¿Quién ha sido?

Algunas notas sobre hidrogeología forense

J. Jaime Gómez-Hernández
Catedrático de Ingeniería Hidráulica de la ETS de Ingeniería de Caminos, Canales y Puertos de la Universitat Politècnica de València y responsable del Grupo de Hidrogeología del Instituto de Ingeniería del Agua y Medio Ambiente
Doctor Ingeniero de Caminos, Canales y Puertos

Resumen
La hidrogeología forense persigue generar información susceptible de ser usada en un juzgado, como, por ejemplo, identificar el lugar y el momento en el que un vertido ha sido introducido en un acuífero. Este es un problema que ha recibido la atención de la comunidad científica desde 1981 pero que, hoy por hoy, se ha resuelto en condiciones poco realistas con asunciones difíciles de justificar. Solo recientemente se han desarrollado técnicas basadas en los filtros de Kalman que dan una solución más realista, pero que todavía dista de poder ser usada ante un tribunal.

Palabras clave
aguas subterráneas, vertidos, simulación, optimización, filtro de Kalman de conjuntos, incertidumbre, errores de medición

Summary
Forensic hydrogeology aims to generate information that can be used in a court of law, such as identifying where and when a discharge has been introduced into an aquifer. This is a problem that has received the scientific community's attention since 1981 but has been solved under unrealistic conditions with assumptions that are difficult to justify. Only recently have techniques based on Kalman filters been developed that give a more realistic solution, but they are still far from being usable in court.

Keywords
groundwater, discharges, simulation, optimization, ensemble Kalman filter, uncertainty, measurement errors

Introducción

Estás en tu casa en el campo, en la que tienes un pozo de abastecimiento para cubrir las necesidades domésticas y quizá regar un pequeño huerto para autoconsumo, y notas un olor impropio en el agua del pozo que no habías notado nunca antes. Llevas una muestra a analizar a un laboratorio homologado y te informan que en el agua del pozo se han encontrado concentraciones muy por encima de los máximos admisibles de tetracloroetileno (TCE). Las concentraciones no disminuyen en el tiempo y te ves obligado a clausurar el pozo y buscar una nueva localización para construir otro que te proporcione el

mismo servicio que el abandonado y que no esté afectado por la contaminación por TCE. A la par, las preguntas que surgen son ¿quién ha sido?, ¿a quién puedo reclamar los daños causados por la contaminación del pozo? Ahí es donde entra en juego la hidrogeología forense, cuya definición más formal sería la rama de la hidrogeología que puede proporcionar información que podría usarse en un juzgado.

La hidrogeología forense

Las referencias a la hidrogeología forense son escasas, una búsqueda de literatura solo da dos coincidencias exactas en el título, una corresponde a un estudio realizado en el río Colorado con el objeto de identificar la fuente de contaminación por petróleo de un tramo de unos 200 metros de río junto al campo petrolífero del condado de Wharton, EE. UU. (Smyth, Dutton y Gutierrez, 2002) y otra a un artículo publicado en la Enciclopedia del Agua de Leonhart y Hargis (2005). No hay referencias a la hidrogeología forense en España o en español. Ampliando la búsqueda a la hidrología forense o a la geología forense, se encuentran algunas referencias más, aunque no demasiadas.

Con el aumento de la concienciación ambiental y la proliferación de legislación orientada a la protección del medio ambiente, la necesidad de una rama forense en las ciencias de la tierra, y concretamente en hidrogeología, parece necesaria. Cada vez son más las veces en las que los hidrogeólogos son llamados como peritos para dar su opinión experta en juicios donde se requiere evaluar afecciones inducidas por perforaciones ilegales, dirimir sobre la existencia de riesgos para las aguas subterráneas derivados de actuaciones en el entorno de un acuífero, o identificar dónde y cuándo un contaminante ha sido introducido en el subsuelo.

Lejos del entorno judicial, cualquier actividad forense puede describirse como la búsqueda, por métodos científicos, de las causas que han producido un cierto efecto. Y, en este sentido, uno de los objetivos de la hidrogeología forense es la búsqueda del agente que produjo la contaminación en un acuífero, cuando esta se detecta en uno o varios pozos. Es en la solución de este problema específico que el Grupo de Hidrogeología de la Universidad Politécnica de Valencia lleva trabajando desde el año 2016, cuando se publicó el trabajo de Xu y Gómez-Hernández, y sobre lo que versará el resto de este artículo.

