

## TRATAMIENTO NO DETERMINISTA DEL PROCESO DE CALIBRACIÓN DE LOS MODELOS DE CALIDAD DEL AGUA EN REDES DE ABASTECIMIENTO. EL ANÁLISIS DE INCERTIDUMBRE.

P. Amparo López Jiménez<sup>1</sup>, Joaquín Izquierdo Sebastián<sup>1</sup>, F. Javier Martínez Solano<sup>1</sup>,  
José Luis Díaz Arévalo<sup>1</sup>.

**Resumen** – Los resultados de los modelos matemáticos de calidad en redes de abastecimiento, por regla general, son de carácter determinista. Esto implica que tanto para caudales y presiones, como para valores de concentración, proponen soluciones únicas para cada línea e instante de tiempo. Sin embargo, la realidad es un proceso aleatorio, normalmente relacionado con las mediciones, que poseen un error inherente. En este artículo se propone una discusión sobre cómo los errores de las mediciones y de la estimación de los parámetros de la calibración de los modelos de calidad de redes, se transmiten a los resultados del mismo. Para ello se propone un análisis de incertidumbre basado en la determinación de la dispersión de los resultados en forma de varianza estadística. Este análisis combinado con el de sensibilidad, se convierten en una herramienta que juega un papel importante en la calibración del modelo.

**Abstract** – The results proposed by supply networks quality models are generally deterministic. This means that we will obtain either for pressure and flows, and quality parameters unique values for each line and time instant. Nevertheless, reality is not deterministic, normally related to measurements. In this article we discuss about how the input measurement errors are transmitted to output results in this sort of quality models. We propose an uncertainty analysis based on determination of dispersion of results by means of statistical variance of predicted values. This analysis combined with a sensitivity analysis, become a very important tool of decision in validation and calibration of the bulk model.

**Palabras clave:** Análisis de sensibilidad, incertidumbre, intervalo de confianza, modelo de calidad de agua en redes de abastecimiento, calibración.

**Keywords:** Sensitivity and uncertainty analysis, confidence region, supply networks quality models, validation.

---

<sup>1</sup> Universidad Politécnica de Valencia – Departamento de Ingeniería Hidráulica y Medio Ambiente – Grupo Multidisciplinar de Modelación de Fluidos – Camino de Vera S/N – CP E46022 – Valencia (España) – Tel: (+34) 96 3879890 – Fax: (+34) 96 3877981 E-mail: palopez@gmmf.upv.es, jizquier@gmmf.upv.es, jmsolano@gmmf.upv.es, jldiaz@gmmf.upv.es.

## INTRODUCCIÓN.

La mayoría de los modelos matemáticos de las redes de abastecimiento, ya se trate de aspectos hidrodinámicos o de aspectos de calidad de aguas, son de carácter determinista, en el sentido de que proponen soluciones exactas para las magnitudes objeto del estudio en todo instante y posición de la red que se trata de modelar.

De esta forma, la predicción que realizan los modelos matemáticos computacionales del comportamiento de una red de agua, presenta las presiones y caudales que van a estar existiendo en cada uno de los nudos y líneas de una red, para todo instante de tiempo; y los niveles de calidad (concentración de cloro, tiempo de permanencia del agua, estado de mezcla, por ejemplo) en la misma. Esta presentación de resultados siempre tiene un valor determinado, concreto para cada instante de tiempo y posición.

Sin embargo, para poder especificar las condiciones iniciales es necesario medirlas con algún aparato. Debemos conocer las presiones en los puntos de presión conocida, los caudales de consumo, o las concentraciones de cloro en puntos concretos. Por otra parte, en el modelo no sólo intervienen estas magnitudes que son resultados de una medición, sino que además como modeladores, debemos asignar valores a ciertos parámetros que intervienen en las ecuaciones matemáticas que representan el comportamiento del sistema; factores de fricción, coeficientes de pérdidas menores, coeficientes representativos del comportamiento de las bombas, coeficientes de reacción para el cloro, de difusividad, etc.... que deben ser estimados en intervalos conocidos, pero no con valores completamente prefijados a priori. Este tipo de análisis ha sido ya considerado en modelos de calidad de aguas para cauces receptores superficiales (Reichert 2001, Qual2E manual, Chapra 1997), sin embargo, la aplicación de este análisis es relativamente reciente para modelos de predicción de calidad en redes a presión (Pasha, 2005)

Como resultado de esta metodología por todos seguida para la calibración se obtienen unos valores concretos. Pero éstos números contienen incertidumbres, ya que en el proceso de medición en que se obtienen o estiman hay factores que, en general, no se pueden controlar. Por ejemplo, si se mide el peso de un cuerpo con una balanza, pueden ocurrir errores en la lectura de la aguja, debido a alguna vibración que produzca el paso de un vehículo, o a causa de alguna corriente de aire, las presiones dependen de la calidad de los manómetros, y lo mismo pasa con las concentraciones medidas, etc. Es decir, siempre que se mide algo hay errores. Por tanto, como resultado de una medición se debe dar el número obtenido así como los límites de los errores que se puedan cometer.

Por supuesto que mientras menores sean los errores que se cometan, mejor será el resultado de la medición. Lo más que se puede hacer es lograr que el intervalo dentro del cual se encuentran las mediciones disminuya, pero no se puede eliminar.

