

Métodos geoestadísticos para la elaboración de mapas de probabilidad de riesgo hidrogeotóxico (HGT) por altas concentraciones de As en las aguas subterráneas. Aplicación a la distribución de HGT en la provincia de Ávila (España)

Geostatistical methods to map the probability of hydrogeotoxic risk by high As concentrations in groundwater. Case study in Ávila province (Spain)

Guardiola-Albert, C.^{a1}, Pardo-Igúzquiza, E.^{a2}, Giménez-Forcada, E.^{b1}

^aInstituto Geológico y Minero de España (IGME) C/ Ríos Rosas, 23, 28003 Madrid (España).
E-mail: ^{a1} c.guardiola@igme.es, ^{a2} e.pardo@igme.es

^aInstituto Geológico y Minero de España (IGME) Unidad de Salamanca C/Azafranal, 48, 37001 Salamanca (España).
E-mail: ^{b1} e.gimenez@igme.es

Recibido: 02/11/2016

Aceptado: 11/01/2017

Publicado: 31/01/2017

RESUMEN

La presencia de As en las aguas subterráneas es un problema prioritario de salud pública e impone serias restricciones en el agua de consumo. Los mapas de probabilidad de superar el umbral permitido por la Organización Mundial de la Salud, OMS (10 µg/L) permiten delimitar las áreas que más riesgo presentan en relación con este parámetro. Las técnicas geoestadísticas constituyen una herramienta de uso común para elaborar estos mapas, aunque lamentablemente no hay un acuerdo sobre qué técnica es la más adecuada. El presente estudio recopila distintos criterios para decidir qué método presenta resultados más robustos. Se utilizan siete métodos de *kriging* no paramétrico en la estimación del mapa de probabilidad de que la concentración de As en manantiales de la provincia de Ávila supere el límite de 10 µg/L. La validación revela que uno de los mejores resultados es del *simplicial indicator kriging*, nunca antes tenido en cuenta en estudios sobre presencia de As geogénico en aguas subterráneas.

Palabras clave | arsénico; aguas subterráneas; riesgo hidrogeotóxico HGT; *kriging* de indicadores; validación cruzada.

ABSTRACT

The presence of As in groundwater is a priority public health issue because it imposes serious restrictions on drinking water. Mapping probabilities of exceedance of the threshold permitted by the World Health Organization, WHO (10 µg/L) allow delimiting the most vulnerable areas. The existing geostatistical techniques are a common tool for the evaluation of these maps, though, there is no agreement on which of the methods is the best. In this study different comparison criteria are illustrated. Seven non-parametric kriging methods are used to estimate the map of probability of exceeding the As concentration the limit of 10 mg/L in groundwater at the province of Ávila. Performed validation reveals that one the best results correspond to the simplicial indicator kriging, never before compared in studies of presence of geogenic As in groundwater.

Key words | arsenic; groundwater; hydrogeotoxicity risk; indicator kriging; cross validation.

INTRODUCCIÓN

En muchas regiones áridas y semiáridas el agua subterránea es la única fuente de agua potable disponible. Las altas concentraciones de Elementos Traza Geogénicos Potencialmente Tóxicos (ETGPT) en el agua de bebida pueden suponer un riesgo para la salud humana e imponer restricciones a la disponibilidad de agua apta para consumo.

Para estudiar la concentración de un determinado elemento traza, en relación con la concentración límite establecida para el agua de consumo, es necesario introducir el concepto HidroGeoToxicidad (HGT). La HGT se define como el cociente entre la concentración de un determinado ETGPT en una muestra de agua concreta y el valor límite máximo para ese elemento en el agua potable de acuerdo con la OMS (Giménez-Forcada *et al.*, 2016):

$$HGT_{ETGPT} = \frac{[ETGPT_{agua}]}{[ETGPT_{OMS}]}$$

La presencia de arsénico (As) en el agua subterránea y el riesgo potencial por HGT_{As} asociado se ha convertido en un tema prioritario de salud ambiental porque compromete la calidad del recurso y limita su uso como agua potable. El control tan estricto sobre la concentración de As en agua de ingesta se basa en su reconocimiento como cancerígeno por la Agencia Internacional de Lucha contra el Cáncer (IARC) (WHO, 2009).

La presencia de As geogénico en las aguas subterráneas en varias áreas sedimentarias cenozoicas españolas y los problemas de abastecimiento asociados (Aragón Sanz *et al.*, 2001; Sahún *et al.*, 2004; Mayorga *et al.*, 2013) se constataron especialmente a partir de la entrada en vigor del Real Decreto RD 140/2003, que limitaba la concentración de As de acuerdo a las nuevas recomendaciones de la OMS y fijaba la nueva concentración máxima en 10 µg/L. Una de las zonas más afectadas es el sur de la Cuenca del Duero donde un buen número de abastecimientos con agua subterránea superan el nuevo valor umbral. Aunque generalmente estos valores están solo ligeramente por encima de 10 µg/L, puntualmente se pueden alcanzar concentraciones de varias centenas de µg/L (Giménez-Forcada y Smedley, 2014). La provincia de Ávila es un buen ejemplo de lo que sucede en la zona. En ella se han constatado altas concentraciones de As en aguas de diversos manantiales. Su distribución se considera representativa de la concentración de As en las aguas subterráneas de la provincia.

Identificar las áreas con riesgo potencial por HGT_{As} es una herramienta de gran utilidad para los técnicos y gestores de las aguas subterráneas en la toma de decisiones. Para ello es de gran utilidad elaborar mapas de probabilidad de superar el valor umbral del As en las aguas subterráneas, y que permitan además relacionar espacialmente la presencia de As con distintos factores, tanto naturales como antrópicos.

La generación de estos mapas de probabilidad no es una tarea sencilla (Ryker, 2001). Por un lado, se debe de disponer de una red de control con una amplia cobertura espacial que, depende muchas veces de la posibilidad de tomar muestras en pozos privados. Por otra parte, la técnica matemática de interpolación elegida para estimar las concentraciones de As en los puntos no medidos influye mucho en el mapa resultante. A los problemas mencionados, hay que añadir que todo valor interpolado tiene una incertidumbre asociada. Una vez interpolados los datos hay que decidir cómo se expresan los resultados, con distintos colores o estadísticos. Esto puede influir mucho en la interpretación y la toma de decisiones, ya que afectará a la delimitación final de las zonas con riesgo por HGT_{As} . Utilizando un solo mapa, una de las aproximaciones más usada para evaluar la presencia de As en el agua subterránea de una zona, es la generación de un mapa de probabilidad de superar el umbral permitido de As en agua potable. Las técnicas geoestadísticas son las más utilizadas para obtener este tipo de mapas ya que a partir de los datos de concentración de As calculan directamente el mapa de probabilidades (Goovaerts, 1997).

En geoestadística, el enfoque habitual para la modelación de la incertidumbre local consiste en considerar los valores de una variable natural z desconocida como una realización de una variable aleatoria Z . En el caso aquí mostrado se considera que $z(\mathbf{u})$ son los valores de As medidos en laboratorio a partir de muestras de agua subterránea para n localizaciones $\mathbf{u}_j, j = 1, \dots, n$. El modelo de incertidumbre de z se obtiene a través de la función de distribución de probabilidad. La probabilidad de que z esté por debajo de un cierto umbral z_0 viene dada por la proporción de las mediciones que están por debajo de dicho umbral, que es la función de distribución acumulativa condicional (ccdf) F de esa variable z :

$$F(\mathbf{u}; z | (n)) = \text{Prob}\{Z(\mathbf{u}) \leq z_0 | (n)\} = \text{proporción de } z_j \leq z_0, \quad j=1, \dots, n \quad (1)$$

donde ' $| (n)$ ' expresa que z está condicionado a la información local, a las n mediciones de $z(\mathbf{u}_j)$.