Perspectiva histórica

El problema de la determinación de la fuente causante de una contaminación en un acuífero no es nuevo, como no lo son los intentos de resolverlo. Ya en 1981, Gorelick, en su tesis doctoral, presentó un método para identificar qué secciones de una conducción, por la que circulaban líquidos tóxicos y que estaba en contacto con un acuífero libre, estaban rotas y producían una inyección continua de contaminante en un acuífero usando las concentraciones de este observadas en una serie de pozos ubicados aguas abajo del vertido. La solución se basaba en un modelo de flujo de agua subterránea y de transporte de masa a partir del cual se obtenían unas funciones de influencia que servían para definir una función objetivo dependiente de las concentraciones en los segmentos de la conducción susceptibles de ser los responsables. Esa función objetivo, que predecía las concentraciones en los puntos de observación en función de las concentraciones en los puntos de inyección, se minimizaba hasta encontrar las concentraciones que deberían entrar en el acuífero desde la conducción resultantes en la mejor reproducción posible de las concentraciones medidas en los pozos de observación. En la figura 1 se muestra el caso sintético resuelto por Gorelick, en el que usó un modelo con muy pocas celdas, y en la que se ve la tubería con nueve

posibles secciones de las que solo dos (la III y la VIII) están rotas. Los pozos numerados S1 a S4 miden las concentraciones en el acuífero y a partir de esas mediciones, y suponiendo que todos los parámetros que describen el acuífero son perfectamente conocidos, Gorelick era capaz de identificar las dos secciones con rotura y las concentraciones de inyección en cada sección.

El planteamiento del problema y la posibilidad de resolverlo con modelos de simulación-optimización atrajo la atención de la comunidad investigadora que comenzó a estudiar el problema desde distintos ángulos y con distintas fórmulas de resolución, tal y como recogen Gómez-Hernández y Xu (2022). En este artículo, los autores muestran el crecimiento exponencial en artículos publicados sobre el tema, especialmente a partir de 1990 (ver figura 2), y revisan las publicaciones clave que supusieron un avance en la solución del problema. Los autores se quejan de que el progreso hacia una solución que pueda realmente servir para su aplicación en la práctica del hidrogeólogo forense no ha sido suficiente. Desde el trabajo de Gorelick, se han resuelto problemas de identificación de fuentes de contaminación con geometrías mucho más complejas, con discretizaciones mucho más densas, y usando algoritmos de simulación-optimización mucho más sofisticados, pero eso no ha supuesto que se hayan acercado a las condiciones que permitan que esas propuestas sean aplicables en la práctica, mucho menos que puedan llevarse ante un tribunal y usarse como evidencia incuestionable en un litigio. Son muy pocos los artículos que muestran aplicaciones a casos reales, y la mayor parte de ellos usan modelos de acuíferos donde parámetros, condiciones de contorno, y bombeos y recargas son conocidos y las únicas incógnitas se refieren a la localización espaciotemporal de la fuente y su intensidad.

Aproximación a una aplicación real

El primer artículo en el que se da un paso de gigante hacia una posible aplicación real de las técnicas de identificación de fuentes (puntuales) de contaminación en acuíferos es el de Xu y Gómez-Hernández (2018) donde, por primera vez, se propone un algoritmo que permite la identificación simultánea de la fuente y de la distribución espacial heterogénea de las conductividades hidráulicas. En los pocos artículos anteriores en los que se había abordado la identificación simultánea de fuente y conductividades hidráulicas, la distribución espacial de estas se limitaba a unas pocas zonas homogéneas, pero Xu y Gómez-Hernández plantean la necesidad de tratar la heterogeneidad de la conductividad hidráulica con la variabilidad observable en la naturaleza y por ello plantean tres ejercicios de identificación que se ven en la figura 3.

Esta figura muestra la geometría y distribución espacial de conductividades en dos acuíferos heterogéneos con características marcadamente diferentes, también muestra la discretización espacial, y la posición del punto de inyección. No muestra las condiciones de contorno, pero estas son tales que inducen un flujo de izquierda a derecha. Los tres casos abordados son: (S1) la identificación de un campo de conductividades hidráulicas heterogéneas con características gaussianas, junto con las características de la fuente ubicada en un punto con conductividad intermedia, (S2) ídem pero ahora el campo es no gaussiano con la presencia de un "canal" de conductividades altas embebido en una matriz de conductividades bajas y con la fuente situada en dicho "canal", y (S3) igual que el anterior pero con la fuente en la matriz.