En general, podemos afirmar que no se puede hacer una medición con precisión absoluta. Siempre se tendrá un intervalo de error dentro del cual cae el "verdadero" valor. Los límites de error experimental son estimaciones cuantitativas de la importancia de las perturbaciones que muchos factores externos, prácticamente imposibles de controlar, provocan en la medición. La determinación del intervalo de error experimental es parte del trabajo cotidiano de un científico experimental. Mientras menor sea el intervalo de error, más precisa será la medición efectuada. Lo mismo ocurre para los valores de los parámetros que intervienen en las ecuaciones y que son objeto de estimación en el proceso de calibración:

siempre se parte de unos valores típicos, sin embargo, en cada caso particular habrá un valor verdadero que difícilmente llegaremos a conocer nunca.

Como consecuencia de esta reflexión, se infiere que en el modelo matemático de la red, también hay un error inherente en la modelación, consecuencia de que intervienen en la misma valores de parámetros que han sido estimados como consecuencia de una medición o de una calibración del modelo. Obviamente, si estos valores son más precisos, también nuestros resultados de la modelación lo serán, pero en buena lid deberíamos reconocer que si imprecisos son estos valores que intervienen en la consideración del modelo, también imprecisos serán los resultados del mismo. Se propone en este documento un tratamiento no determinista de los resultados del modelo, mediante el análisis de incertidumbre sobre el mismo.

## **EL ANÁLISIS DE INCERTIDUMBRE: EL ESTABLECIMIENTO DE INTERVALOS DE CONFIANZA.**

### **El análisis de incertidumbre.**

La incertidumbre de parámetro se debe al conocimiento incompleto del valor verdadero del mismo y se origina en la necesidad de establecer inferencias para toda una población a partir de muestras pequeñas sobre los procesos que tienen un cierto grado de aleatoriedad. Por ejemplo: la imprecisión o error en las muestras analíticas, las inferencias hechas con una base de datos limitada u otros errores sistemáticos en la colección y agregación de datos (Evans 2002). Cuando estos parámetros intervienen en la estimación mediante modelos, el error se transporta a las soluciones de salida y por tanto también las mismas deben ser considerada en forma de intervalos de predicción, y no de un valor único.

Por otro lado, la incertidumbre de modelo se debe a la necesidad de depender de una analogía, que es necesariamente falsa, en la modelación de un cierto proceso (EPA 1987). La incertidumbre de modelo puede originarse en las suposiciones y simplificaciones hechas de los procesos reales para propósitos de manejabilidad, y también del mal uso o aplicación de un modelo o la utilización de datos sustitutos. Además, la incertidumbre de modelo se puede introducir cuando un modelo se basa o está validado para una porción de espacio de un parámetro y se utiliza para rangos fuera de los cuales el modelo es válido. Un ejemplo de esto es la extrapolación de dosis mayores a dosis menores de los modelos de dosis-respuesta. Otros factores que pueden introducir incertidumbre incluyen la resolución del modelo, la validación y las condiciones de contorno e iniciales que intervienen en la modelación.

El tratamiento de la incertidumbre es muy usual en los estudios de modelación matemática. En el contexto del tema que nos ocupa, los modelos de calidad de agua, el análisis de incertidumbre tiene por objetivo el establecimiento de intervalos de confianza para el parámetro final modelado alrededor de los valores deterministas del mismo. En este caso: la concentración de cloro a lo largo de las líneas y del tiempo de la simulación, aunque hay otros indicadores de la calidad que deberían ser tratados igualmente.

En muchas ocasiones puede ser especialmente interesante realizar la estimación de cómo se propagan estos errores incluso antes de realizar la campaña de medición o en la primera fase de calentamiento o inicialización del proceso de calibración, con los primeros datos del cauce receptor con objeto de programar fases posteriores de la medición.

Es importante conocer si los errores en una determinada condición de contorno van a ser anulados o amplificados. Esto nos ayudaría a seleccionar métodos de mediciones, distribución espacial y temporal de campañas futuras y sobre todo aplicación de recursos para mejorar la precisión de las variables de entrada.

La estimación del efecto relativo previamente presentado informa de la influencia de una variable en la solución, así como de la linealidad de este efecto. Como complemento el análisis de incertidumbre informa sobre la influencia de la dispersión de un parámetro de entrada sobre la variación de los resultados: en definitiva sobre la precisión de la modelación.

Independientemente del cuidado con el que se desarrollan los experimentos, el procedimiento humano y el propio aparato, tienen un error inevitable. Asimismo, las imperfecciones se deben además a la aleatoriedad de las medidas y a que la propia suposición de estados permanentes no existe en la naturaleza.

El error experimental (llamado algunas veces ruido) es debido a la fluctuación o discrepancias en la repetición de los experimentos. Este error no implica que los datos no sean buenos, o no sean reales, se refiere a que las medidas tienen fluctuaciones debidas a imperfecciones en el proceso de medición: errores de la instrumentación, variación de las condiciones ambientales (especialmente en los procesos considerados permanentes), habilidades del personal de laboratorio y otros. A pesar de ello, estas variaciones deben ser conocidas, valoradas y a su vez, minimizadas.

