

LA CALIBRACION AUTOMATICA DE PARAMETROS EN LOS MODELOS HIDROLOGICOS DETERMINISTICOS

TEODORO ESTRELA MONREAL (*)

RESUMEN. En la implementación de los modelos hidrológicos determinísticos una de las fases que siempre ha suscitado un mayor interés ha sido la calibración de los parámetros. En los últimos años, en un intento de facilitar y objetivizar este proceso, las técnicas de estimación automática de parámetros han experimentado un avance importante. En este artículo se hace una revisión de su problemática, ilustrándose ésta mediante tests con series sintéticas de caudales. Para ello se ha utilizado un modelo determinístico de balance de humedad que opera a escala mensual. Con él se han generado, para unos parámetros prefijados y a partir de registros de precipitación y estimas de la evapotranspiración potencial, datos de caudales libres de error. Posteriormente esos datos se han contaminado con diferentes tipos de error y se ha procedido al ajuste del modelo utilizando criterios objetivos tales como mínimos cuadrados (SLS), mínimos cuadrados generalizados (GLS) y mínimos cuadrados ponderados (WLS). Finalmente se han calibrado los parámetros del modelo en un caso real.

ABSTRACT. *Parameter calibration is one of the more interesting stages in the hydrologic models implementation. Trying to ease and objectivize this process we have to say that the automatic estimation techniques have experimented an important development in the latest years. In this article we make a brief review of the problems that arise from automatic estimation techniques taking as an example for it some tests with synthetic runoff data. A water balance model, operating on a monthly base and for parameters values given, has been used to generate synthetic error-free runoff data from precipitation records and potential evapotranspiration estimates. Those values being contaminated with different kinds of error. Afterwards the model has been fitted according to objective criterions such as: Simple Least Squares (SLS), Generalized Least Squares (GLS) and Weighted Least Squares (WLS). Finally, parameter values have been calibrated in a real case.*

INTRODUCCION

Los modelos determinísticos y los estocásticos forman los dos grandes grupos en los que se dividen los modelos matemáticos utilizados en Hidrología (Clarke, 1973). En los determinísticos las variables vienen determinadas por leyes físicas consideradas como exactas y que explican su variabilidad, mientras que en los estocásticos las variables son regidas, en todo o en parte, por las leyes del azar y, por tanto, caracterizadas en términos de probabilidad. Los modelos determinísticos suelen tener un mayor potencial que los estocásticos, generalmente cajas negras, para predecir los efectos de cambios en el uso de la cuenta o para simular caudales en cuencas no aforadas, dada la posible relación de sus parámetros con las características físicas de las cuencas (Mein and Brown, 1978). Este tipo de modelos ha sido tradicionalmente utilizado por los hidrólogos con un mayor agrado, al reflejar los procesos esenciales del ciclo hidrológico.

Las fases en las que puede dividirse el proceso de implementación de un modelo matemático determinístico son: a) formación del modelo conceptual, b) construcción del modelo, c) calibración de los parámetros del modelo y d) validación del modelo.

Por formación del modelo conceptual se entiende la idea, fundamental cualitativa, que tiene el hidrólogo del funcionamiento del sistema. La etapa de construcción del modelo es aquella en la que se forman las ecuaciones que describen el modelo conceptual, se establecen las relaciones entre ellas y se procede a su resolución. La calibración es la etapa del proceso de implementación que obliga a que el funcionamiento del modelo matemático reproduzca la historia del sistema real estudiado y es seguida por la validación, con lo cual se comprueba el adecuado funcionamiento del modelo con una serie de observaciones no utilizadas en la calibración.

Una de las etapas más controvertidas del proceso de implementación del modelo es la calibración. En muy pocas ocasiones se describen las propiedades de los residuos o diferencias entre los valores simulados por el modelo y los valores observados. En la mayoría de los

(*) Centro de Estudios Hidrográficos.

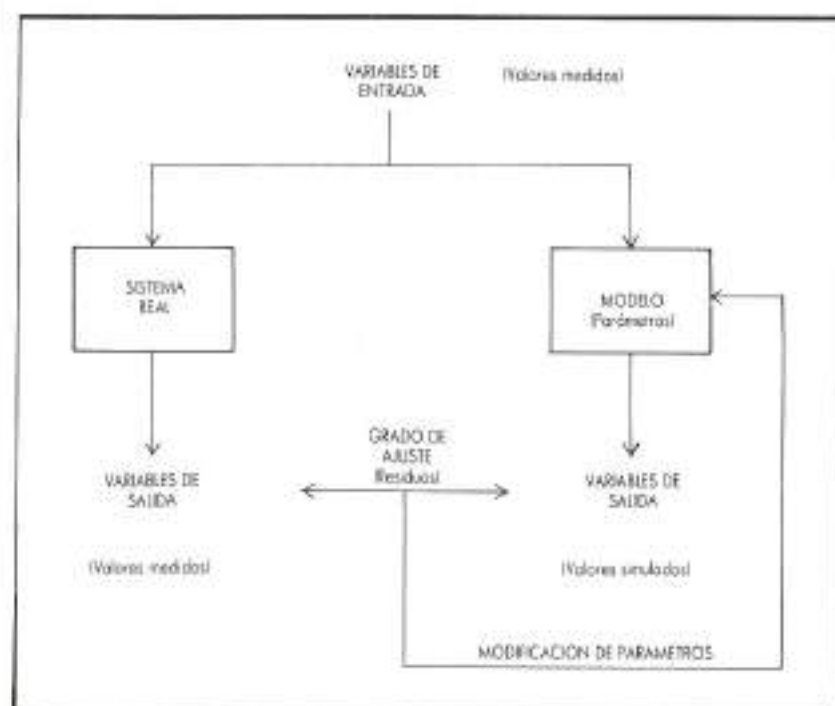


FIGURA 1. Proceso de calibración.

estudios únicamente se presentan gráficos comparando las series simuladas y observadas de caudales y alguna medida de la bondad del ajuste. Este artículo pretende profundizar en la fase de calibración haciendo una revisión del tipo de errores habituales en la modelación lluvial-caudal, de las funciones de error u objetivo utilizadas para su tratamiento y de las técnicas de optimización que minimizan esas funciones, para seguidamente analizar la problemática de la optimización mediante una serie de tests con datos sintéticos y reales.

CONCEPTO DE CALIBRACION: MANUAL Y AUTOMATICA

La calibración consiste en la determinación de los valores de los parámetros del modelo que conducen a una reproducción adecuada del funcionamiento observado del sistema y son compatibles con el rango de validez física de los mismos. Algunos parámetros se pueden estimar mediante el estudio de las características de la cuenca, pero otros, generalmente relacionados con los subprocesos internos que tienen lugar en la cuenca, no pueden ser determinados de esa manera y, por tanto, deben ser estimados indirectamente en la fase de calibración. Aun en la situación hipotética de tener información sobre todos los parámetros, no sería conveniente proceder a la utilización del modelo sin haberse asegurado de que todos los datos disponibles (variables y parámetros) son cualitativa y cuantitativamente compatibles.

La calibración es «per se» un proceso iterativo, en el que se empieza con los valores medidos o supuestos de

los parámetros, y se van modificando sucesivamente a fin de reducir las diferencias entre los valores simulados y observados de las variables de control (figura 1). Este proceso se puede hacer manual o automáticamente.

Cuando se calibra manualmente un modelo es habitual utilizar gráficos o mapas de los valores simulados y observados, para poder deducir a partir de ellos cuáles son los parámetros que deben modificarse y en qué dirección deben hacerlo (figuras 2a y 2b). Este procedimiento es muy subjetivo y, dependiendo de quien sea el modelador, conducirá a unos resultados u otros, aunque presenta la ventaja de que aquél, durante ese proceso, consigue una gran familiaridad con el papel y sensibilidad de cada parámetro. Por otra parte, la calibración manual requiere un esfuerzo de tiempo considerable, sobre todo cuando el número de parámetros o variables de control es muy elevado.

Otra posibilidad es la calibración automática, donde mediante la utilización de métodos numéricos se obtienen los valores de los parámetros que minimizan los residuos o diferencias entre los valores simulados y observados. Cuando se utilizan este tipo de técnicas es conveniente, en cualquier caso, visualizar gráficamente los resultados y hacer un análisis de sensibilidad de los mismos respecto a los parámetros.