Los estudios sobre riesgo potencial por HGT_{As} se centran en la probabilidad de que z esté por encima de un cierto umbral z_0 , o valor límite de concentración para agua potable (en As, 10 $\mu\text{g/L}$). Esta probabilidad se calcula como el complementario a 1 de la cdf:

$$\text{Prob}\{Z(\mathbf{u}) > z_0 | (n)\} = 1 - \text{Prob}\{Z(\mathbf{u}) \leq z_0 | (n)\} = 1 - \text{proporción de } z_j \leq z_0, \quad j=1, \dots, n \quad (2)$$

Hay que distinguir entre los métodos geoestadísticos paramétricos y los no paramétricos. En la geoestadística paramétrica, la distribución de los valores que puede tomar z se supone conocida y generalmente gaussiana. La geoestadística no paramétrica no es tan restrictiva ya que construye los modelos de distribución de probabilidad transformando los datos originales a una variable indicadora i (Gómez-Hernández, 1991) para el umbral z_0 :

$$i(z_0; z_j) = \begin{cases} 0, & \text{si } z_j \leq z_0 \\ 1, & \text{si } z_j > z_0 \end{cases} \quad (3)$$

Con esta transformación la función de distribución cdf (Ecuación (1)) se puede reescribir como el valor esperado condicional de una variable aleatoria indicadora I (cuya realización es i), dada la información (n) (Journel, 1983):

$$1 - F(\mathbf{u}; z_0 | (n)) = \text{Prob}\{Z(\mathbf{u}) > z_0 | (n)\} = E\{I(\mathbf{u}); z_0 | (n)\} \quad (4)$$

Los métodos para estimar el valor esperado de la variable indicadora i dan lugar a las distintas aproximaciones geoestadísticas paramétricas y no paramétricas para obtener el mapa de probabilidad de que no se supere el umbral z_0 . Los trabajos publicados comparando la diversidad de interpoladores geoestadísticos con indicadores (Brus *et al.*, 2002; Cattle *et al.*, 2002) muestran que lo más común es no encontrar grandes diferencias en la fiabilidad de los distintos resultados (Papritz y Dubois, 1999; Lark y Ferguson, 2004; Goovaerts *et al.*, 2005). Como consecuencia de ello no existe un acuerdo sobre cuál es el mejor método para la estimación de mapas de probabilidad de superar cierto umbral de concentración de un ETGPT.

El presente trabajo recopila de forma sencilla las pautas que los técnicos y gestores de las aguas subterráneas en materia de calidad del agua subterránea pueden seguir para decidir el método más fiable para elaborar estos mapas. La decisión sobre cuál de las técnicas empleadas estima mejores predicciones de la probabilidad de superar este umbral se realiza considerando las incertidumbres de la medida. Esto implica que se obtendrá el mejor mapa para diferenciar las zonas potencialmente peligrosas de las zonas no afectadas por riesgo $\text{HGT}_{\text{ETGPT}}$ y, por tanto, seguras. Los criterios se basarán en los valores de validación cruzada y los píxeles clasificados erróneamente.

Para mostrar los pasos a seguir y ver qué método arroja resultados más fiables se ha elegido como ejemplo la concentración de As en el agua de manantiales de la provincia de Ávila (España) y a partir de estos datos se ha elaborado el mapa de probabilidad de superación 10 $\mu\text{g/L}$ en esta provincia, utilizando siete interpoladores geoestadísticos de indicadores. Cuatro de ellos pertenecen a los métodos no paramétricos: (i) *simple indicator kriging* con estimación manual del variograma (Gómez-Hernández, 1991), (ii) *simple indicator kriging* con estimación automática del variograma utilizando el código AUTO-IK con un umbral (Goovaerts, 2009), (iii) *simple indicator kriging* con media local variable (Goovaerts, 1997) y (iv) *simplicial indicator kriging* (Tolosana-Delgado *et al.*, 2005; 2008); y tres a los métodos paramétricos: (v) *bayesian indicator estimation* (Guardiola-Albert y Pardo-Igúzquiza, 2011), (vi) *compositional bayesian indicator estimation* (Guardiola-Albert y Pardo-Igúzquiza, 2001) y (vii) *disjunctive kriging* (Rivoirard, 1994). Los métodos *simplicial indicator kriging*, *bayesian indicator estimation* y *compositional bayesian indicator estimation* son poco conocidos a pesar de sus buenos resultados.

MATERIAL Y MÉTODO

Interpolación de indicadores (*indicator kriging*)

El término general de *kriging* designa al interpolador geoestadístico que ofrece el mejor estimador como una combinación lineal de los datos disponibles, sin sesgo y con varianza de error mínima. Para calcular los pesos del estimador lineal, el *kriging* utiliza un modelo de variabilidad espacial a través de la función denominada variograma, que describe el grado de dependencia espacial de la variable aleatoria. La función de variograma utilizada en el *kriging* (modelo de variograma) se ajusta a los valores de variograma calculados experimentalmente con los datos. Cuando lo que interesa es la probabilidad de superar un umbral, y no tanto la variable en sí misma, se emplea el *indicator kriging*. Como se ha explicado en la Introducción, éste método se basa en la transformación de unos datos continuos a una variable binaria en función de un umbral. La estimación del valor esperado de esta nueva variable indicadora i , es decir, su mapa interpolado, es directamente la probabilidad que se quiere obtener. En todos los métodos se han utilizado el mismo número máximo de puntos cercanos utilizados en la estimación (12).

Otra de las ventajas de los métodos geoestadísticos es que la estimación (interpolación) de una variable con pocas medidas se puede mejorar utilizando una variable secundaria de la que se disponga de datos en todo el dominio de estudio. En este sentido, algunas de las técnicas geoestadísticas aplicadas en el presente estudio tendrán en cuenta la relación de la presencia As con la geología. La provincia de Ávila es un ejemplo para establecer esta relación ya que en ella los materiales ígneos y metasedimentarios del Sistema Central contrastan con los sedimentos terciarios del sur de la Cuenca del Duero.

Dependiendo de las asunciones tomadas sobre las variables aleatorias Z e I se han desarrollado distintos métodos de estimación puntual que han sido aplicados en el presente trabajo:

- (i) *Simple indicator kriging* (Gómez-Hernández, 1991): método no paramétrico que asume que la media de la variable indicadora I es constante y conocida. El modelo de variograma se ha ajustado con Sgems (Remy *et al.*, 2009).
- (ii) *Simple indicator kriging* (Gómez-Hernández, 1991): método no paramétrico que asume que la media de la variable indicadora I es constante y conocida y modela automáticamente el variograma utilizando el código AUTO-IK (Goovaerts, 2009) con un umbral.
- (iii) *Simple indicator kriging* con media local variable (Goovaerts, 1997): método no paramétrico en el que la media puede variar localmente. Esta media se calcula en función de otra variable, en este caso la geología del área de estudio. En este caso se asume que la media es constante dentro de cada afloramiento geológico. Esta hipótesis se puede tomar si se considera que existe relación entre la presencia de As y la litología de la zona. Se ha calculado con el código Sgems (Remy *et al.*, 2009).
- (iv) *Simplicial indicator kriging* (Tolosana-Delgado *et al.*, 2005; 2008): método no paramétrico que identifica el espacio muestral de los vectores de probabilidad con el simplex, el espacio muestral de las composiciones. Además este método tiene en cuenta el error de la propia clasificación de la muestra, asumiendo los 0's como valores cercanos a 0 y los 1's como valores cercanos a 1. Una vez en el espacio del simplex, el *kriging* se ha obtenido utilizando Sgems (Remy *et al.*, 2009).
- (v) *Bayesian indicator estimation* (Guardiola-Albert y Pardo-Igúzquiza, 2011): método paramétrico que tiene en cuenta el error de la propia muestra (el 0 y el 1 puede que no sean exactos) e información previa como la relación con el contexto geológico. Define el estimador dentro del marco bayesiano como la mediana de la distribución *a posteriori*. Se ha utilizado un código propio en FORTRAN para aplicación de este método.
- (vi) *Compositional bayesian indicator estimation* (Guardiola-Albert y Pardo-Igúzquiza, 2001): similar al método anterior pero considerando la estimación como una composición de dos partes, cuyo dominio de definición es el simplex. Se ha utilizado un código propio en FORTRAN para aplicación de este método.
- (vii) *Disjunctive kriging* (Rivoirard, 1994): es un método no lineal que asume que la distribución univariada y bivariada de la variable es conocida. Para este método se ha utilizado la herramienta implementada en Arcmap (ESRI, 2011).