## **La incertidumbre de los resultados calculados mediante el modelo**

En el tema que nos ocupa, la magnitud que vamos a querer calibrar es la concentración de cloro a lo largo de las líneas del modelo de la red de abastecimiento en una modelación en varios intervalos de tiempo. Esta va a ser considerada como una magnitud normal en un análisis estadístico no determinista. Si se conoce la varianza de los valores calculados por el modelo, o lo que es lo mismo, la desviación típica de los mismos, podremos conocer la región de confianza de la predicción que realiza el modelo y esto sería de gran utilidad para la calibración y determinación de la fiabilidad del mismo. Se trata, por tanto, de conocer cómo se transmite la incertidumbre de los datos de entrada a los valores calculados a través del modelo.

El caso más simple es aquel en que los parámetros externos de entrada se ligan con los de salida mediante funciones lineales. Supóngase que  $y$  es la concentración que buscamos, (cloro), y  $x_i$  las concentraciones medidas en los nudos de aporte ( $x_i$  son concentraciones de diferentes sustancias de las cuales depende la concentración calculada), que se comportan como variables normales. Si la influencia fuera lineal, con tres concentraciones medidas el modelo adoptaría la forma:

$$y = \phi_0 + \phi_1 x_1 + \phi_2 x_2 + \phi_3 x_3 \quad (1)$$

donde  $\phi_i$  serían en este caso parámetros constantes para el análisis que estamos considerando. Los errores en  $x_i$  serían aleatorios y ello convertiría a estas magnitudes en distribuciones normales. El error esperado en la función de salida sería el siguiente:

$$E(y) = \phi_0 + \phi_1 E(x_1) + \phi_2 E(x_2) + \phi_3 E(x_3) \quad (2)$$

Y la varianza esperada para la función y sería:

$$\begin{aligned} Var(y) = & \phi_1^2 Var(x_1) + \phi_2^2 Var(x_2) + \phi_3^2 Var(x_3) + \phi_1 \phi_2 Cov(x_1, x_2) + \\ & + \phi_1 \phi_3 Cov(x_1, x_3) + \phi_2 \phi_3 Cov(x_2, x_3) \end{aligned} \quad (3)$$

Expresión en la que las covarianzas entre las magnitudes de entrada pueden ser positivas, negativas o nulas. Si dichas magnitudes  $x_i$  no estuvieran correlacionadas, esto es, que la varianza de una no afectará a la varianza de la otra, los términos de covarianza serían nulos, por lo que la expresión (4.8) resultaría:

$$Var(y) = \phi_1^2 Var(x_1) + \phi_2^2 Var(x_2) + \phi_3^2 Var(x_3) \quad (4)$$

Los términos de la derecha representan las contribuciones separadas de cada variable a la varianza total de la magnitud calculada. Dicho valor representa la varianza de la función calculada y puede ser utilizada para definir el intervalo de confianza de la concentración proporcionada por el modelo en este caso.

El hecho de que las variables que se han considerado de entrada no estén correlacionadas adopta pleno sentido físico en el caso del modelo de dispersión de contaminantes y cuando las magnitudes consideradas de entrada son, como poníamos por caso, las concentraciones en los puntos de aporte. En este caso, la varianza de una de ellas no afectará a la varianza de la concentración de los demás, a no ser que el sistema no estuviera bien ramificado.

El mismo tipo de razonamiento podría hacerse si la función no fuera lineal. Supongamos que la función  $y=f(x_1, x_2)$  se reemplaza por su aproximación lineal alrededor de los valores medios:

$$Y = \phi_0 + \phi_1(x_1 - \bar{x}_1) + \phi_2(x_2 - \bar{x}_2) \quad (5)$$

Donde Y es el desarrollo de Taylor de la función y en el que los términos de mayor orden han sido ignorados. En él, la aproximación se encuentra centrada alrededor de los valores medios de las variables de entrada. Los coeficientes lineales y las derivadas de y con respecto a  $x_1, x_2$  evaluadas en las cercanías de los valores medios  $\bar{x}_1$  y  $\bar{x}_2$  se relacionan por medio de las expresiones:

$$\phi_1 = \left( \frac{\partial y}{\partial x_1} \right)_{\bar{x}_1} ; \quad \phi_2 = \left( \frac{\partial y}{\partial x_2} \right)_{\bar{x}_2} \quad (6)$$

Así, la varianza de la función objetivo será, tal como se trató para la función lineal:

$$Var(Y) = \left( \frac{\partial y}{\partial x_1} \right)_{\bar{x}_1}^2 Var(x_1) + \left( \frac{\partial y}{\partial x_2} \right)_{\bar{x}_2}^2 Var(x_2) + \left( \frac{\partial^2 y}{\partial x_1 \partial x_2} \right)_{\bar{x}_1 \bar{x}_2} Cov(x_1, x_2) \quad (7)$$

Esto puede ser generalizado para el caso de  $n$  variables medidas:

$$Var(Y) = \sum_{i=1}^n \left( \frac{\partial y}{\partial x_i} \right)_{x_i}^2 Var(x_i) + \sum_{i=1}^n \sum_{j=i+1}^n \left( \frac{\partial^2 y}{\partial x_i \partial x_j} \right)_{x_i x_j} Cov(x_i, x_j) \quad (8)$$

Habr , por lo tanto,  $n(n-1)/2$   tems de covarianza, cada uno positivo, negativo o nulo. Ello puede hacer crecer o decrecer la varianza de la funci n objetivo. Si los valores medidos son independientes entre s , los t rminos de covarianza ser n nulos. Se suele tomar siempre una interdependencia nula pero, en el caso general, esto habr a que justificarlo.