ANALISIS DE ERRORES

La fase de calibración tiene como objetivo minimizar los errores del modelo. Antes de plantear la función de error, más comúnmente llamada función objetivo, es

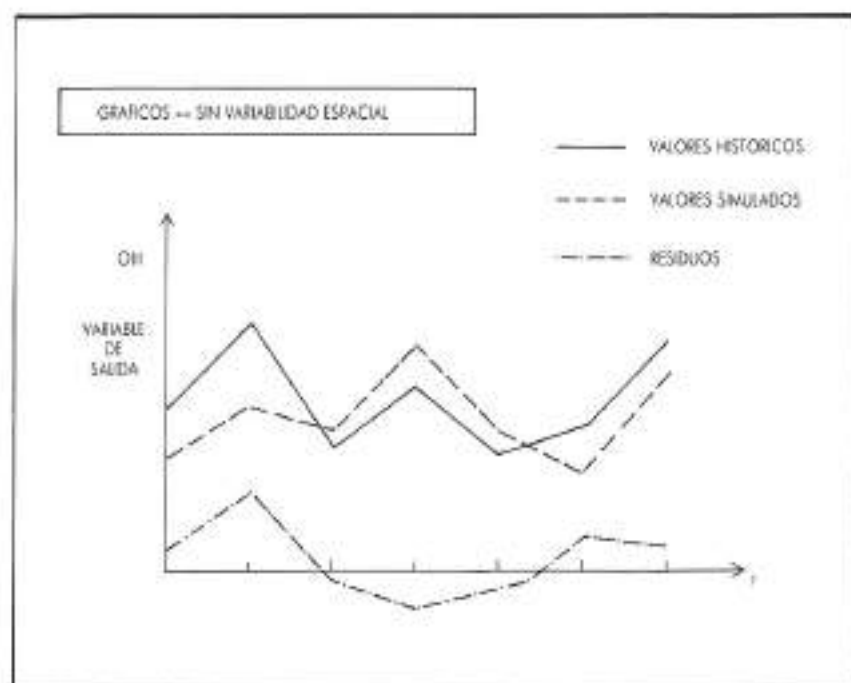


FIGURA 2a. Gráfico de calibración. Distribución temporal de una variable.

conveniente analizar los tipos de errores habituales en la modelación hidrológica.

Sea un modelo de transformación lluvia-caudal con K parámetros $\theta_1, \theta_2, \dots, \theta_K$, que simula caudales sobre n periodos sucesivos de tiempo (ej.: días, semanas, meses). Expresado en forma matemática:

$$q_t^* = q_t(\theta) + e_t(\theta) \quad 1 \leq t \leq n \quad [1]$$

donde q_t^* es el caudal observado para el periodo de tiempo t , $q_t(\theta)$ es el simulado por el modelo con el conjunto de parámetros θ y $e_t(\theta)$ es el residuo.

Las fuentes posibles de error son las siguientes:

- La falta de adecuación del propio modelo, que al ser una simplificación de la realidad siempre tiene asociado un cierto error estructural.
- Los datos. Pueden existir errores en los datos de entrada (ej.: precipitaciones, evapotranspiraciones potenciales, etc.), lo cual afectará a la salida del modelo. Por otra parte, los caudales observados con los cuales se compara los resultados del modelo tendrán también un cierto grado de error.

Sería deseable que estos errores cumplieren una serie de requisitos:

1. Que los errores $e_t(\theta)$ sean estacionarios con media, $E(e_t) = 0$ y varianza constante $E(e_t^2) = \sigma_e^2$.

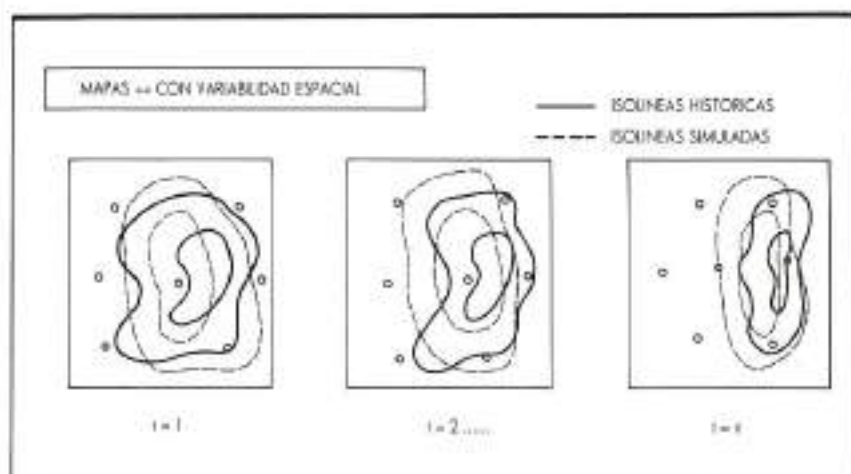


FIGURA 2b. Mapas de calibración. Distribuciones espacial y temporal de muchas variables.

2. Que los errores no presenten autocorrelación temporal.

$$E(e_t e_{t-1}) = 0 \quad \text{para todo } t \neq 0$$

3. Que los errores estén normalmente distribuidos.

$$e_t \in N(0, \sigma_e)$$

Si los errores presentasen esa clase de estructura sería obvio que un simple número, σ_e^2 , resumiría como de segura sería la predicción con nuestro modelo. Desafortunadamente, en muchas situaciones los errores no obedecen esas hipótesis.

Una examen del gráfico de los errores como una función de las predicciones revela, en ocasiones, que éstos muestran frente a las predicciones una dependencia en la media y que la variabilidad de los mismos (varianza) no es la misma para cualquier magnitud del caudal predicho. Los cambios en la varianza suelen estar asociados al hecho de que a los eventos mayores le corresponden errores de predicción también mayores. Ello se debe, entre otras causas: a) la falta de adecuación del modelo para representar esos eventos, b) la forma de la curva de gasto (relación nivel-caudal) y la dificultad en la estimación de su tramo alto, c) la propagación del error a partir de los datos de precipitación; un mismo porcentaje de error afecta menos a los caudales predichos cuando el valor de la precipitación es menor, etc. Por otra parte, si se examina el gráfico de evolución temporal de los residuos se observa que en los períodos de recesión suelen aparecer secuencias largas de residuos con el mismo signo. Aparte de la explicación obvia de la inadecuada modelación del proceso de transformación lluvia-caudal, un factor que también puede contribuir a ello es la correlación entre caudales, la cual se introduce automáticamente cuando éstos son traducidos de una curva de gasto errónea (Clarke, 1973).

Tanto la falta de estacionariedad como la autocorrelación serial aumentan a medida que el intervalo entre medidas sucesivas se hace más corto. A escalas temporales pequeñas (horas, días) estos efectos son mucho más apreciables que a escalas superiores (mes). Ello se debe a la suavización que tiene lugar cuando se agregan y promedian los datos temporalmente.

La hipótesis de normalidad es difícil de satisfacer. Los residuos suelen presentar distribuciones algo sesgadas, explicables en parte por la propia distribución sesgada que siguen los caudales. En el caso de que se requieran regiones de confianza para las estimas de los parámetros (Clarke, 1973), será preciso hacer hipótesis sobre la distribución de probabilidad de los residuos.

CRITERIOS O FUNCIONES OBJETIVO

El análisis de los tipos de errores en la modelación lluvia-caudal ha propiciado la aparición de diversos tipos de funciones objetivo. Son expresiones matemáticas que miden el grado de ajuste entre los valores simulados y los observados, como una función de la secuencia temporal de los residuos.

Uno de los criterios más simple y utilizado hasta la fecha ha sido el de mínimos cuadrados (SLS). Sea q_1^* , q_2^* , ..., q_n^* una secuencia de n caudales observados. Los valores del conjunto de parámetros θ serán aquellos que minimizan la función objetivo $U_n(\theta)$.

$$U_n(\theta) = n^{-1} \sum_{t=1}^n e_t(\theta)^2 \quad [2]$$

o lo que es lo mismo,

$$U_n(\theta) = n^{-1} \sum_{t=1}^n (q_t^* - q_t(\theta))^2 \quad [3]$$

Una forma más general que las expresiones [2] y [3] es:

$$U_n(\theta, j) = n^{-1} \sum_{t=1}^n |q_t^* - q_t(\theta)|^j \quad [4]$$

donde j es cualquier número positivo. Cuando $j = 1$, se tiene una función objetivo de mínimos absolutos y cuando $j = 2$ coincide con la expresión [3].

Un estudio de sensibilidad de la función objetivo respecto al exponente j fue hecho por Johnston and Pilgrim (1976) con un modelo de balance de humedad simplificado, concluyendo que los valores estimados de los parámetros no dependían del exponente, pero sí que lo hacía la forma de la superficie de la función objetivo y, por tanto, la facilidad o dificultad para encontrar el mínimo. Un valor del exponente $j = 2$ presentaba una superficie más fácil de optimizar.

Según la experiencia de algunos investigadores, el criterio de mínimos cuadrados da lugar a estimas insesgadas y de mínima varianza sólo si se satisfacen las hipótesis siguientes: a) los errores son estacionarios en media y varianza; b) los errores no presentan autocorrelación. Si éstos siguen además una distribución normal, entonces los parámetros serán estimas de máxima verosimilitud (ML).

Tradicionalmente, para manejar casos de error no estacionarios en varianza se han seguido dos caminos: a) usar transformaciones de los caudales que estabilicen la varianza (p. ej.: la log-transformación) y aplicar posteriormente el criterio SLS, b) utilizar directamente algún criterio de mínimos cuadrados ponderados (WLS).