La Figura 1 muestra los pasos que pueden seguirse para conseguir un mapa que defina las zonas más afectadas por riesgo hidrogeotóxico por HGT_{As} , tomando como ejemplo los datos de concentración de As en manantiales de la provincia de Ávila.

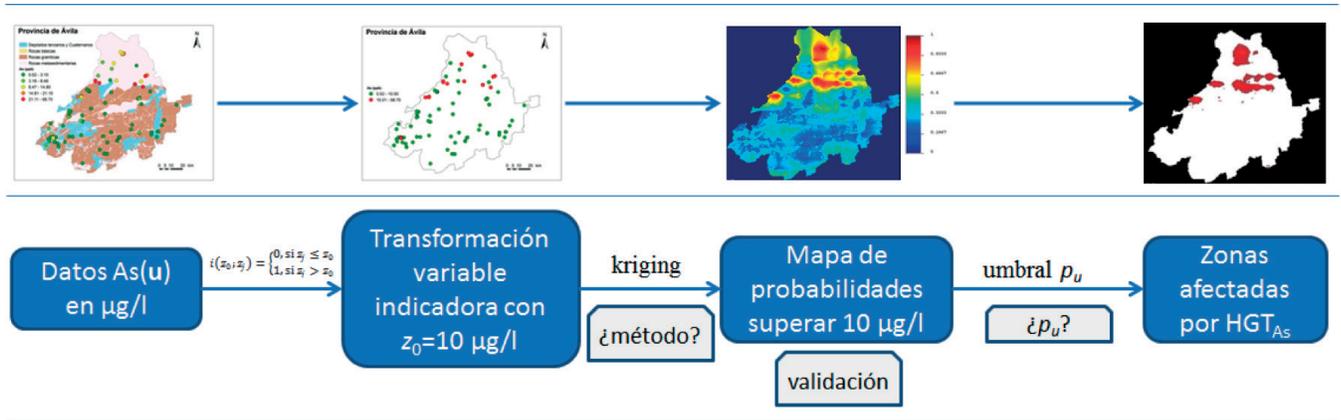


Figura 1 | Diagrama donde se indican de forma esquemática todas las etapas para la consecución de un mapa de posibles zonas afectadas por riesgo hidrogeotóxico (HGT_{As}), tomando como ejemplo el mapa de probabilidad de HGT_{As} en la provincia de Ávila.

Validación

Para la validación y comparación de los siete mapas de probabilidad obtenidos con los diferentes métodos descritos se ha aplicado el método de validación cruzada (Isaaks y Srivastava, 1989; Goovaerts, 1997). Esta técnica consiste en eliminar una observación, $i(\mathbf{u}_j)$, y re-estimar su valor, $i^*(\mathbf{u}_j)$, con los valores vecinos. Esto se ha realizado sucesivamente con todos los datos, obteniéndose un valor del error de estimación para cada dato. A partir de estos errores se ha calculado el error medio (EM) y el error medio absoluto (EMA):

$$EM = \frac{1}{n} \sum_{j=1}^n [i^*(\mathbf{u}_j) - i(\mathbf{u}_j)] \tag{5}$$

$$EMA = \frac{1}{n} \sum_{j=1}^n |i^*(\mathbf{u}_j) - i(\mathbf{u}_j)| \tag{6}$$

Falivene *et al.* (2010) demuestran que las comparaciones realizadas utilizando como criterio el mínimo error de la validación cruzada pueden ser completamente erróneas. Estos autores demuestran que la comparación es mucho más efectiva si se busca el método que dé el mínimo error y el menor suavizado que preserve la mayor variabilidad espacial. El suavizado resultante de los mapas interpolados se mide con la desviación estándar de los valores estimados en la validación cruzada ($Dst VC$):

$$Dst VC = \sqrt{\frac{\sum_{j=1}^n (i^*(\mathbf{u}_j) - \bar{i^*}(\mathbf{u}))^2}{n}} \tag{7}$$

donde $\bar{i^*}(\mathbf{u}) = \frac{1}{n} \sum_{j=1}^n i^*(\mathbf{u}_j)$ es la media de los valores estimados en la validación cruzada.

Si $Dst VC$ es más pequeño reflejará que las estimaciones son más suaves espacialmente. Un algoritmo de interpolación será mejor si puede proporcionar menos suavizado (Isaaks y Srivastava, 1989; Olea y Pawlowsky, 1996; Journel *et al.*, 2000), y así podrá preservar mejor la variabilidad lateral de los valores de los datos. Siguiendo este criterio cuanto mayor es la $Dst VC$, mejor es el método.

Otro criterio de comparación, con la misma filosofía de la validación cruzada, es el de los residuos ortonormales (Kitanidis, 1991). A diferencia de esta última, los residuos ortonormales se calculan siguiendo una secuencia concreta en la eliminación del punto a estimar. En el presente estudio los residuos ortonormales se han calculado ordenando los datos siguiendo la dirección nortesur, ya que visualmente los datos presentan una tendencia en esta dirección. Kitanidis (1991) sugiere que en estos casos este orden es el adecuado. La distribución de los residuos ortonormales, considerados como una variable aleatoria, deberían tener una media ($Q1$) cercana a 0, una varianza ($Q2$) próxima a 1 y no deberían estar correlacionados.

Tabla 1 | Resumen de los criterios utilizados en la validación y comparación de los mapas de probabilidad obtenidos. Estos estadísticos están basados en los errores obtenidos de la validación cruzada y los residuos ortonormales.

	Abreviatura en este trabajo	Referencia	Valor óptimo	Valor deficiente
Basados en los errores (residuos) de la validación cruzada				
Error Medio	<i>EM</i>	Isaaks y Srivastava (1989)	0	alto
Error Medio absoluto	<i>EMA</i>	Isaaks y Srivastava (1989)	0	alto
Desviación estándar de los valores estimados	<i>Dst VC</i>	Falivene <i>et al.</i> (2010)	mayor	menor
Tasa de clasificación errónea de píxeles	<i>% clasificación errónea</i>	Lark y Ferguson (2004)	0	100
Basados en los errores (residuos) ortonormales				
Media de los errores	<i>Q1</i>	Kitanidis (1991)	0	alto
Varianza de los errores	<i>Q2</i>	Kitanidis (1991)	1	lejos de 1

Además, se debe tener en cuenta el problema de la clasificación errónea de píxeles, que puede tener consecuencias graves a la hora de tomar decisiones. Es por esto que se consideran mejores aquellos métodos con un número mínimo de píxeles mal clasificados, porque se evitan errores a la hora de establecer las áreas más vulnerables por riesgo de HGT. Esta decisión se toma cuando la probabilidad estimada (p^*) supera cierto umbral de probabilidad (p_u). La selección de este umbral de probabilidad es la que marcará qué regiones serán declaradas como susceptibles de sufrir un riesgo potencial HGT y cuáles no. En la práctica se elegirá este umbral de tal forma que el riesgo de una clasificación errónea sea aceptable. Según proponen Goovaerts (1997) y Lark y Ferguson (2004) se eligen una serie de valores p_u y se evalúa la tasa de clasificación errónea con los resultados de la validación cruzada. En general para valores de p_u pequeños la tasa de clasificación errónea será mayor, ya que habrá más datos de validación que erróneamente se hayan considerados como menores que el umbral. Si p_u es grande, también la tasa de clasificación errónea será grande al haber también muchos datos clasificados como mayores que el umbral. Los valores intermedios de p_u serán los que minimicen la tasa de clasificación errónea. Tal y como propone Goovaerts (1997) se ha realizado un gráfico con los datos de la validación cruzada representando las tasas de clasificación errónea frente a los distintos umbrales p_u entre 0 y 1 variando en incrementos de 0.05. De esta forma se puede elegir aquel para el cual los errores en la clasificación sean mínimos. Siguiendo este razonamiento, el mejor método será aquel que de menos errores de clasificación para un rango mayor de p_u . Este método para evaluar los errores en la declaración de zonas con riesgo HGT potencial o no, es más simple y claro que otros existentes (Rousseau, 1980; Juang *et al.*, 2004).