Si la funci n proviene de un modelo de simulaci n, a partir de un sistema de ecuaciones resueltas en diferencias finitas, o no es posible la evaluaci n de la derivada de la funci n en el entorno del punto medio, esta derivada puede ser sustituida por evaluaciones incrementales. Esta linealizaci n ser  v lida para variaciones “peque as” en el intervalo de estudio, por ejemplo, dos o tres desviaciones t picas de las funciones de entrada (Mac Berthouex y Brown, 1994).

Esta metodolog a es la que se implementa para hacer un an lisis de sensibilidad llamado “de primer orden” (Chapra, 1997). El an lisis de error de primer orden utiliza la aproximaci n de primer orden para la relaci n del c lculo de varianzas en situaciones multivariantes (EPA, 1987) (Canal de Isabel II, 1992). Se supone que las varianzas de entrada act an independientemente (se ignoran las covarianzas) y el modelo es lineal (se omiten t rminos de orden superior en el desarrollo de Taylor). Estas aproximaciones son generalmente buenas para el caso del modelo de dispersi n que nos ocupa.

En la contribuci n anterior, ha sido definido el concepto de efecto relativo. Este es un valor normalizado. Supongamos que con el establecimiento de este efecto relativo hemos conocido que un incremento  $\Delta k$  en valor del par metro de entrada genera un incremento  $\Delta C$  en la concentraci n de salida.

Para definir la influencia que la varianza de los valores de entrada tiene en los de salida, se considera que los componentes de varianza de cada variable de salida  $C_j$  son los porcentajes de la varianza de salida atribuibles a cada variable de entrada, par metro  $k_i$ , calculadas de la siguiente manera:

$$Var(C_j) = \sum_i Var(k_i) \left( \frac{\Delta C_j}{\Delta k_i} \right)^2 \quad (9)$$

donde:  $Var(C_j)$ : Varianza de las variables de salida  $C_j$

$Var(k_i)$ : Varianza de los par metros de entrada  $k_i$

De esta manera cada t rmino es un sumando de la varianza de la variable de salida  $C_j$  debido a la variable de entrada  $k_i$ .

Los componentes de la varianza de salida representan una ponderaci n de las varianzas de entrada por el cuadrado de la sensibilidad del modelo a la entrada. As , una variable de entrada en particular puede contribuir mucho o poco a la varianza de la salida en funci n del efecto que tenga sobre la misma, y su propia varianza.

Pueden perturbarse todas las variables de entrada o solamente algunas dependiendo de los parámetros que resulten de interés.

Las necesidades para la realización de este análisis de error de primer orden son la magnitud de la perturbación de la variable de entrada, la perturbación que esta genera en la de salida, (habrá sido establecida en el efecto relativo que hemos determinado anteriormente), y el valor de la varianza de la magnitud de entrada.

Conocida, pues, la varianza de la concentración de salida, podremos establecer en cada caso, cual es la región de confianza de la predicción del modelo.

El conocimiento del efecto que la variación de las magnitudes de entrada tiene en las variables de salida, solo tiene sentido en los parámetros externos, puesto que son los que están sujetos a distribuciones estadísticas, y nos informa de las magnitudes en las que hay que encontrar mayor precisión. Si la varianza de un dato de entrada afecta mucho a la varianza de los valores de salida, aquel parámetro deberá ser evaluado con mayor precisión.

En el sentido estadístico, la precisión depende de lo que se llama la región de confianza (Mac Berthouex y Brown, 1994), esto es, puede ser definida por el usuario, pero que suele ser la región en la que se encuentran el 95% de los valores. El tamaño de esta región de confianza se desea que sea lo menor posible, y ello dependerá de la precisión del aparato de medida, del número de observaciones realizado y de la localización de las mediciones respecto a la variable independiente. No siempre un mayor número de mediciones mejora la varianza, puesto que aunque las poblaciones sean muy grandes siempre hay una muestra que las representa con cierta validez.

### **Regiones o intervalos de confianza**

Las regiones (o intervalos) de confianza de una variable estadística es aquel rango de valores entre los cuales existe una cierta probabilidad de que se encuentren las mediciones para una población, en este caso, de datos experimentales (Tze Ling Ng, 2005). En este sentido puede ser considerado, como se ha anunciado anteriormente, el intervalo de confianza del 95%, (esto es, el 95% de los valores experimentales, considerada la magnitud que se mide como una variable normal, se encontrarán dentro del intervalo), una banda de dos desviaciones típicas alrededor de los valores medios (Mac Berthouex y Brown, 1994,).

Esta estimación para la confianza de los valores puede tener sentido tanto para las magnitudes medidas que son condiciones de contorno en los nudos de entrada del sistema (concentraciones y temperaturas) como para las que son referencia para la calibración a lo largo del mismo. Y entendiendo que las magnitudes que calcula el modelo a su vez son concentraciones reales, también para los resultados del modelo.

Establecer la región de confianza de los resultados del modelo es interesante desde el punto de vista de la calibración, porque cuantifica la fiabilidad de las predicciones y representa de manera local la influencia de la incertidumbre en las magnitudes registradas en los nudos de aporte. Dicha región de confianza se establecerá con el conocimiento de la varianza en los resultados.