Si se asume que los errores tienen media nula, no presentan autocorrelación y siguen una distribución normal, la función máxima verosimilitud (Sorooshian and Dracup, 1980) coincide con el criterio de mínimos cuadrados ponderados (WLS). Esto es:

$$U_n(\theta) = n^{-1} \sum_{t=1}^n w_t (q_t^* - q_t(\theta))^2 \quad [5]$$

donde w_t son los pesos, que deben ser iguales al inverso de la varianza del error (σ_t^{-2}). La dificultad del método estriba precisamente en la determinación de esos pesos. Generalmente se han estimado de forma subjetiva, como, por ejemplo, hace el modelo HEC-1 del U.S. Army Corps of Engineers (1973), donde los pesos se asignan de la forma siguiente:

$$w_t = \frac{q_t^* + \bar{Q}}{2 \bar{Q}} \quad t = 1, \dots, n \quad [6]$$

siendo $\bar{Q}$ el caudal medio de la serie histórica. Esta regla tiende a mejorar la reproducción de los caudales punta. Es digno de mencionar el hecho de que con este esquema de ponderación se está en conflicto directo con la experiencia, que dice que los caudales más altos llevan asociados mayores errores.

Otra forma sencilla de asignar los pesos sería suponer que la desviación típica del error, σ_t , es proporcional al valor observado, q_t^* .

$$\sigma_t = K q_t^* \quad [7]$$

Admitiendo que,

$$w_t = \sigma_t^{-2} \quad [8]$$

de las expresiones [7] y [8] se obtiene:

$$w_t = (K q_t^*)^{-2} \quad [9]$$

y la expresión [5] se convierte en:

$$U_n(\theta) = K^{-2} n^{-1} \sum_{t=1}^n \left(\frac{q_t^* - q_t(\theta)}{q_t^*} \right)^2 \quad [10]$$

En esencia, este criterio considera el error cuadrático relativo en lugar del error cuadrático absoluto.

Sooroshian y Dracup (1980) desarrollaron un estimador de máxima verosimilitud ML en el que los pesos se introducían de forma objetiva tras plantear relaciones más realistas entre la desviación típica del error, σ_t , y la magnitud del valor observado, q_t^* . Este estimador mostró su superioridad frente a los criterios tradicionales de mínimos cuadrados (SLS) y mínimos cuadrados ponderados (WLS).

En el caso de errores estacionarios en varianza, pero autocorrelacionados, hay bastante acuerdo por parte de la mayoría de los investigadores en las modificaciones a realizar en la función objetivo. Williams y Yeh (1983) sugirieron el uso del estimador por mínimos cuadrados generalizados (GLS) para eliminar el efecto de autocorrelación serial. Para ello se requiere que algún proceso razonable de autocorrelación sea presupuesto para los errores. Es habitual asumir un esquema autorregresivo

(Box y Jenkins, 1970) de primer orden para representar los errores. Esto es,

$$e_t = \Phi e_{t-1} + v_t \quad t = 1, \dots, n \quad [11]$$

donde Φ es el primer coeficiente de autocorrelación serial para los errores y $(v_1, v_2, \dots, v_n)^T$ es un vector de componentes aleatorias.

En este caso la nueva función objetivo adopta la expresión:

$$U_n(\theta) = n^{-1} \sum_{t=1}^n v_t^2 = n^{-1} \sum_{t=1}^n (e_t - \Phi e_{t-1})^2 \quad [12]$$

o lo que es lo mismo:

$$U_n(\theta) = n^{-1} \sum_{t=1}^n [(q_t^* - q_t(\theta)) - \Phi (q_{t-1}^* - q_{t-1}(\theta))]^2 \quad [13]$$

Cuando se presentan conjuntamente errores no estacionarios en varianza y autocorrelacionados puede utilizarse la expresión [13], ponderando los términos del sumatorio con diferentes pesos, w_t . Kuzcera (1982) utilizó una aproximación bayesiana para estimar los parámetros y su incertidumbre en este tipo de situaciones.

Es digno de mención el hecho de que la elección de una u otra función objetivo puede afectar a los valores estimados de los parámetros, ya que cada criterio pone un énfasis distinto en la diferencia entre los valores de los caudales observados y simulados. Diskin y Simon (1977) ya indican este hecho en un trabajo en el que comparan un gran número de funciones objetivo.

ALGORITMO DE OPTIMIZACION

La minimización de las funciones objetivo planteadas en el apartado anterior requiere la utilización de algoritmos numéricos de optimización. La elección de éstos ha sido discutida por muchos investigadores (Ibitt y O'Donnell, 1971; Johnston y Pilgrim, 1976; Sorooshian y Dracup, 1980; entre otros). Prácticamente hay coincidencia entre todos ellos en que las técnicas de búsqueda directa se han mostrado superiores a las técnicas de gradiente. En estas últimas la optimización de la función objetivo requiere la evaluación de las derivadas de ésta respecto a los parámetros, realizándose la búsqueda según la dirección del gradiente (figura 3a). La forma de la superficie de respuesta de las funciones objetivo utilizadas es generalmente compleja y la obtención de expresiones analíticas de las derivadas es tarea inabordable, requiriéndose la aplicación de métodos numéricos de perturbación para su obtención. Por otra parte, Johnston y Pilgrim (1976) señalan que los métodos basados en la evaluación de las derivadas se enfrentan con el problema de las discontinuidades en la superficie respuesta, lo cual afecta adversamente al funcionamiento de este tipo de métodos. Mediante ejemplos sencillos

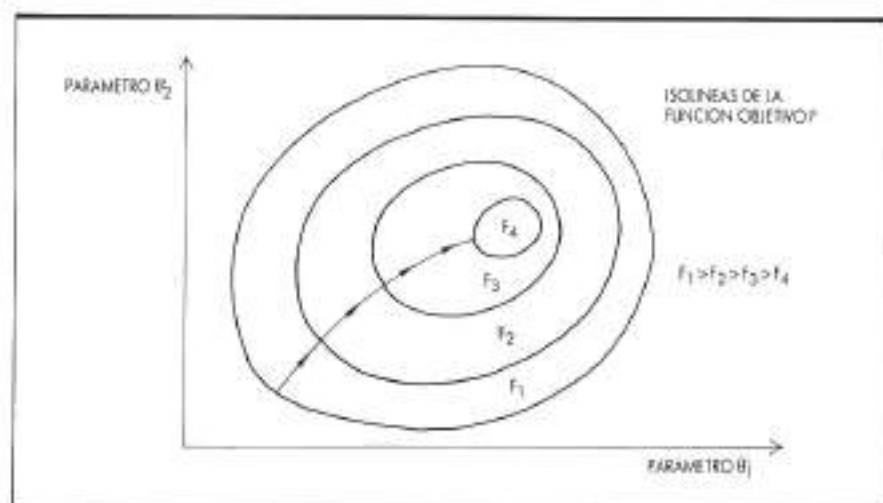


FIGURA 3a. Minimización en dos dimensiones. (Método del gradiente).

analizan cómo los umbrales que definen las capacidades de almacenamiento en los distintos elementos considerados (ej.: intercepción, almacenamiento en depresión, suelo, etc.) crean superficies discontinuas de la función objetivo.

El método de búsqueda directa más utilizado es el de búsqueda con rotación de ejes (Rosenbrock, 1960). Está basado en el de búsqueda univariada, en el cual se realiza la optimización sucesiva de la función objetivo, dejando en movimiento cada vez uno de los parámetros y fijando el resto. Cuando todas las «n» direcciones según los parámetros han sido investigadas sucesivamente, termina una iteración y la búsqueda es repetida empezando desde los mejores valores encontrados para los parámetros. La mayor debilidad del método de búsqueda univariada es que no puede optimizar satisfactoriamente problemas donde las curvas de nivel de la super-

ficie respuesta presentan formas alargadas e inclinadas respecto a los ejes de los parámetros. Tales situaciones son normales cuando existe algún grado de dependencia o correlación entre parámetros. El proceso degenera entonces en una búsqueda en zig-zag de progresión muy lenta hacia el óptimo (figura 3b). El método de búsqueda con rotación de ejes es similar en su primera fase, pero cuando se producen estructuras como la anteriormente comentada recurre a rotaciones de los ejes de los parámetros, o lo que es lo mismo, a realizar la búsqueda del mínimo de la función según combinaciones lineales de los parámetros. Este método progresa muy rápidamente en esas estructuras, evitando la debilidad del método de búsqueda univariada. En este trabajo se ha utilizado el algoritmo de optimización de Rosenbrock con pequeñas modificaciones operacionales relativas a los criterios de convergencia.