Área de estudio y datos de concentración de As

La zona de estudio elegida para contrastar los distintos métodos es la provincia de Ávila (Figura 2), donde los materiales ígneos y metasedimentarios del Sistema Central contrastan con los sedimentos detríticos terciarios y cuaternarios de la Cuenca del Duero y de otras áreas sedimentarias menores, como el valle del Amblés o el Valle del Corneja, ambas asociadas al Dique Messejana-Plasencia (DMP). Este dique dolerítico se extiende a lo largo de más de 500 km, desde el S de Portugal y desaparece

bajo los sedimentos terciarios de la cuenca, en la provincia de Ávila. A esta estructura se asocian las cuencas *pull-apart* terciarias del Valle Amblés y Corneja, que forman parte también de la zona de estudio. Se han detectado anomalías significativas de As en la prolongación de esta estructura bajo el relleno terciario de la cuenca del Duero (Giménez-Forcada y Smedley, 2014; Rey-Moral *et al.*, 2016). Aproximadamente la cuarta parte de la provincia de Ávila corresponde a los depósitos terciarios de la Cuenca del Duero, mientras que el resto del territorio se encuentra sobre el Sistema Central (Figura 2). Ambos dominios entran en contacto por el cabalgamiento del Sistema Central sobre la Cuenca del Duero.

Las investigaciones realizadas hasta la fecha coinciden en atribuir la presencia de As a las reacciones agua-sedimento en determinado ambiente de pH y redox (García-Sánchez *et al.*, 2005; Gómez *et al.*, 2006; Delgado *et al.*, 2009; Giménez-Forcada y Smedley, 2014, entre otros) y atribuyen las altas concentraciones de As y otros ETGPT formadores de aniones a procesos de movilización por desorción de los oxi-hidróxidos de Fe y Mn presentes en el sedimento, en condiciones oxidantes y pH alcalino. Por otra parte, Giménez y Smedley (2014) establecen una relación entre la distribución de As y la estructura de la Cuenca del Duero. Recientemente se ha podido comprobar el control estructural de la distribución de anomalías de As en base a datos de gravimetría y magnetismo (Rey-Moral *et al.*, 2016).

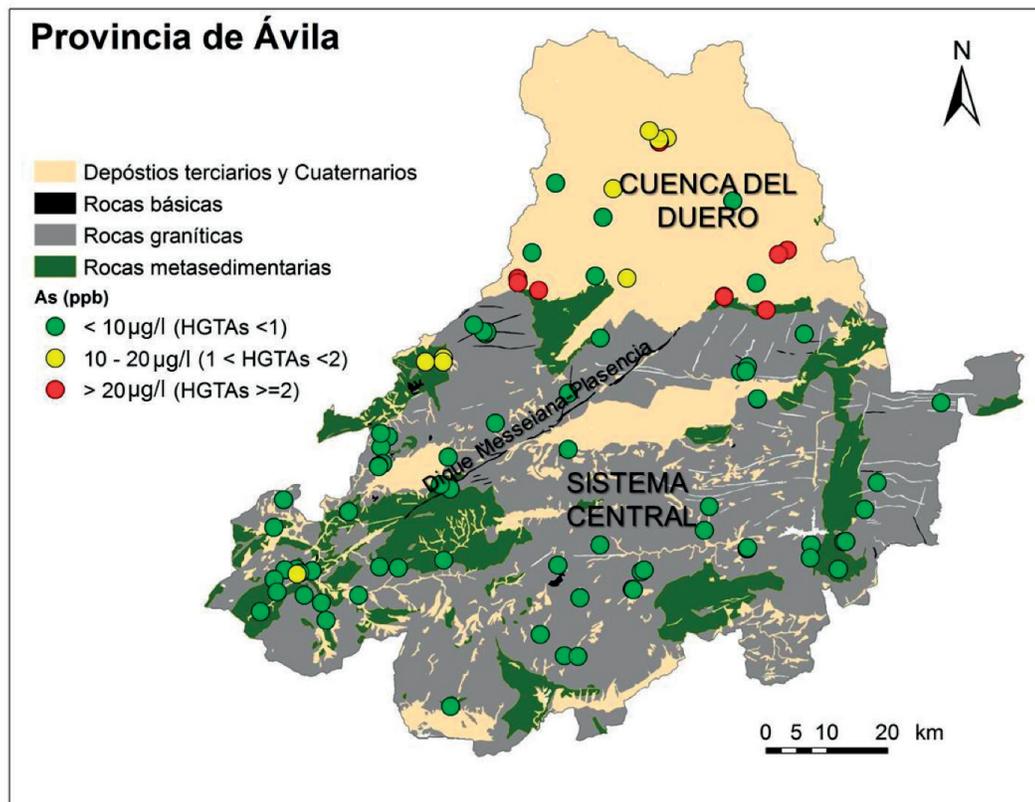


Figura 2 | Área de estudio: provincia de Ávila. En color de fondo se diferencian las distintas litologías presentes. Los círculos reflejan los valores de concentración de As. Los valores en verde indican ausencia de riesgo HGT_{As}, los amarillos un cierto grado de peligrosidad por HGT_{As} y los rojos que existe un riesgo hidrotóxico potencial por HGT_{As}.

El presente estudio se basa en una campaña de muestreo de 87 manantiales, entre junio y agosto de 2006, sobre un área de 8050 km². Previamente a la recogida de la muestra se registraron in situ valores de temperatura, pH y potencial redox, así como las coordenadas geográficas del punto de muestreo. Se tomaron dos alícuotas de agua: una para el análisis de cationes (acidificada con HNO₃, 1% v/v) y otra para análisis de aniones. En ambos casos las muestras fueron previamente filtradas a través de filtros de celulosa de 0.45 µm. Las muestras se conservaron a 4°C hasta su análisis. La alcalinidad (ion bicarbonato) se

determinó por volumetría, usando naranja de metilo como indicador. Los iones cloruro, nitrato y sulfato fueron determinados con un Autoanalizador ALKEMP 501. Los metales calcio, magnesio, sodio y potasio fueron analizados con Espectrometría de Absorción Atómica (AAS). El As fue determinado por ICP-MS (límite de detección 1 µg/L). En todos los casos el error de balance iónico fue igual o inferior a 3%.

En la Figura 3 se observa que las concentraciones de As registradas en las aguas subterráneas de la provincia de Ávila son elevadas en el área correspondiente a los depósitos terciarios y cuaternarios de la Cuenca del Duero y otras depresiones terciarias, mientras que en el ámbito ígneo-metamórfico del Sistema Central se identifican valores de concentración de As disuelto más bajos. Aunque en los materiales ígneo-metamórficos ocasionalmente también se pueden encontrar valores altos de As, en general puede verse que son menores ya que su tercer cuartil es casi igual que el primer cuartil del As en los depósitos terciarios y cuaternarios. Esto muestra el control geológico del contenido de As. Gracias a esta relación, entre el As y la litología, la estimación de los mapas de probabilidad puede mejorarse utilizando el mapa geológico de la provincia de Ávila como variable secundaria relacionada con la concentración de As.