Para el conocimiento de la varianza de las magnitudes calculadas utilizamos la expresión anteriormente presentada. Consideremos el efecto relativo como un parámetro adimensional de la sensibilidad de la forma siguiente: el incremento de la concentración

reducido al valor de la concentración ( $C$ ), cuando se incrementa una magnitud externa, también reducida al valor centrado de dicha magnitud ( $k$ ).

Así pues, la variación absoluta de la magnitud de la que pretendemos conocer su valor, se relaciona con la magnitud que perturbamos de la siguiente manera:

$$ER \frac{\Delta k}{k} = \frac{\Delta C}{C} \quad \Rightarrow \quad \frac{\Delta C}{\Delta k} = ER \frac{C}{k} \quad (10)$$

Y por lo tanto, la expresión para la varianza transmitida a la magnitud calculada:

$$Var(C_j) = \sum_i Var(k_i) \left( \frac{\Delta C_j}{\Delta k_i} \right)^2 = \sum_i Var(k_i) \left( ER_{ij} \frac{C_j}{k_i} \right)^2 \quad (11)$$

Esto es, conocido el efecto relativo de la variación de una magnitud externa sobre la concentración que buscamos, y conocida la varianza propia de dicha magnitud, se puede conocer la varianza de la concentración calculada mediante la expresión (11).

Esta expresión contiene implícita la suposición de que la influencia de una magnitud sobre otra se comporta de forma lineal, y ello no está asegurado. El efecto relativo si el comportamiento es no lineal puede ser variable dependiendo de la magnitud de la perturbación. Aunque también se ha despreciado las covarianzas suponiendo un modelo lineal. Por lo tanto, esta expresión adopta pleno significado en ER constantes, influencias lineales y perturbaciones no correlacionadas. Esas son las limitaciones de la aplicación.

Asimismo podrá hablarse de la desviación típica de la magnitud calculada:

$$\sigma(C_j) = \sqrt{Var(C_j)} = \sqrt{\sum_i Var(k_i) \left( \frac{\Delta C_j}{\Delta k_i} \right)^2} = \sqrt{\sum_i Var(k_i) \left( ER_{ij} \frac{C_j}{k_i} \right)^2} \quad (12)$$

Supongamos que se conoce la desviación típica relativa de las magnitudes que se miden. Esta se relaciona con la varianza de la siguiente manera:

$$\begin{aligned} \sigma_R(k) &= \frac{\sigma(k)}{\bar{k}} & \sigma(k) &= \bar{k} \sigma_R(k) & ; \\ Var(k) &= [\sigma_R(k) \bar{k}]^2 = [\sigma_R(k)]^2 \bar{k}^2 \end{aligned} \quad (13)$$

Esto es:

$$\begin{aligned} \sigma(C_j) &= \sqrt{Var(C_j)} = \sqrt{\sum_i Var(k_i) \left( \frac{\Delta C_j}{\Delta k_i} \right)^2} = \sqrt{\sum_i Var(k_i) \left( ER_{ij} \frac{C_j}{k_i} \right)^2} \\ &= \sqrt{\sum_i [\sigma_R(k_i)]^2 \bar{k}_i^2 \left( ER_{ij} \frac{C_j}{k_i} \right)^2} \end{aligned} \quad (14)$$

En este caso,  $\bar{k}_i$  es el valor medio de la magnitud perturbada,  $i$ , o sea:



$$\sigma(C_j) = \sqrt{\sum_i [\sigma_R(k_i)]^2 k_i^2 \left( ER_{ij} \frac{C_j}{k_i} \right)^2} = \sqrt{\sum_i [\sigma_R(k_i)]^2 ER_{ij}^2 C_j^2} \quad (15)$$

Así, deberá ser conocida la desviación típica relativa de la magnitud perturbada y el efecto relativo que esta perturbación provoca alrededor de los valores calculados. Y (conocido el valor medio), podemos conocer la desviación típica de la magnitud calculada por el modelo, y con ella la región de confianza. Hay que recalcar que la magnitud perturbada es cada una de las magnitudes de entrada medidas en cada uno de los aportes en que se ha medido, por lo que en sistemas con muchos aportes y muchas mediciones, este sumatorio es considerable.

Para conocer la amplitud de la región de confianza, o más bien la banda de confianza alrededor del valor calculado por el modelo, conocida la desviación de las magnitudes calculadas a partir de todas las desviaciones de las magnitudes medidas y el correspondiente efecto relativo, se determina la desviación típica relativa de la magnitud calculada ( $\sigma_R$ ), dividiendo la desviación típica calculada mediante la expresión (15) por el valor medio para dicha concentración:

$$\sigma_R(C_j) = \frac{\sigma(C_j)}{C_j} = \sqrt{\sum_i [\sigma_R(k_i)]^2 ER_{ij}^2}$$

Dicha amplitud, como se ha dicho, es de dos desviaciones típicas alrededor del valor calculado (considerado entonces como un valor medio), esto es:

$$\text{Region de confianza} = \text{Magnitud calculada} \cdot (1 \pm 2\sigma_R)$$

De esta manera, la misma formulación podrá servir para el régimen no permanente, una vez conocida la desviación típica relativa y el valor calculado por el modelo para cada instante, se evaluará la amplitud de la región de confianza que también será variable en el tiempo.