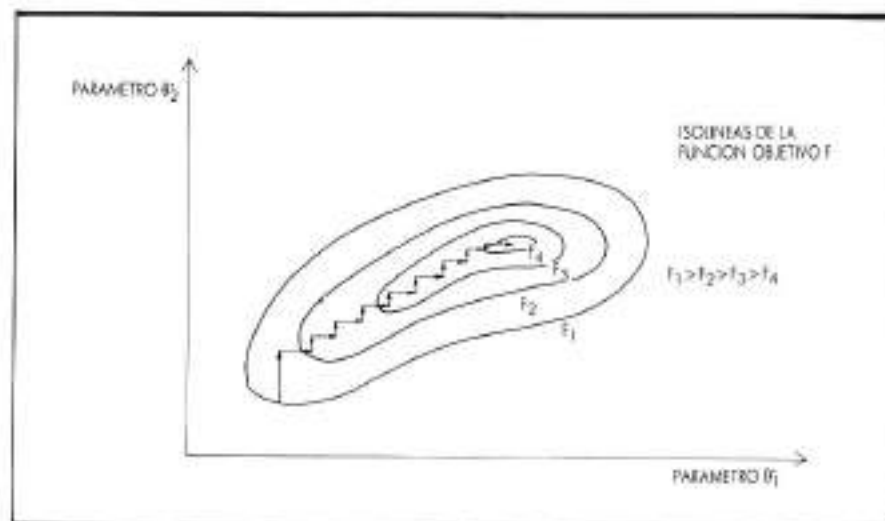


FIGURA 3b. Minimización en dos dimensiones. (Método de búsqueda univariada).

PROBLEMATICA DE LA OPTIMIZACION

Algunas de las dificultades que se encuentran los métodos de optimización en su búsqueda del mínimo de la función objetivo son:

A. Indiferencia de la función objetivo a los valores de los parámetros. Esto tiene lugar cuando las salidas del modelo y, por tanto, el valor de la función objetivo no es afectada por cambios en el valor de un parámetro, bien porque éste representa procesos que no son esenciales en la modelación, bien porque no es activado por la secuencia de datos utilizada en la calibración. Si al aumentar la secuencia de datos no se consigue activar el parámetro, ello revela que en la etapa de formación conceptual del modelo, el proceso introducido asociado a ese parámetro, tiene escasa importancia.

B. Interdependencia entre parámetros. Se produce cuando un cambio en el valor de un parámetro puede compensarse por cambios en uno o más del resto de parámetros. Un gran número de combinaciones de los valores de los parámetros dará valores similares en la función objetivo. Esta circunstancia es muy habitual en modelos de muchos parámetros, donde se crean valles de fondo alargado en la forma de la superficie de la función objetivo, que incrementan la dificultad en la optimización. El algoritmo de Rosenbrock intenta facilitar la optimización en esos casos. Sin embargo, este hecho suele indicar una formulación no adecuada en la estructura del modelo. La obtención de la matriz de correlación de las estimas de los parámetros puede ser muy útil cuando el número de parámetros es superior a dos y el dibujo de gráficos de la superficie respuesta es una tarea dificultosa.

El efecto de interdependencia (o correlación) en la modelación lluvia-caudal es discutido por Johnston y Pilgrim (1976) y Mein y Brown (1978). Sorooshian y Arfi (1982) proponen varias medidas de interacción o interdependencia de los parámetros.

C. Óptimos locales. Estos son puntos de la superficie de la función objetivo que tienen valores más bajos de ésta que cualquiera de los puntos que los rodea, pero más grandes que cualquier otro punto o puntos en otra región de la superficie respuesta. La búsqueda puede, por tanto, terminar en un punto que no es el óptimo absoluto. Este problema se presenta habitualmente en la modelación hidrológica y es más acusado a medida que crece el número de parámetros. La única solución para acercarse al mínimo absoluto es realizar un barrido repitiendo el proceso de optimización con «m» conjuntos de valores iniciales de los parámetros, e incluso en ese caso no se sabrá si se ha alcanzado el óptimo. El no alcanzarlo no es problemático si el ajuste es satisfactorio y las propiedades de los residuos son aceptables.

D. Identificabilidad y unicidad de la solución. La cuestión de la identificabilidad y unicidad de parámetros es considerada en detalle por Gupta y Sorooshian (1983) y Sorooshian, Gupta y Fulton (1983). Ellos argumentan que los parámetros umbrales, técnicamente identifica-

bles para un período de calibración suficientemente largo, pueden en algunos casos no ser activados por series más pequeñas de datos, haciendo que estimas únicas no sean posibles de obtener. Este hecho, claramente intuitivo, no parece ser la única causa. Troutman (1985) señala que el problema de la identificación y unicidad está estrechamente ligado con el de la correlación entre parámetros. Un ejemplo de un conjunto de parámetros que no son identificables sería b_1 y b_2 , si estos dos parámetros aparecieran en el modelo sólo como suma $b_1 + b_2$ o producto $b_1 b_2$. En ese caso no serían identificables, ya que se podrían modificar sus valores manteniendo la suma o el producto constante y no alterando, por tanto, los resultados de la predicción. Tal modelo se dice estar sobreparametrizado. En ocasiones, expresiones de los parámetros más complejas, que en principio no parecen mostrar interdependencia, pueden enmascarar relaciones como las arriba expuestas para ciertas secuencias de datos.

E. Estimaciones de parámetros fuera de rangos físicos aceptables. Este problema podría resolverse planteando algoritmos de optimización con restricciones, aunque ello no es aconsejable, pues cuando se produce suele indicar que: a) la estructura del modelo es incorrecta, b) el criterio objetivo no es adecuado, c) los datos de calibración presentan fuertes incoherencias entre sí. Si los valores óptimos de los parámetros no estuviesen dentro del subespacio de los parámetros considerado como factible, habría que cuestionar y analizar las situaciones anteriores. Troutman (1985) indica que los errores en los datos de entrada parecen ser la causa más relevante de las estimas no realistas de los parámetros.

F. Escasez de datos de calibración. Muchos investigadores han sugerido utilizar series muy largas para mejorar los resultados de la calibración. Sin embargo, no es tanto la longitud de las series de datos, sino la información contenida en ellas (variabilidad hidrológica) lo que es importante.

Toda esta problemática, que no sólo se plantea en la calibración automática, sino también en la manual, puede dificultar la validación del modelo e inducir a error si se está inmerso en un proceso de regionalización de parámetros. Una calibración rigurosa de los mismos dirigirá hacia una interpretación más realista de sus valores, lo que facilitará el establecimiento de correlaciones con las características físicas de las cuencas.

CASO DE ESTUDIO SINTETICO

En este apartado se presenta un caso de estudio con datos sintéticos que ilustra la revisión realizada en apartados anteriores sobre el análisis de errores, criterios objetivos y problemática de la optimización.

Se ha generado una serie sintética de caudales libres de error con un modelo determinístico en el que se ha prefijado un conjunto de parámetros y se ha introducido una secuencia de datos meteorológicos (precipitaciones y evapotranspiraciones potenciales).

El modelo formulado por Témez (1977) ha sido usa-

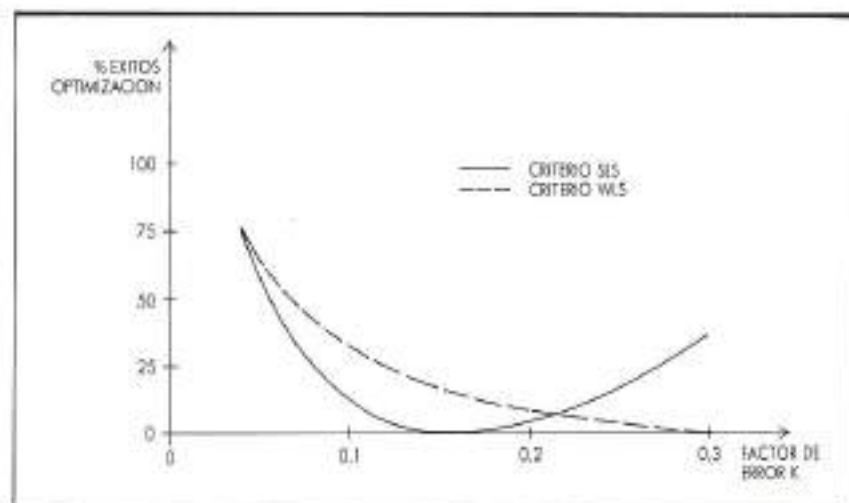


GRÁFICO 2. Caso de error estacionario y no correlacionado.

criterio WLS, ya que tiene en cuenta, en la ponderación que efectúa, la no estacionariedad del ruido y más en este caso, donde para contaminar los datos se ha utilizado una expresión equivalente a la que define el criterio WLS.

El gráfico 2 muestra que si bien con errores pequeños el criterio WLS funciona mejor que el SLS, cuando éstos crecen sucede lo contrario. Esto, que en un primer análisis resulta extraño, es explicable si se tiene en cuenta que este criterio considera errores relativos y a medida que el error crece tiende a sacrificar el ajuste de los valores altos incidiendo más en la reproducción de la forma de la secuencia de caudales, con lo cual empeora el funcionamiento del proceso de optimización y las estimas de los parámetros se alejan del óptimo.