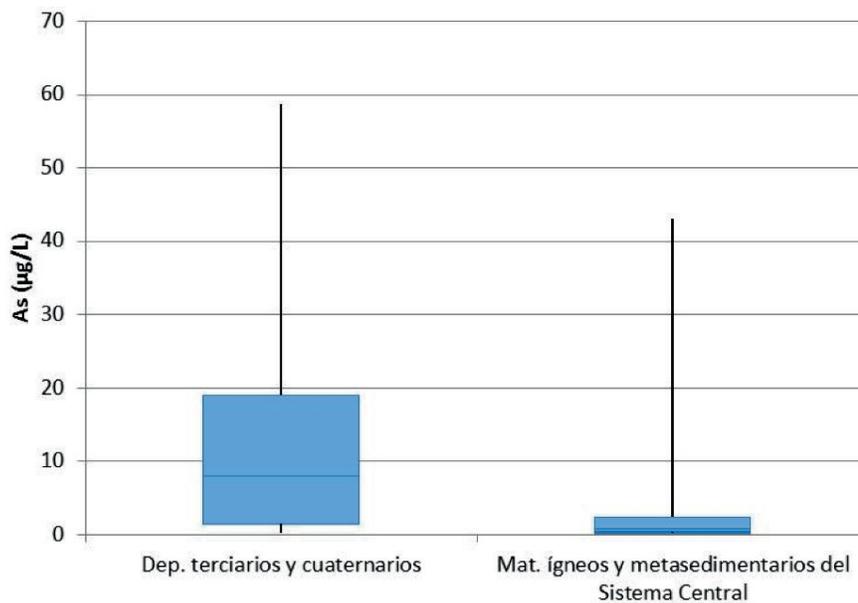


Figura 3 | Diagrama de cajas y bigotes de los valores de concentración de As disuelto, diferenciando entre los manantiales situados sobre depósitos terciarios y cuaternarios y los localizados sobre los materiales ígneo-metamórficos del Sistema Central. El tamaño de las cajas se determina con la diferencia entre el primer y el tercer cuartil. Los bigotes se extienden hasta los valores máximo y mínimo. La situación de la línea azul dentro de la caja corresponde al valor de la mediana.

RESULTADOS

Tras la aplicación de los siete métodos descritos se han obtenido siete mapas de probabilidad de riesgo potencial por HGT_{As} (Figura 4). Como puede observarse hay diferencias considerables entre los distintos mapas obtenidos. Comparando los mapas 4(a) (ajuste manual del variograma con alcance 18 km) y 4(b) (ajuste automático del variograma con alcance 6 km), se puede observar que el ajuste automático modela un alcance que probablemente sea demasiado pequeño, lo que se ve reflejado en las formas circulares del mapa de probabilidades. Esto significa que las técnicas de ajuste automático pueden dar lugar a mapas pocos realistas con artefactos espaciales que nada tienen que ver con la realidad.

Los mapas obtenidos con las técnicas que consideran el mapa geológico como información previa, 4(c), 4(e), y 4(f), son diferentes a los demás ya que se observa que incluyen parte del comportamiento espacial de la geología presente en la zona. Si

se tiene en cuenta el carácter composicional de las probabilidades, 4(d) y 4(f), en general se obtienen valores más extremos de probabilidad y menos valores intermedios.

Los estadísticos de validación (EM , EMA , $Dst VC$, $Q1$ y $Q2$) también presentan mucha variabilidad y dificultad para elegir qué mapa es el mejor y qué método se debe aplicar (Tabla 2). Para el método de *disjunctive kriging* no fue posible obtener los residuos ortonormales por la dificultad en la programación del camino a seguir en la validación en el software utilizado para este caso (ArcGIS): Con respecto a los errores de los distintos métodos, todos presentan valores similares de EM y EMA , muy próximos a 0. El *simplicial indicator kriging* es el que presenta mejores valores con un EM de -0.02 y un EMA de 0.11. Por otro lado, el método que da peores errores es el *compositional bayesian indicator kriging* con 0.35 de EM y 0.42 de EMA . Observando los valores de $Dst VC$ tampoco hay grandes diferencias entre los métodos usados, siendo de nuevo el más óptimo, el que da menos suavizado ($Dst VC$ más grande), el *simplicial indicator kriging* ($Dst VC$ igual a 0.4). Para los estadísticos de los residuos ortonormales, el método más robusto, con respecto a ambos estadísticos $Q1$ y $Q2$, es el *simple indicator kriging* con ajuste manual del variograma (-0.28 de $Q1$ y 1.11 de $Q2$) y con media local variable (0.15 de $Q1$ y 1.20 de $Q2$). El resto de métodos no serían muy robustos según estos dos estadísticos con valores de $Q1$ lejanos de cero o de $Q2$ muy diferentes a 1.

Tabla 2 | Valores de los estadísticos utilizados en la comparación de los mapas de probabilidad según el estudio de validación cruzada realizado: EM (error medio), EMA (error medio absoluto), $Dst VC$ (desviación estándar de los valores predichos) y $Q1$ y $Q2$ (estadísticos de los residuos ortonormales). En negrita se destacan los mejores valores obtenidos para dichos estadísticos.

	ME	MAE	$Std CV$	$Q1$	$Q2$
<i>Simple indicator kriging</i> variograma manual	-0.02	0.13	0.34	-0.28	1.11
<i>Simple indicator kriging</i> variograma automt.	-0.02	0.15	0.32	-0.13	0.15
<i>Simple indicator kriging</i> media variable	-0.05	0.15	0.37	0.15	1.20
<i>Simplicial indicator kriging</i>	-0.02	0.11	0.40	0.46	0.23
<i>Bayesian indicator kriging</i>	0.23	0.34	0.23	-1.41	4.92
<i>Compositional bayesian indicator kriging</i>	0.35	0.43	0.24	-1.63	5.11
<i>Disjunctive kriging</i>	0.01	0.22	0.20		
Todos los puntos se declaran no tóxicos	0.79	0.79			
Todos los puntos se declaran tóxicos	-0.21	0.21			
Máxima entropía: Todos los puntos tienen una probabilidad del 50% de ser tóxicos	0.29	0.5			

La Figura 5 representa el porcentaje de clasificación errónea en función del umbral de probabilidad (p_u) elegido para la decisión de declarar una zona con riesgo potencial por HGT_{As} o segura. El mínimo de píxeles mal clasificados (6.9%) se obtiene con el método más sencillo (*simple indicator kriging*), con una probabilidad de corte (p_u) de 0.3. Las tasas de clasificación errónea más bajas se obtienen para los métodos de *simple indicator kriging*, *simple indicator kriging* con media local variable y *simplicial indicator kriging*. En general, y para todos los métodos, la p_u que da menos píxeles mal clasificados es la de 0.5. Para los distintos valores de p_u se obtienen una media de un 15% de clasificación errónea en estos tres métodos.