### **Utilidad de la consideración del análisis de incertidumbre en el proceso de modelación.**

El proceso de modelación es un proceso complejo, por el cual conseguimos que la resolución de unas ecuaciones matemáticas se represente el comportamiento de un sistema real. Es interesante considerar, de forma inherente al los propios resultados del modelo, cómo el error en la fase de toma de datos de entrada o de estimación de parámetros que intervienen en la calibración, va a trasladarse a los resultados del modelo.

Una de las utilidades más claras de la proposición de intervalos de confianza para las magnitudes calculadas, o de desviaciones típicas para las mismas, se encuentra en el proceso de calibración, más en concreto, en la fase o criterio de terminación del proceso de calibración (López, 2001; Shoemaker, 2004).

La calibración en su fase final, determinará que un modelo ha sido calibrado cuando se ajuste a ciertas mediciones consideradas como de referencia, cuando las predicciones se ajusten a la realidad en un número determinado de ocasiones. Sin embargo, también estas mediciones de referencia tendrán una cierta precisión, por tanto la comparación no es totalmente unívoca, sino que debe ser considerada asimismo de forma no determinista.

Una vez determinada la región de confianza para las magnitudes calculadas, y comparadas con las medidas, en caso de que se disponga de ellas, el modelador debe tomar una decisión sobre si el proceso de calibración ha terminado o se debe continuar destinando esfuerzos al mismo.

Pueden existir diversos criterios para la terminación del proceso de calibración, que deberán ser dispuestos por el usuario en función de la precisión que se desee para el modelo concreto. Por ejemplo se puede acotar el error relativo medio o la dispersión de los resultados obtenidos respecto a los calculados considerando esta como una varianza maestra. De esta manera, decidiremos que la calibración ha terminado cuando los intervalos de confianza de las predicciones y las modelaciones, se intersecten en una proporción determinada, por ejemplo en un 90% de las veces (López, 2001). Este es un criterio objetivo que puede resultar de gran utilidad al modelador y que permite la consideración del error tanto de las medidas de referencia como de las magnitudes calculadas, transmitido a las mismas a lo largo de la modelación matemática.

## EJEMPLO PRÁCTICO.

Para ilustrar el concepto de intervalo de confianza, se presenta de nuevo una modelación de la concentración de cloro en una línea de un sistema determinado a lo largo de un período de simulación diario.

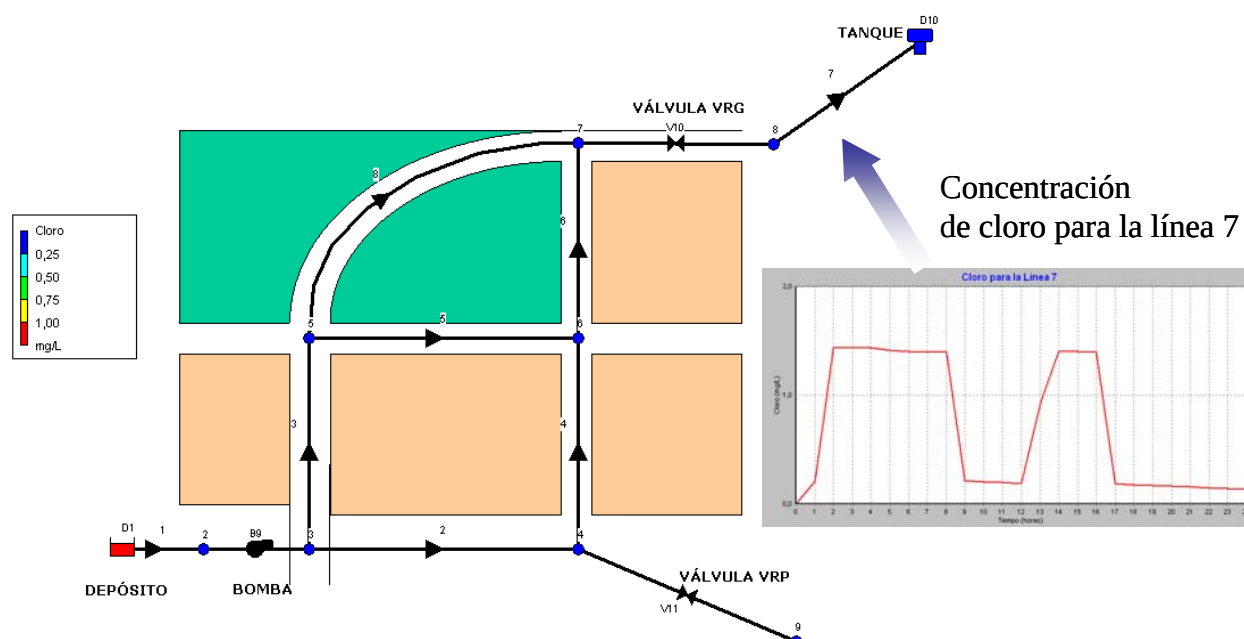


Figura 1. Ejemplo práctico para la aplicación del análisis de incertidumbre

La estimación de la concentración de cloro sobre la línea 7, descrita en la Figura 1 es, tal como muestra el modelo EPANET, una para cada instante de tiempo.