En cualquier caso, ambos criterios (SLS y WLS) proporcionan peores resultados que cuando los errores son estacionarios. La aplicación de otros criterios de mini-

mos cuadrados ponderados más sofisticados, que planteen en detalle la función que relaciona la desviación típica del error y el caudal, podría proporcionar mejores resultados. No se pretende en este artículo realizar este análisis, sino únicamente poner de manifiesto que la aplicación del criterio WLS de forma sencilla, como habitualmente se hace, o bien lo que es en la práctica equivalente, realizar transformaciones logarítmicas de los caudales y aplicar el criterio SLS, puede dar lugar a estimas no adecuadas de los parámetros.

EFFECTO DEL CRITERIO OBJETIVO EN LA OPTIMIZACION CON SERIES SINTÉTICAS CONTAMINADAS CON RUIDOS ESTACIONARIOS AUTOCORRELACIONADOS

La comparación de los criterios SLS y GLS, cuando existe autocorrelación de la serie de residuos, indica que el criterio GLS es claramente superior (gráfico 3). Curiosamente se observa que cuando mayor es el coefi-

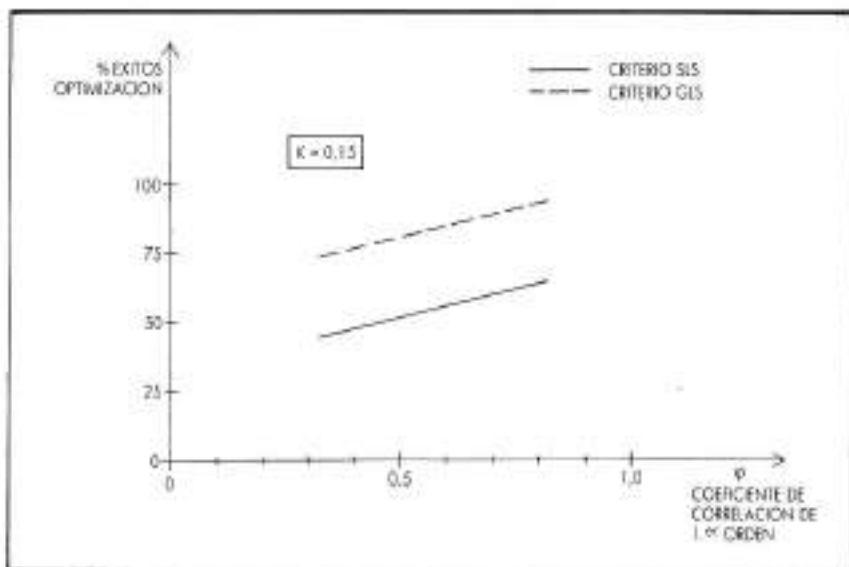


GRÁFICO 3. Caso de error estacionario y correlacionado.

ciente de autocorrelación se obtienen mejores resultados. Quizá ello pueda ser debido a que al existir una mayor correlación entre errores, el proceso de optimización gana en facilidad para identificar los subprocesos de reparto de flujo y, consecuentemente, los parámetros a ellos asociados.

CONCLUSIONES DEL CASO DE ESTUDIO SINTETICO

Los resultados obtenidos, aunque no deben generalizarse a todas las situaciones, pues responden a unos datos y valores de los parámetros concretos, sí permiten hacer las consideraciones siguientes: a) la existencia de errores estacionarios y no correlacionados en las series de caudales no produce un empeoramiento en el proceso de optimización, incluso para valores altos de los mismos; b) con errores no estacionarios ninguno de los criterios analizados funciona bien. El criterio WLS no debería ser aplicado sin un estudio previo para la determinación de los pesos. La introducción de éstos de forma subjetiva no proporciona buenos resultados y es, quizá, preferible aplicar el criterio SLS; c) con errores estacionarios es siempre aconsejable la utilización de criterios que tengan en cuenta la posible autocorrelación temporal de la secuencia de residuos.

CASO DE ESTUDIO REAL

La cuenca objeto de estudio es la misma que en el caso sintético. Junto con las series de precipitación y evapo-

transpiración potencial (gráfico 4) se ha introducido, para calibrar el modelo, la secuencia correspondiente de caudales mensuales observados (gráfico 5). Para reducir el efecto de los errores en las condiciones iniciales, estado de humedad y descarga subterránea, se realizaron unos tanteos previos, adoptándose finalmente los valores de 75 mm y 0 m³/seg, respectivamente.

El criterio objetivo utilizado para la calibración ha sido el SLS y se ha partido de los mismos conjuntos iniciales de los parámetros que en el estudio sintético. Tras aplicar el algoritmo de optimización de Rosenbrock se observó que aproximadamente en el 50 % de los casos se obtenían valores de los parámetros muy similares, lo que se puede calificar como muy satisfactorio (tabla 2). La inspección visual de los gráficos de ajuste y los valores de las funciones objetivo alcanzados en cada uno de los casos sirvieron para elegir el conjunto de parámetros óptimo: $H'_{\max} = 243$, $c = 0,2$, $I_{\max} = 72$ y $\alpha = 0,043$.

Con este conjunto de parámetros se simularon los caudales y se obtuvieron los residuos del modelo, para seguidamente proceder al análisis de sus propiedades. Según se desprende del gráfico 6, donde se dibujan los residuos frente a los caudales simulados, puede admitirse que aquéllos constituyen una serie estacionaria, dado que para todo el rango de caudales la media y varianza (dispersión de los puntos respecto a la media) es prácti-

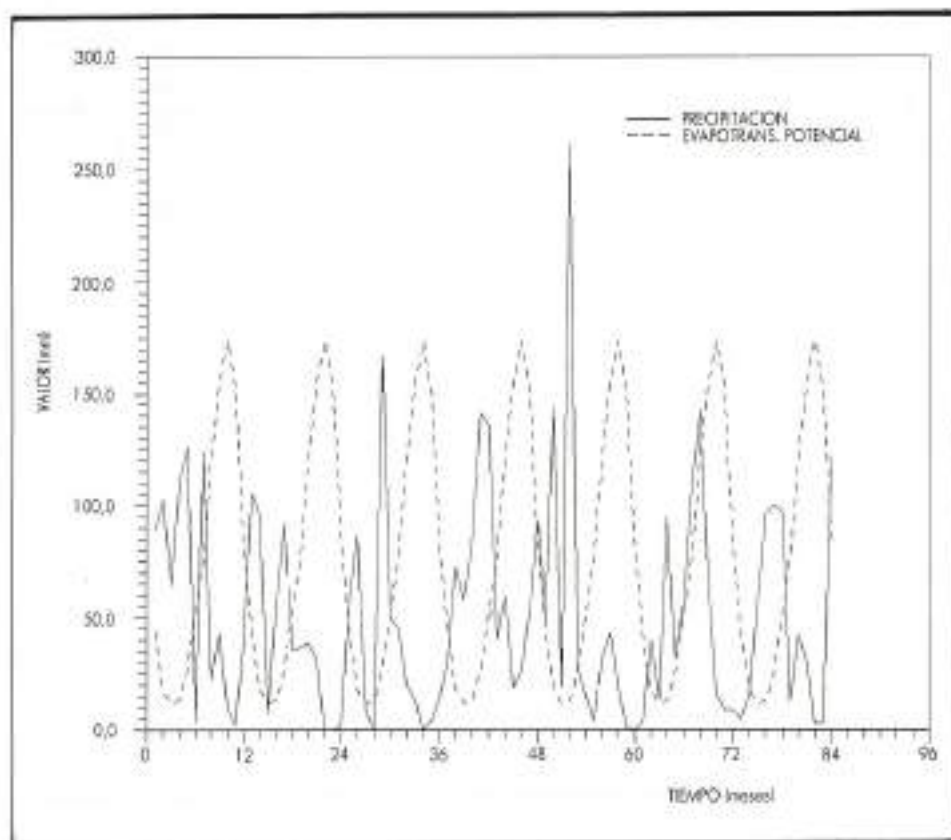


GRAFICO 4. Series históricas de precipitación y evapotranspiración potencial.