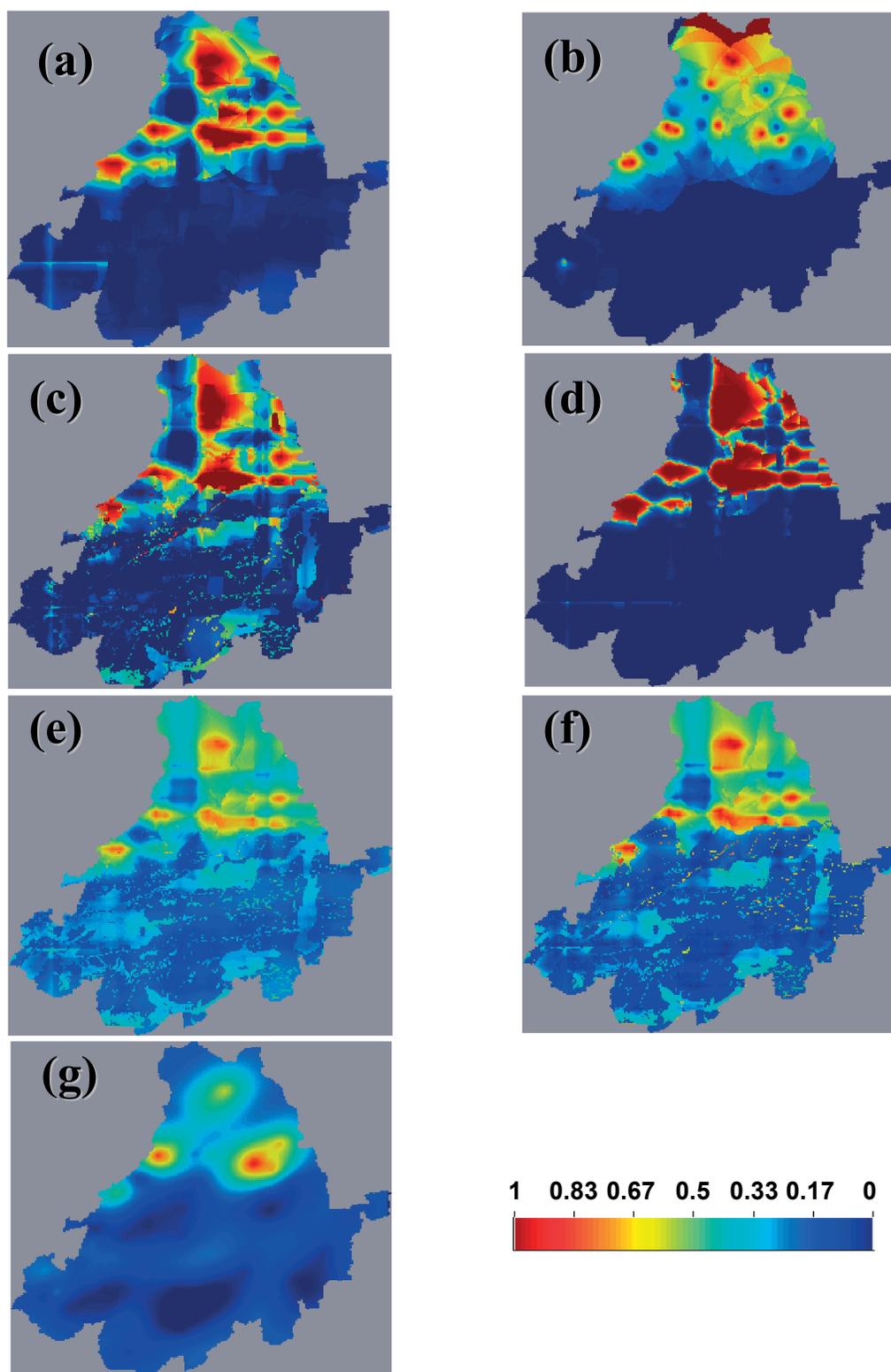


Figura 4 | Mapas de mapas de probabilidad riesgo HGT_{AS} en la provincia de Ávila. (a) *simple indicator kriging* con ajuste manual del variograma, (b) *simple indicator kriging* con ajuste automático del variograma, (c) *simple indicator kriging* con media local variable, (d) *simplicial indicator kriging*, (e) *bayesian indicator kriging*, (f) *compositional bayesian indicator kriging* y (g) *disjunctive kriging*.

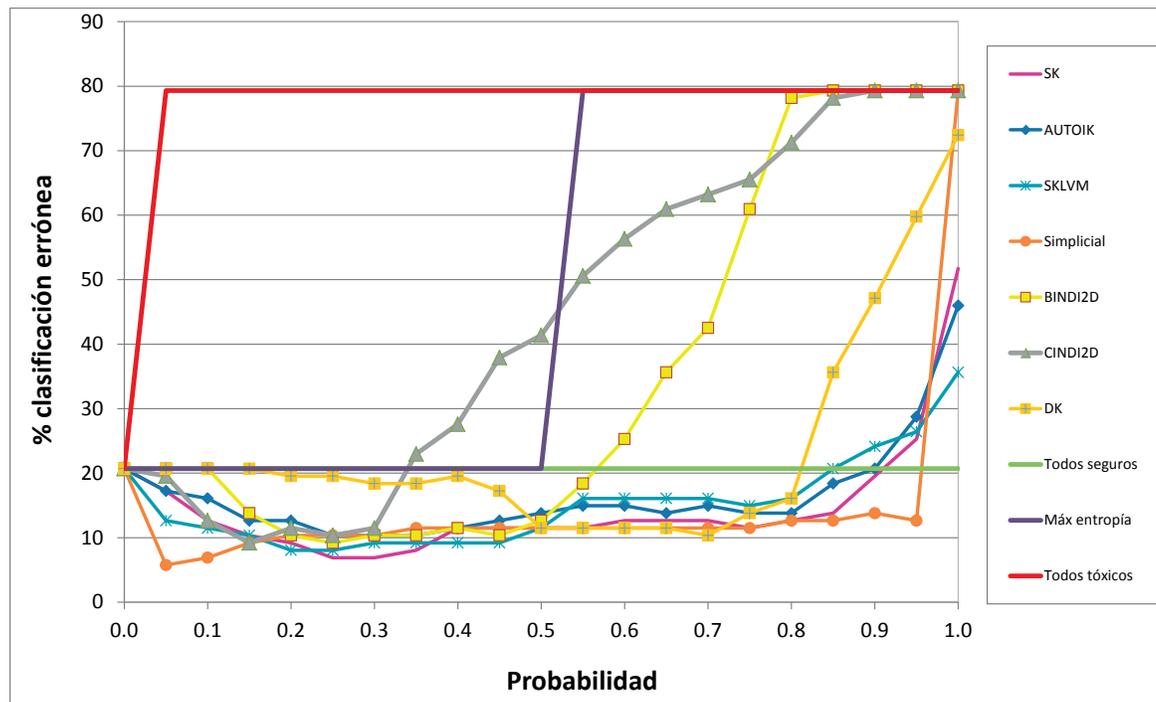


Figura 5 | Porcentaje de píxeles mal clasificados como con riesgo HGT_{AS} o seguros según cada uno de los métodos utilizados. SK: *simple indicator kriging* con ajuste manual del variograma, AUTOIK: *simple indicator kriging* con ajuste automático del variograma, SKLVM: *simple indicator kriging* con media local variable, Simplicial: *simplicial indicator kriging*, BINDI2D: *bayesian indicator kriging*, CINDI2D: *compositional bayesian indicator kriging*, DK: *disjunctive kriging*.

DISCUSIÓN Y CONCLUSIONES

El presente trabajo quiere presentar una serie de pautas y recomendaciones para los técnicos y responsables de la administración pública en materia de control de calidad de las aguas, que les ayude a decidir qué método o métodos son más adecuados para elaborar mapas de probabilidad de riesgo HGT. A continuación se discuten los resultados obtenidos y se describe el proceso de análisis que permite elegir el mapa de probabilidades óptimo.

En un primer análisis, según algunos de los criterios elegidos para validar los mapas de probabilidad obtenidos por los distintos métodos (EM, EMA y *Std CV*), se considera como mejor método el *simplicial indicator kriging*. Este método no había sido utilizado anteriormente en estudios sobre distribución de As en agua subterránea y la HGT_{AS} asociada. D'Or *et al.* (2008) señalan lo importante que puede ser tener un número importante de muestras por debajo del límite de detección. El *simplicial indicator kriging*, así como el *bayesian indicator kriging* y el *compositional bayesian indicator kriging*, permiten tener en cuenta en cierto modo este porcentaje ya que no asignan un valor exacto de 0 o 1 a la variable indicadora, asumiendo un pequeño porcentaje de error en las medidas.