A continuación se propone un análisis de incertidumbre sobre este ejemplo sintético teniendo en cuenta tres posibilidades de error de forma simultánea:

- Se considera que la concentración de cloro en el depósito de cabecera (que provendrá de la medición en este punto para cada instante de tiempo) responde a una variable normal y tiene una desviación típica del 5%.
- Se considera que el coeficiente de velocidad de flujo  $K_b$ , que afecta de forma directa a la velocidad de reacción, tiene una variabilidad en su estimación del 20%. En este sentido este coeficiente no es un parámetro que provenga de la medición, sino que se estima a lo largo de la modelación. Este tipo de parámetros no tienen una desviación típica asociada en el sentido estadístico de la medición, sino que se encuentran dentro de unos intervalos que el modelador debe acotar, es en este sentido que se considera la posible variación del mismo
- Se considera que el coeficiente de reacción de cloro con la pared,  $K_w$ , de forma similar al caso anterior, tiene una posible variación de hasta el 20% alrededor de su valor por defecto.

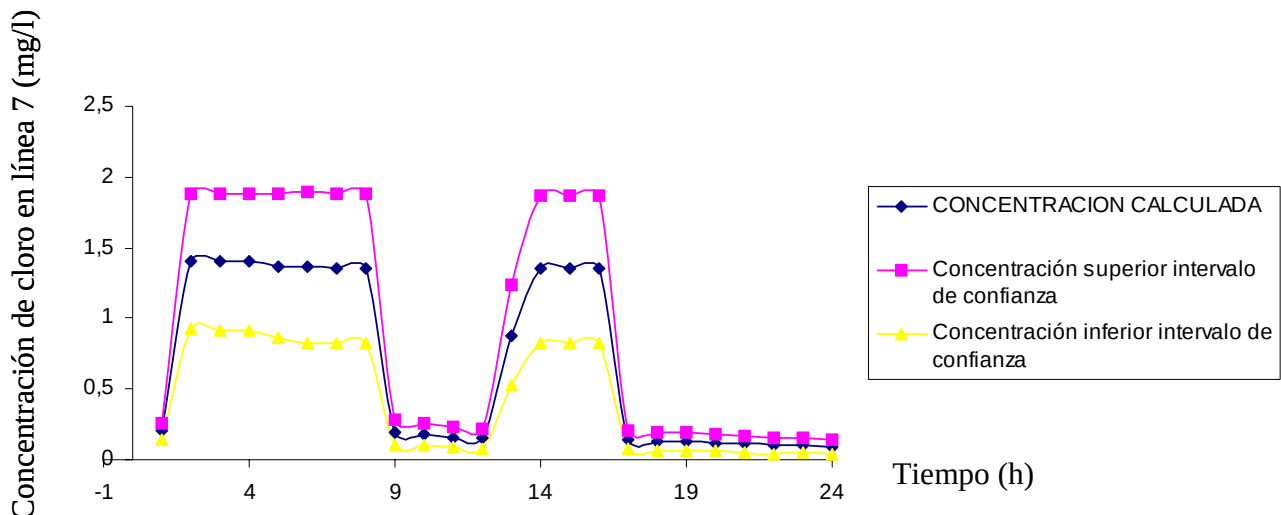
Así, se estudian en primer lugar los efectos relativos sobre la concentración de cloro de estas variaciones, que se presentan en la Tabla 1.

*Tabla 1. Efecto relativo y desviación típica calculada para concentración de cloro en línea 7.*

| INSTANTE<br>(h) | ER de variación en<br>concentración<br>inicial de cloro | ER<br>variación<br>$K_b$ | ER<br>Variación<br>$K_w$ | Desviación típica calculada<br>para concentración de cloro<br>en línea 7 |
|-----------------|---------------------------------------------------------|--------------------------|--------------------------|--------------------------------------------------------------------------|
| 1               | 1                                                       | 0                        | 0,5                      | 0,15                                                                     |
| 2               | 1                                                       | 0,04285714               | 0,61428571               | 0,169663533                                                              |
| 3               | 1                                                       | 0,04285714               | 0,62857143               | 0,171339262                                                              |
| 4               | 1                                                       | 0,04285714               | 0,64285714               | 0,172998761                                                              |
| 5               | 1,02189781                                              | 0,04379562               | 0,74452555               | 0,184628251                                                              |
| 6               | 1,029411765                                             | 0,05882353               | 0,82352941               | 0,194596113                                                              |
| 7               | 1,037037037                                             | 0,05925926               | 0,82962963               | 0,195315509                                                              |
| 8               | 1,037037037                                             | 0,05925926               | 0,84444444               | 0,196826677                                                              |
| 9               | 1,052631579                                             | 0,31578947               | 0,94736842               | 0,230559959                                                              |
| 10              | 0                                                       | 0,33333333               | 0,88888889               | 0,221108319                                                              |
| 11              | 1,25                                                    | 0,25                     | 0,875                    | 0,21937411                                                               |
| 12              | 1,333333333                                             | 0,26666667               | 0,93333333               | 0,226568606                                                              |
| 13              | 1,136363636                                             | 0,06818182               | 0,88636364               | 0,202540681                                                              |
| 14              | 1,037037037                                             | 0,04444444               | 0,81481481               | 0,192257543                                                              |
| 15              | 1,037037037                                             | 0,05925926               | 0,81481481               | 0,193792558                                                              |
| 16              | 1,037037037                                             | 0,05925926               | 0,81481481               | 0,193792558                                                              |
| 17              | 1,428571429                                             | 0,42857143               | 1                        | 0,24640269                                                               |
| 18              | 1,538461538                                             | 0,46153846               | 0,92307692               | 0,243373723                                                              |
| 19              | 0                                                       | 0,61538462               | 1,07692308               | 0,260177454                                                              |
| 20              | 0                                                       | 0,5                      | 1                        | 0,244948974                                                              |
| 21              | 1,818181818                                             | 0,54545455               | 1,09090909               | 0,264575131                                                              |
| 22              | 2                                                       | 0,6                      | 1,2                      | 0,277488739                                                              |
| 23              | 0                                                       | 0,6                      | 1,2                      | 0,268328157                                                              |
| 24              | 2,222222222                                             | 0,66666667               | 1,11111111               | 0,276887462                                                              |