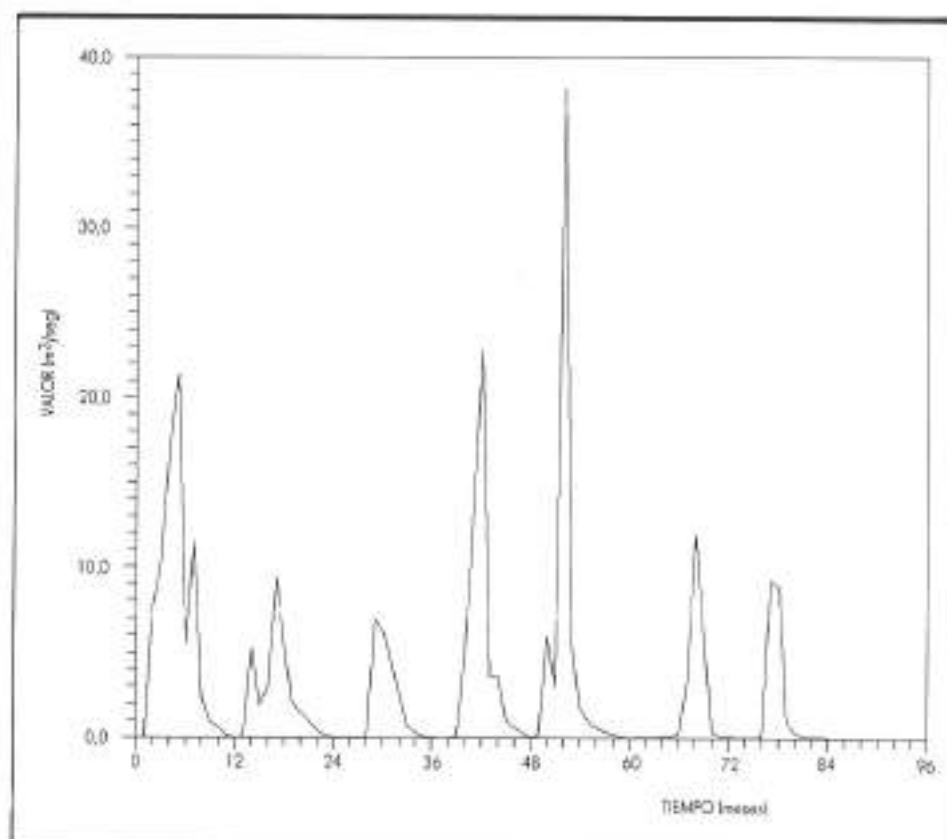


GRAFICO 5. Serie histórica de caudales.

CONJUNTO PARAMETROS	INICIAL				FUNCION OBJETIVO	FINAL				FUNCION OBJETIVO
	$H_{m\acute{o}d}$	C	$I_{m\acute{o}d}$	α		$H_{m\acute{o}d}$	C	$I_{m\acute{o}d}$	α	
1	200	0,20	100	0,004	5,973	242	0,20	75	0,043	2,073
2	100	0,50	200	0,01	9,624	241	0,18	81	0,044	2,096
3	200	0,60	50	0,005	4,803	243	0,20	72	0,043	2,073
4	200	0,40	300	0,004	13,588	243	0,20	80	0,046	2,077
5	300	0,60	150	0,10	11,377	242	0,20	70	0,041	2,074
6	150	0,20	200	0,03	6,863	232	0,26	83	0,045	2,180
7	50	0,30	150	0,02	28,309	240	0,20	60	0,037	2,106
8	150	0,30	200	0,10	11,107	240	0,30	227	0,101	2,523
9	50	0,20	500	0,50	65,989	237	0,35	1731	0,161	2,938
10	100	0,80	1000	0,004	22,167	161	1,29	152	0,012	5,416
11	200	0,80	150	0,04	5,714	241	0,21	107	0,056	2,141
12	150	0,40	300	0,05	7,710	244	0,24	94	0,063	2,176

TABLA II.

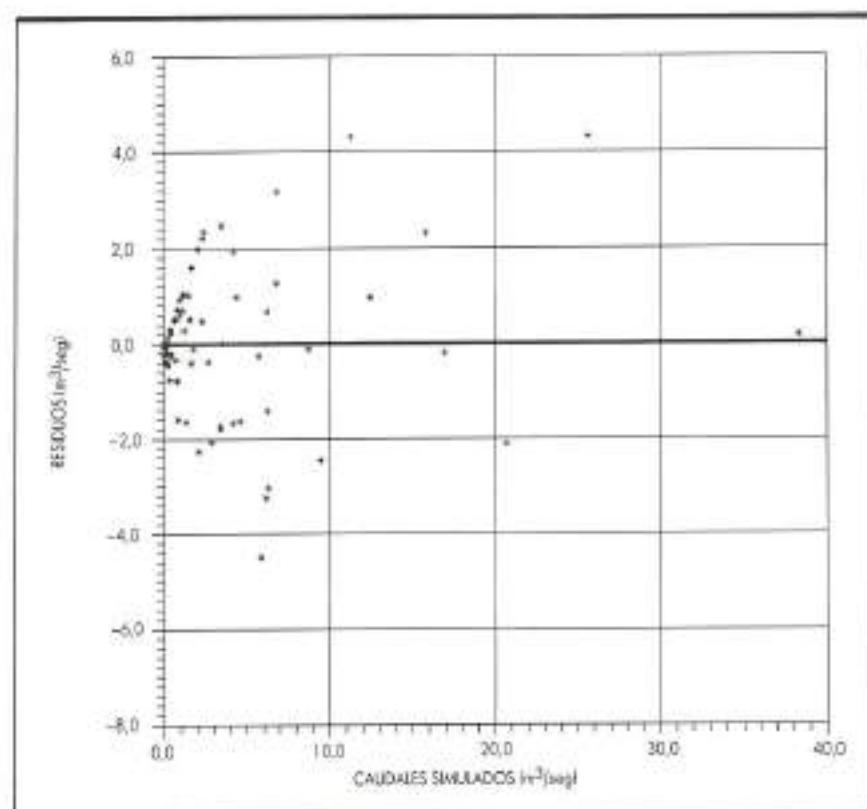


GRAFICO 6. Residuos versus caudales simulados.

camente constante. El dibujo de la secuencia temporal de residuos (gráfico 7) y de su correlograma (gráfico 8) indica que no existe autocorrelación temporal. Por otra parte, en el gráfico 9 se observa que los residuos siguen una distribución moderadamente normal sin sesgos importantes. El coeficiente de sesgo muestral es de 0,094. En el gráfico 10 se muestra el ajuste entre valores simulados e históricos.

En líneas generales se puede afirmar que el ajuste realizado es satisfactorio, y, dadas las propiedades de los residuos, no es necesario aplicar otros criterios objetivo como el WLS o el GLS.

CONCLUSIONES Y LINEAS FUTURAS DE INVESTIGACION

En este artículo se ha hecho una breve revisión del tipo de errores con los que se suele enfrentar la modelación lluvia-caudal, las funciones objetivo utilizadas para su tratamiento y las técnicas de optimización que sirven para minimizar esas funciones. Para ilustrar la problemática de la optimización se han realizado una serie de tests con datos generados sintéticamente y contaminados, posteriormente, con diferentes tipos de error. Finalmente se han calibrado los parámetros del modelo en un caso real. Las conclusiones alcanzadas y las líneas de investigación propuestas pueden agruparse en los siguientes puntos:

1. Problemática en la optimización: Interdependencia de parámetros y mínimos locales. Independientemente de la función objetivo e incluso para series de datos libres de error, se observa que en el proceso de optimización siempre aparecen mínimos locales en la función objetivo. En estas situaciones debe repetirse el proceso de optimización con diferentes conjuntos iniciales de los parámetros. Un análisis de los resultados, no sólo teniendo como finalidad la búsqueda del valor mínimo de las funciones objetivo resultantes, sino también conseguir una estructura adecuada de los residuos, permitirá acercarse al conjunto óptimo de los parámetros.

Entre las diversas causas que producen la aparición de mínimos locales, una de las que tiene mayor efecto es la propia estructura conceptual del modelo, dada la dificultad para evitar la interdependencia o correlación entre parámetros para cualquier tipo de situación climática o características físicas de la cuenca. Para mejorar esta representación conceptual una línea de investigación sería aquella en la cual se realizase una estimación de las matrices de covarianza y correlación de los parámetros.

2. Influencia de la función objetivo en la calibración. Las funciones objetivo utilizadas para series estacionarias de residuos, sin autocorrelación (mínimos cuadrados) o con autocorrelación (mínimos cuadrados generalizados), han proporcionado buenos resultados, incluso con la introducción de errores importantes en los datos. No suce-

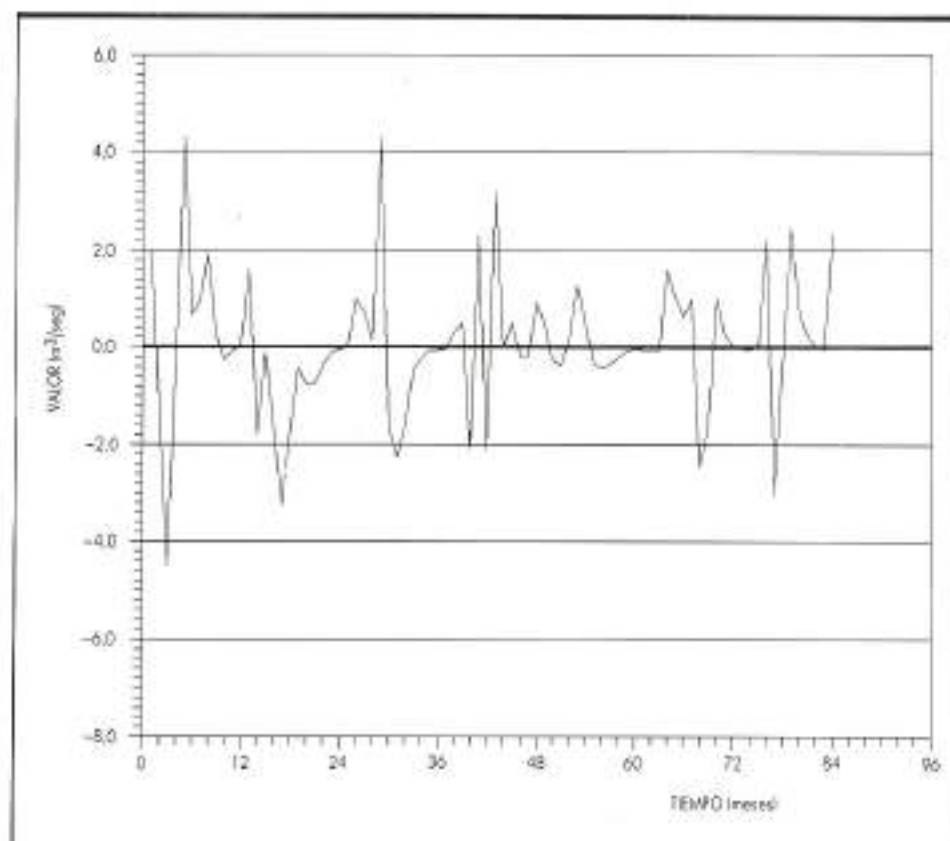


GRAFICO 7. Serie de residuos (caudales simulados-caudales observados).