Una de las críticas que se hacen a menudo a la validación cruzada es que no tiene en cuenta el agrupamiento de las muestras (*clustering*). Por ello, se decidió utilizar otros estadísticos como los residuos ortonormales, cuyos mejores resultados son los correspondientes al *simple indicator kriging*, el método más sencillo, y *simple indicator kriging* con media local variable, que tiene en cuenta la relación entre la concentración de As en el agua subterránea y la geología del entorno. Este último método utiliza la geología para el cálculo de la media local variable, dando como resultado un mapa más detallado que los métodos que utilizan sólo las concentraciones de As (Figura 4(c)). Por otro lado todos los mapas que han tenido en cuenta la litología de la zona estudiada (Figuras 4 (c), (e) y (f)) reflejan un mayor contraste entre las áreas sedimentarias (con altas concentraciones de As en las aguas subterráneas) y las del Sistema Central, donde generalmente los valores de concentración de As son más bajos, a excepción de las

fracturas donde se aprecia un HGT_{As} mayor, especialmente en torno al trazado del dique Messejana-Plasencia, corroborando de este modo la hipótesis sobre el control estructural de la distribución de As (Giménez-Forcada y Smedley, 2014; Rey-Moral *et al.*, 2016).

El porcentaje de píxeles mal clasificados también aporta una valiosa información sobre el poder de predicción de los diferentes métodos. Los métodos más sencillos tienen unas tasas de clasificación errónea bastante buenas, que oscilan entre el 10 y el 15%. Es interesante ver la comparación entre las tasas de mala clasificación con los casos extremos en los que se declarara a todos los puntos de muestreo sin riesgo HGT_{As} (línea verde de la Figura 5) o todos con riesgo potencial HGT_{As} (línea roja de la Figura 5). Parece que en este caso de estudio se cometería menos error siendo optimistas y declarando toda la zona como segura que optar por ser alarmistas. El caso denominado de máxima entropía (línea morada de la Figura 5) equivaldría a darle a todos los puntos de muestreo la misma probabilidad de ser potencialmente tóxicos, como de no serlo.

Otra forma de profundizar en la capacidad de predicción de los distintos métodos es diferenciar entre la declaración de píxeles con cierto grado de peligrosidad por HGT_{As} o seguros, a la hora de evaluar la mala clasificación de los mismos. Esto se realiza con tablas de contingencia como las de la Figura 6. Aplicando dos umbrales de probabilidad, $p_u = 0.3$ y $p_u = 0.5$ y la mencionada tabla de contingencia se han obtenido los resultados resumidos en la Tabla 3. En general el método con mejores predicciones y menos falsas alarmas y esperanzas es el *simple indicator kriging*, aunque el resto de métodos tampoco tienen porcentajes importantes de mala predicción. Destaca quizás el *disjunctive kriging*, por su mayor tasa de falsa esperanza y su menor probabilidad de falsas alarmas. Es curiosa la diferencia en el éxito del método si tenemos en cuenta por igual tanto el éxito en la predicción como las clasificaciones erróneas (Figura 5) o si las diferenciamos (Tabla 3). Los peores métodos, según la Figura 5, son valorados como de los mejores en la Tabla 3. En todo caso la reflexión más importante es que los métodos sencillos como el *simple indicator kriging*, dan muy buenos resultados según todos los criterios utilizados, siempre que el ajuste del variograma se haga de forma manual o que se compruebe que el ajuste automático tiene sentido físico. Por ello debe ser considerado como buen método para el estudio de la distribución de probabilidades de riesgo HGT_{As} .

Número de píxeles clasificados		Verdad	
		Seguro	Riesgo HGT
Predicción	Seguro	SS	SR
	Riesgo HGT	RS	RR
Probabilidad de falsa esperanza		$\frac{SR}{SR + SS}$	
Probabilidad de falsa alarma		$\frac{RS}{RS + RR}$	

Figura 6 | Tabla de contingencia para el estudio de las falsas alarmas y declaraciones erróneas de puntos de abastecimiento seguros.

De todos los análisis realizados sobre los distintos métodos los autores concluyen que el mejor método de los aplicados es el *simple indicator kriging*, con media local variable. Con él se obtienen los mejores resultados de todos los estadísticos de validación, ofrece una baja tasa de clasificación errónea de píxeles, la posibilidad de incluir información externa y relacionada con la distribución de concentraciones de As, tiene capacidad de dar baja probabilidad de falsas zonas seguras o potencialmente tóxicas, así como una alta probabilidad de éxito en la detección de zonas seguras y con riesgo potencial por HGT_{As} .

Tabla 3 | Probabilidades obtenidas siguiendo la tabla de contingencia de la Figura 6. SK: *simple indicator kriging* con ajuste manual del variograma, AUTOIK: *simple indicator kriging* con ajuste automático del variograma, SKLVM: *simple indicator kriging* con media local variable, Simplicial: *simplicial indicator kriging*, BINDI2D: *bayesian indicator kriging*, CINDI2D: *compositional bayesian indicator kriging*, DK: *disjunctive kriging*. En negrita se destacan los mejores valores para una mejor gestión de la HGT_{As}: mayor probabilidad de éxito y menor probabilidad de clasificación falsa.

		SK	AUTOIK	SKLVM	Simplicial	BINDI2D	CINDI2D	DK
$p_u = 0.3$	Probabilidad falsa esperanza	-7%	11%	7%	6%	6%	3%	18%
	Probabilidad éxito zona segura	93%	89%	93%	94%	94%	97%	82%
	Probabilidad falsa alarma	7%	9%	19%	26%	26%	33%	25%
	Probabilidad éxito zona potencialmente tóxica	93%	91%	81%	74%	74%	67%	75%
$p_u = 0.5$	Probabilidad falsa esperanza	6%	9%	6%	6%	3%	3%	11%
	Probabilidad éxito zona segura	94%	91%	94%	94%	97%	97%	89%
	Probabilidad falsa alarma	30%	33%	30%	30%	30%	36%	17%
	Probabilidad éxito zona potencialmente tóxica	70%	67%	70%	70%	64%	33%	83%

El paso siguiente a este tipo de estudios sería el estudio de la posible asociación estadística entre la concentración de As en el agua y la distribución de patrones de salud. Este tipo de estudios podrían permitir contemplar la influencia de anomalías de HGT en la distribución geográfica de determinadas patologías, teniendo en cuenta que las aguas naturales juegan un papel básico en la transferencia de ETGPT desde el medio físico a la biosfera, y representa por tanto un nudo crítico de interrelación dinámica entre el medio ambiente y la salud humana.

AGRADECIMIENTOS

Los autores agradecen a la Obra Social de Caja de Ávila el apoyo a la investigación, al financiar el proyecto “Manantiales de la provincia de Ávila (2006-2007)” y a los revisores anónimos por los comentarios realizados.

