.; Hawkins, R.H. "Runoff curve number: has it reached Maturity?". Journal of Hydrologic Engineering, January 1996. pp 11-19.
- (11) CEDEX. "Aspectos prácticos de la definición de la máxima crecida ordinaria. Asistencia técnica para la ordenación de cauces y márgenes inundables". Informe para la Dirección General de Calidad de las Aguas. 1994.