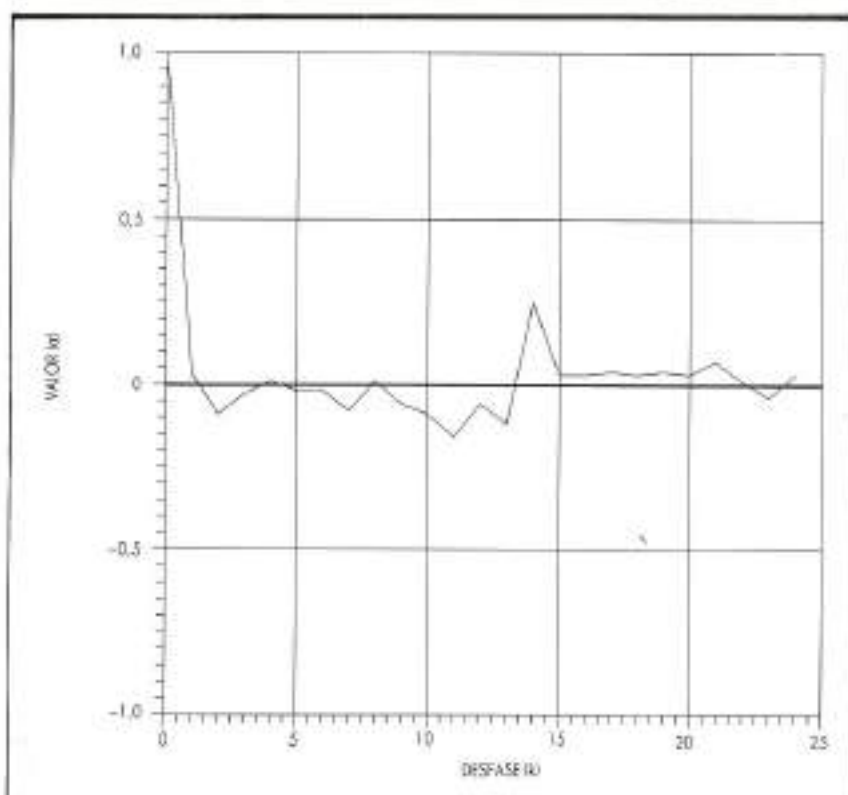


GRAFICO 8. Correlograma serie residuos.

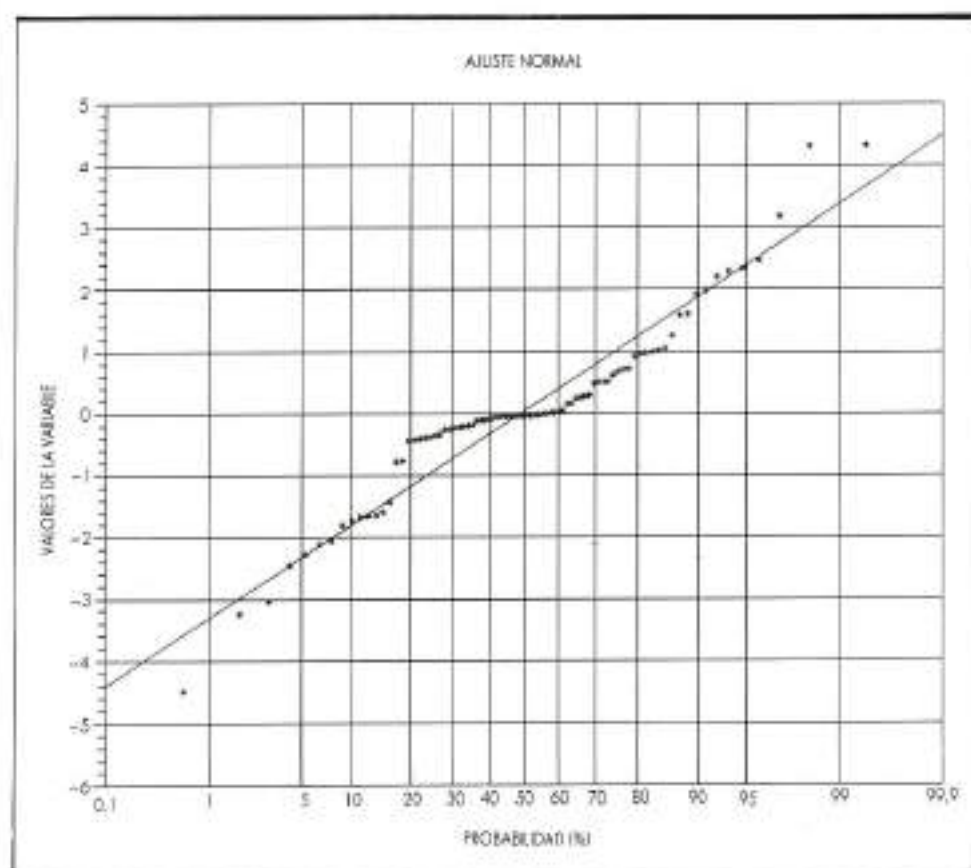


GRAFICO 9. Función de distribución residuos.

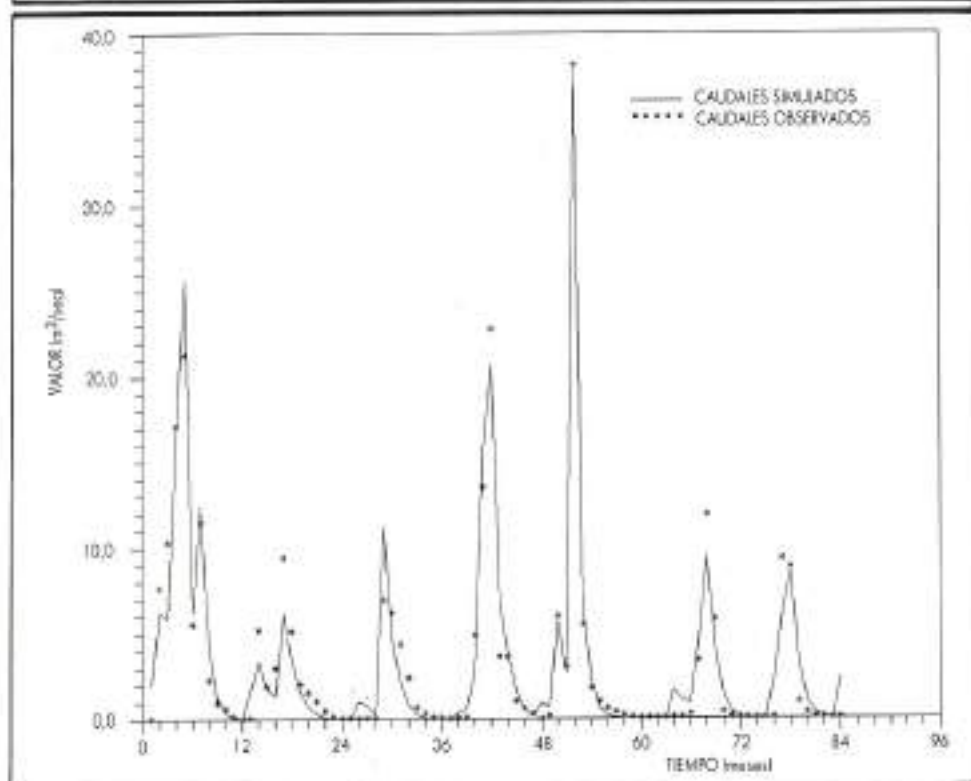


GRAFICO 10. Periodo completo de calibración.

de así cuando los errores no son estacionarios en varianza. Un criterio de mínimos cuadrados ponderados tradicionalmente utilizado, como es el error cuadrático relativo, no parece ser la forma más idónea para tratar este tipo de errores. Debería investigarse el funcionamiento de otros criterios similares con una asignación de pesos más realista.