REFERENCIAS

- Aragonés Sanz, N., Palacios Diez, M., Avello de Miguel, A., Gómez Rodríguez, P., Martínez Cortés, M., Rodríguez Bernabeu, M.J. 2001. Nivel de arsénico en abastecimientos de agua de consumo de origen subterráneo en la Comunidad de Madrid. *Revista Española de Salud Pública*, 75, 421-432.
- Barroso, J.L., Lillo, J., Sahún, B., Tenajas, J. 2002. Caracterización del contenido de arsénico en las aguas subterráneas de la zona comprendida entre el río Duero, el río Cega y el Sistema Central. In: *Presente y Futuro del agua subterránea en España y la Directiva Marco Europea*. Zaragoza, Spain, 77-84.
- Brus, D.J., Gruijter, J.J., Walvoort, D.J.J., de Vries, F., Bronswijk, J.J.B., Römkens, P.F.A.M., de Vries, W. 2002. Mapping the probability of exceeding critical thresholds for cadmium concentrations in soils in the Netherlands. *Journal of Environmental Quality*, 31, 1875-1884. doi:10.2134/jeq2002.1875
- Cattle, J.A., McBratney, A.B., Minasny, B. 2002. Kriging method evaluation for assessing the spatial distribution of urban soil lead contamination. *Journal of Environmental Quality*, 31, 1576-1588. doi:10.2134/jeq2002.1576

- Delgado, J., Medina, J., Vega, M., Carretero, C., Pardo, R. 2009. Los minerales de la arcilla y el arsénico en los acuíferos de la Tierra de Pinares. *Revista de la Sociedad Española de Mineralogía*, 11, 75-76.
- D'Or, D., Demougeot-Renard, H., Garcia, M. 2008. Geostatistics for contaminated sites and soils: some pending questions. *geoENV VI – Geostatistics for Environmental Applications*, 15, 409-420. doi:10.1007/978-1-4020-6448-7_34
- ESRI. 2011. *ArcGIS Desktop: Release 10*. Redlands, CA: Environmental Systems Research Institute.
- Falivene, O., Cabrera, L., Tolosana-Delgado, R., Sáez, A. 2010. Interpolation algorithm ranking using cross-validation and the role of smoothing effect. A coal zone example. *Computers & Geosciences*, 36(4), 512-519.
- García-Sánchez, A., Alvarez-Ayuso, E. 2003. Arsenic in soils waters its relation to geology mining activities, (Salamanca Province, Spain). *Journal of Geochemical Exploration* 80, 69-79. doi:10.1016/S0375-6742(03)00183-3
- García-Sánchez, A., Moyano, A., Mayorga, P. 2005. High arsenic in groundwater of central Spain. *Environmental Geology*, 47(6), 847-854. doi:10.1007/s00254-004-1216-8
- Giménez-Forcada, E., Smedley, P.L. 2014. Geological factors controlling occurrence and distribution of arsenic in groundwaters from the southern margin of the Duero Basin, Spain. *Environmental Geochemistry and Health*, 36(6), 1029-1047. doi:10.1007/s10653-014-9599-2
- Gómez, J.J., Lillo, F.J., Sahún, B. 2006. Naturally occurring arsenic in groundwater identification of the geochemical sources in the Duero Cenozoic Basin, Spain. *Environmental Geology*, 50, 1151-1170. doi:10.1007/s00254-006-0288-z
- Gómez Hernández, J.J. 1991. *Geoestadística, para el análisis de riesgos: Una introducción a la geoestadística no paramétrica*. Publicación Técnica 04/91, ENRESA.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, USA.
- Goovaerts, P., Avruskin, G., Meliker, J., Slotnick, M., Jacquez, G., Nriagu, J. 2005. Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. *Water Resources Research*, 41(7), W07013. doi:10.1029/2004WR003705
- Goovaerts, P. 2009. AUTO-IK: A 2D indicator kriging program for the automated non-parametric modeling of local uncertainty in earth sciences. *Computers & Geosciences*, 35(6), 1255-1270. doi:10.1016/j.cageo.2008.08.014
- Guardiola-Albert, C., Pardo-Igúzquiza, E. 2011. Compositional Bayesian indicator estimation. *Stochastic Environmental Research and Risk Assessment*, 25(6), 835-849. doi:10.1007/s00477-011-0455-y
- Isaaks, E.H., Srivastava, R.M. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, USA.
- Journel, A.G. 1983. Non-parametric estimation of spatial distributions. *Mathematical Geology*, 15(3), 445-468. doi:10.1007/BF01031292
- Journel, A., Kyriakidis, P.C., Mao, S. 2000. Correcting the smoothing effect of estimators: a spectral postprocessor. *Mathematical Geology*, 32(7), 787-813. doi:10.1023/A:1007544406740
- Juang, K.W., Chen, Y.S., Lee, D.Y. 2004. Using sequential indicator simulation to assess the uncertainty of delineating heavy-metal contaminated soils. *Environmental Pollution*, 127(2), 229-238. doi:10.1016/j.envpol.2003.07.001
- Kitanidis, P.K. 1991. Orthonormal Residuals in Geostatistics: Model Criticism and Parameter Estimation. *Mathematical Geology*, 23(5), 741-758. doi:10.1007/BF02082534
- Lark, R.M., Ferguson, R.B. 2004. Mapping risk of soil nutrient deficiency or excess by disjunctive and indicator kriging. *Geoderma*, 118(1-2), 39-53. doi:10.1016/S0016-7061(03)00168-X
- Mayorga, P., Moyano, A., Anawar, H.M., García-Sánchez, A. 2013. Temporal variation of arsenic and nitrate content in groundwater of the Duero River Basin (Spain). *Physics and Chemistry of the Earth*, 58-60, 22-27. doi:10.1016/j.pce.2013.04.001
- Olea, R., Pawlowsky, V. 1996. Compensating for estimation smoothing in kriging. *Mathematical Geology*, 28(4), 407-417. doi:10.1007/BF02083653

- Papritz, A., Dubois, J.P. 1999. Mapping heavy metals in soil by (non-)linear kriging: an empirical validation. In: *geoENV II: Geostatistics for Environmental Applications*. Quantitative Geology and Geostatistics, vol. 10. Gomez-Hernandez, J. et al. (eds.), Kluwer Academic Publishing, Dordrecht, 429-440.
- Remy, N., Boucher, A., Wu, J. 2009. *Applied Geostatistics with SGeMS: A User's Guide*. Cambridge University Press. doi:10.1017/CBO9781139150019
- Rey-Moral, C., Gómez Ortiz, D., Giménez-Forcada, E., López-Bahut, M.T. 2016. Modelización gravimétrica y aeromagnética en el SE de la Cuenca del Duero (provincias de Ávila y Segovia). Factores geológicos que controlan la distribución de As (Arsénico) y otros ETGPT (Elementos Traza Geogénicos Potencialmente Tóxicos) en las aguas subterráneas. *IX Congreso Geológico de España*. Huelva 2016.
- Rivoirard, J. 1994. *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*. Oxford Univ. Press, Oxford, UK.
- Rousseau, D. 1980. *Contrôle des previsions, II, Vérification des prévisions de l'occurrence d'un phénomène : Application aux prévisions de précipitations*, report. Étab. d'Études et Rech. de la Météorol./Météo France, Paris, France.
- Ryker, S.J. 2001. Mapping arsenic in ground water: A real need, but a hard problem. *Geotimes*, 46(11), 34-36.
- Sahún, B., Gómez Fernández, J.J., Lillo, J., Olmo, P.D. 2004. Arsénico en aguas subterráneas e interacción agua-roca: un ejemplo en la cuenca terciaria del Duero (Castilla y León, España). *Revista de la Sociedad Geológica de España*, 17(1-2), 137-155.
- Smedley, P.L., Kinniburgh, D.G. 2002 A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, 17(5), 517-568. doi:10.1016/S0883-2927(02)00018-5
- Tolosana-Delgado, R., Pawlowsky-Glahn, V., Egozcue, J.J., van der Boogaart, K.G. 2005. A compositional approach to indicator kriging. In: *2005 annual conference of the IAMG* (Cheng, Q., Bonham-Carter, G., eds.), Toronto, Canada: 651-656.
- Tolosana-Delgado, R., Pawlowsky-Glahn, V., Egozcue, J.J. 2008. Indicator kriging without order relation violations. *Mathematical Geosciences*, 40(3), 327-347. doi:10.1007/s11004-008-9146-8
- WHO. 2009. Chemicals Safety - Activity Report 2009. http://www.who.int/ipcs/about_ipcs/activity_report_2009.pdf. Last access: 24.10.2016.