Utilizando estos datos, se está en condiciones de estimar el intervalo de confianza de las predicciones de concentración de cloro alrededor de los valores calculados por EPANET, esto es lo que se presenta en la Figura 2, marcando los límites superior e inferior de las predicciones para cada instante de tiempo.

De esta manera se convierte lo que es una predicción determinista en una región de predicción, de forma que si se tuviesen mediciones para la comparación en la calibración, deberían considerarse también de forma no determinista, con sus propios intervalos de medición, dando así un carácter aleatorio al modelo.



*Figura 2. Intervalo de confianza alrededor de las magnitudes calculadas para la concentración de cloro en la línea 7.*

## CONCLUSIONES

En el presente trabajo se ha desarrollado la posibilidad de convertir los resultados de un modelo de calidad para las redes de abastecimiento de agua de carácter determinista, en una predicción estocástica comprendida dentro de unos intervalos de confianza con carácter no determinista.

Se ha demostrado el establecimiento de los intervalos de confianza de las predicciones del modelo en cuanto a la concentración de cloro en una de las líneas de un sistema a modo de ejemplo de cómo puede realizarse el propuesto análisis de incertidumbre. Este análisis de incertidumbre refleja, por un lado, cómo los errores en las mediciones de magnitudes que intervienen en la modelación; y por otro los parámetros que se estiman a lo largo de la misma, incide sobre la precisión de los resultados. Asimismo, este análisis debe ser combinado con un análisis de sensibilidad para establecer el efecto que tiene la dispersión de cada uno de los parámetros de entrada.

Esta consideración en base a intervalos de confianza en las predicciones del modelo tiene gran interés en el proceso de calibración, ya que permite una comparación con mediciones de concentración de cloro en la red, considerando éstas asimismo como variables aleatorias y de esta forma puede proponerse un criterio objetivo para terminación del proceso de calibración, independientemente de la experiencia del modelador.

## REFERENCIAS BIBLIOGRÁFICAS

- Canal de Isabel II. (1992). GRYM. *Un modelo para la gestión de la calidad de los ríos de la Comunidad de Madrid*.
- Chapra, S.C. (1997). *Surface Water Quality Modelling*. Mc Graw-Hill.
- EPA, (1987). *Rate, Constants and Kinetics Formulations in Surface Water Quality Modelling*. Second Edition.
- Evans, M. y Rosenthal, J. (2002). *Probabilidad y estadística. La ciencia de la incertidumbre*. Reverté.
- Instituto nacional de Ecología, México. (2005). *Riesgo e incertidumbre*. Consultado en Internet en <http://www.ine.gob.mx/ueajei/publicaciones/libros/400/cap5.html> en noviembre de 2005.
- López, Jiménez, P.A. (2001). “*Metodología para la calibración de modelos matemáticos de dispersión de contaminantes incluyendo regímenes no permanente*”. Tesis doctoral. Universidad Politécnica de Valencia.
- Pasha, M.F.;Lansey, K (2005) “*Analysis of Uncertainty on Water Distribution Hydraulics and Water Quality*” World Water Congress 2005. Impacts of Global Climate Change. World Water and Environmental Resources Congress 2005.Raymond Walton - Editor, May 15–19, 2005, Anchorage, Alaska, USA
- Mac Berthouex, P. y Brown, L. (1994). *Statistics for Environmental Engineers*. Lewis Publishers. U.S.A.
- Reichert P, Vanrolleghem P. (2001). “Identifiability and uncertainty analysis of the river water quality model”, *Water Sci tecn.* Vol 43, pp 329-38. 2005
- Scientific Software Group. 2005. *QUAL2E Manual*. Consultado en Internet en [http://www.scisoftware.com/products/qual2eu\\_overview/qual2eu\\_overview.html](http://www.scisoftware.com/products/qual2eu_overview/qual2eu_overview.html) en noviembre de 2005.
- Shoemaker,C.; Tolson, B.; Mendez, F. 2004. “Automatic Calibration and Uncertainty Analysis of a Watershed Model”. World Water and Environmental Resources Congress 2004. Gerald Sehlke, Donald F. Hayes, David K. Stevens - Editors, June 27 – July 1, 2004, Salt Lake City, Utah, USA
- Tze Ling Ng, J. Wayland Eheart, 2005. “*Effects of Discharge Permit Trading on Water Quality Reliability*,” *J. Water Resour. Plng. and Mgmt.*, Volume 131, Issue 2, pp. 81-88 (March/April 2005)