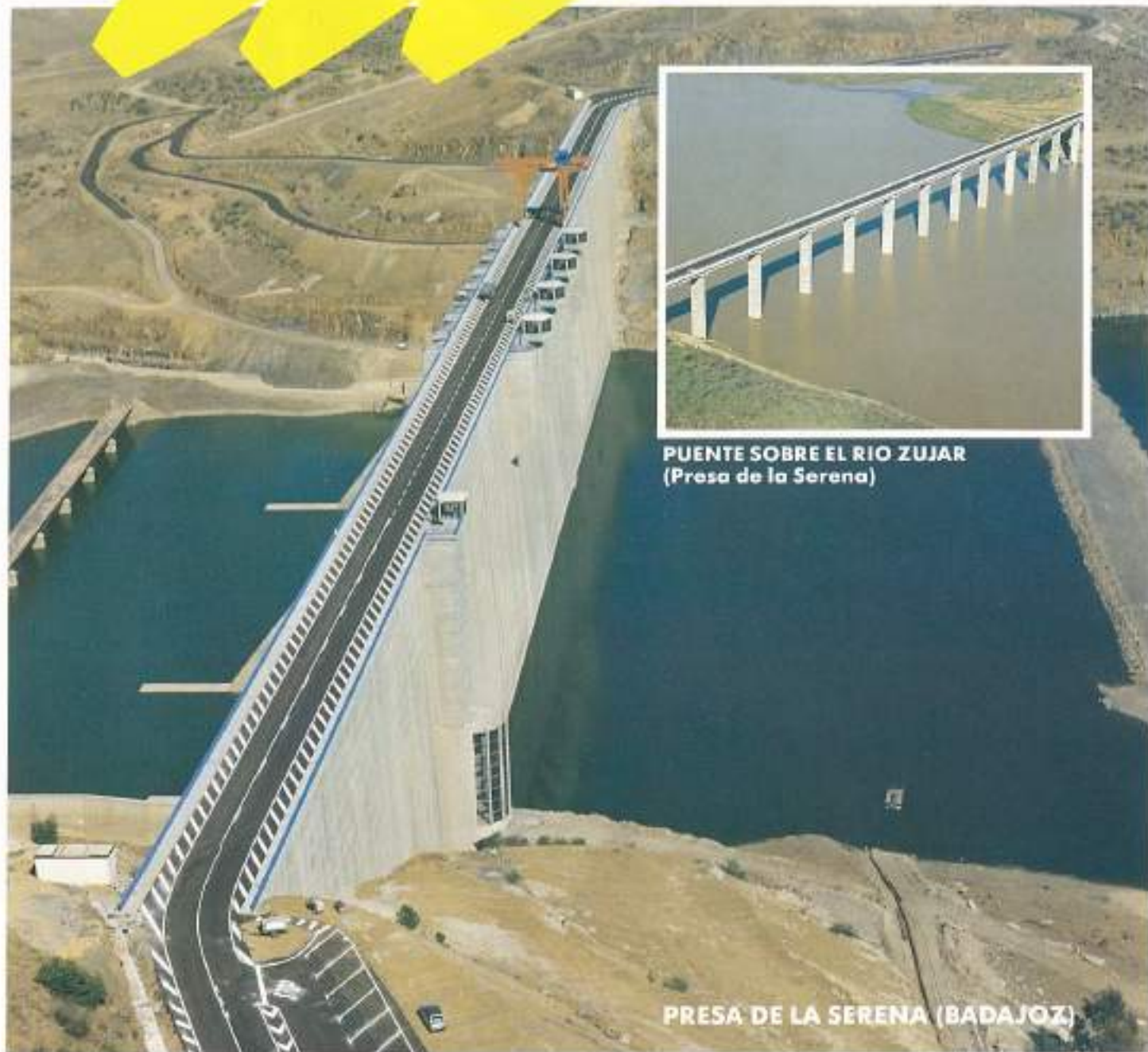
3. Regiones factibles de los parámetros. Uno de los inconvenientes que se argumenta para la utilización de métodos de calibración automática de parámetros es la posible falta de sentido físico de los valores estimados. En los tests sintéticos realizados, y en la aplicación a un caso real, el porcentaje de veces en que las estimas de los parámetros se apartaron de lo que podría considerarse como parámetros factibles fue despreciable. Sin embargo, la experiencia dice que en otras ocasiones esto sí sucede. Ello es debido, generalmente, o a una formulación conceptual no adecuada del modelo, o lo que es más corriente, que los datos utilizados en la calibración no guardan coherencia entre sí.

REFERENCIAS BIBLIOGRAFICAS

- BOX, G. E. P., y JENKINS, G. M. (1970). «Time Series Analysis: Forecasting and Controls». Holden-Day, San Francisco, California.
- CLARKE, R. T. (1973). «A review of some mathematical models used in Hydrology, with observations on their calibration and use». *Journal of Hydrology* 19, págs. 1-20.
- DISKIN, M. H., y SIMON, E. (1977). «A procedure for the selection of objective functions for hydrologic simulation models». *Journal of Hydrology*, 34, págs. 129-149.
- ESTRELA, T. (1980). «Los modelos de simulación integral de Cuenca y su utilización en estudios de recursos hídricos». *Ingeniería Civil*, 72. CEDEX. Madrid.
- GUPTA, V. K., y SOROOSHIAN, S. (1983). «Uniqueness and observability of Conceptual Rainfall-Runoff Model Parameters: The Percolation Process Examined». *Water Resources Research*, vol. 19, núm. 1, págs. 269-276. Febrero.
- IBITT, P., y O'DONNELL, T. (1971). «Fitting methods for conceptual catchment models». *Journal of the Hydraulics Division*. Proceeding of the American Society of Civil Engineers. Septiembre.
- IBITT, R. P. (1972). «Effects of Random Data Errors on the

- Parameter Values for a Conceptual Models». *Water Resources Research*, vol. 8, núm. 1, Febrero.
- JOHNSTON, P. R., y PILGRIM, D. H. (1976). «Parameter optimization for Watershed Models». *Water Resources Research*, vol. 12, núm. 3. Junio.
- KUZCERA, G. (1983). «Improved Parameter Inference in Catchment Models. 1. Evaluating Parameter Uncertainty». *Water Resources Research*, vol. 19, núm. 5, págs. 1151-1162. Octubre.
- MEIN, R. G., y BROWN, B. M. (1978). «Sensitivity of Optimized Parameters in Watershed Models». *Water Resources Research*, vol. 14, núm. 2. Abril.
- PICKUP, G. (1977). «Testing efficiencies of algorithms and strategies for automatic calibration of rainfall-runoff models». *Hydrological Science Bulletin*. 22(2), 257-274.
- ROSENBROCK, H. H. (1960). «An Automatic Method for Finding the Greatest or Least Value of a Functions». *The Computer Journal*, vol. 3, 1960, págs. 175-184.
- SOROOSHIAN, S., y DRACUP, J. A. (1980). «Stochastic Parameter Estimation Procedures for Hydrologic Rainfall-Runoff Models: Correlated and Heteroscedastic Error Cases». *Water Resources Research*, vol. 16, núm. 2, págs. 430-442. Abril.
- SOROOSHIAN, S., y ARFI, F. (1982). «Response Surface Parameter Sensitivity Analysis Methods for Postcalibration Studies». *Water Resources Research*, vol. 18, núm. 5, págs. 1531-1538. Octubre.
- SOROOSHIAN, S., GUPTA, V. K., y FULTON, J. L. (1983). «Evaluation of Maximum Likelihood Parameter Estimation Techniques for Conceptual Rainfall-Runoff Models: Influence of Calibration Data Variability and Length on Model Credibility». *Water Resources Research*, vol. 19, núm. 1, págs. 251-259. Febrero.
- TEMEZ, J. R. (1977). «Modelo matemático de transformación precipitación-aportación». ASINEL. España. Octubre.
- TROUTMAN, B. M. (1985). «Errors and Parameter Estimations in Precipitation-Runoff Modeling. 1. Theory and 2. Case Study». *Water Resources Research*, vol. 21, núm. 8, págs. 1196-1213 y 1214-1222. Agosto.
- U. S. ARMY CORPS OF ENGINEERS, Hydrologic Engineering Center (1973). «HEC-1 Flood Hydrograph Package, User Manual». Davis, California, 1973.
- WILLIAMS, B. J., y YEH, W. W.-G. (1983). «Parameter estimation in rainfall-runoff models». *Journal of Hydrology*. 63 (1983), 273-293.

COMPROMISO DE CALIDAD



PUENTE SOBRE EL RIO ZUJAR
(Presa de la Serena)

PRESA DE LA SERENA (BADAJOZ)

TODO EN SEÑALIZACION Y SEGURIDAD VIAL

• Equipos reflectantes
para Policía Municipal.



• Señales de Tráfico pintadas
y reflexivas.



• Nueva señalización urbana.



• Barreras de Seguridad.

• Señalización para obras.



señalizaciones **VILLAR**

Avda. Valladolid, 58 apdo. 61 42004 **SORIA**

Telf. (975) 22 05 04

Fax (975) 22 06 45