Para la resolución del problema se dispusieron 25 puntos de observación, espaciados regularmente en el acuífero y se hizo una simulación previa usando las conductividades

hidráulicas de la figura 3 resultando en unas series temporales de evolución de la piezometría y de la concentración en los 25 puntos, que se usarían como los datos "reales" que se observarían en el campo caso que se produjera el vertido en los puntos indicados en dicha figura. Estas series de observación, junto con la geometría del sistema y condiciones de contorno es toda la información disponible para intentar resolver el problema: la identificación de la heterogeneidad de la conductividad hidráulica y los parámetros que definen la fuente, es decir, sus coordenadas x e y, el instante de inicio y la duración de la inyección (que se supone uniforme en el tiempo), y su concentración.
El método utilizado y que se describe más adelante es el filtro de Kalman de conjuntos (Evensen, 1994), una variante del filtro de Kalman (Kalman, 1960) que se ha demostrado evita las dificultades del filtro original para los casos en que se trabaje con sistemas que evolucionan en el tiempo de manera no lineal o cuando los parámetros que controlan esa evolución siguen una distribución estadística no gaussiana. El filtro de Kalman de conjuntos ya se había mostrado muy eficiente para la identificación de conductividades hidráulicas a partir de información piezométrica transitoria, y fueron Xu y Gómez-Hernández (2016) quienes descubrieron su potencial para la identificación de fuentes de contaminación. El éxito de ese primer trabajo llevo al intento, con éxito también, de identificar conjuntamente conductividades y fuente.

El filtro de Kalman

El filtro de Kalman se desarrolló con el objetivo de mejorar el guiado de misiles. Cuando se lanza un misil, existe una ecuación de estado que predice cuál va a ser la trayectoria. Una vez lanzado se realizan observaciones cada cierto tiempo que nos indican dónde está; generalmente, las observaciones no coinciden exactamente con las predicciones de la ecuación de estado, ya sea por errores en las mediciones o por errores en la ecuación; el filtro de Kalman lo que pretende es filtrar esos errores (de ahí su nombre) y proponer una posición del misil que se ubicará entre el valor predicho y el valor observado por el aparato de medida. Partiendo del valor reestimado de la posición del misil se recalcula la trayectoria con la ecuación de estado y tras cada nueva observación se volverá a reestimar la posición del misil en función de la discrepancia entre observación y predicción. Y así, sucesivamente.

El filtro de Kalman de conjuntos

El quid de la cuestión en el filtro de Kalman es el cálculo de la corrección, que matemáticamente se expresaría así

$$z_{\text{corregido}} = z_{\text{predicho}} + K\left(z_{\text{observado}} - z_{\text{predicho}}\right)$$

donde z es la variable de estado y K es lo que se conoce como la ganancia de Kalman que, entre otras cosas, depende de los errores de medida y del modelo, de manera que si los errores en el modelo tienden a cero, K tiende a cero y el valor corregido es igual al valor observado y si los errores de medida tienden a cero, K tiende a uno y el valor corregido es igual al valor observado. El cálculo de K no es trivial y este no es el lugar para derivar su expresión matemática, sirva decir que requiere de un modelo estadístico de los errores y es función de las funciones de covarianza de estos y de la variable de estado.
El filtro de Kalman se puede extender a la mejora de la predicción no solo de la variable de estado (la posición del misil en su formulación original) sino también los parámetros de la ecuación de estado que pueden ser los responsables de la desviación entre predicciones y observaciones (por ejemplo, la velocidad del misil). En ese caso, se calcula, simultáneamente con la corrección de la variable de estado, una corrección de los parámetros del modelo tras cada observación según la siguiente expresión

$$p_i = p_{i-1} + K'(z_{\text{observado}} - z_{\text{predicho}})$$

donde p_{i-1} es la estimación del parámetro tras la observación anterior y K' es una nueva ganancia de Kalman que ahora incorpora información de la ecuación de estado que relaciona el parámetro p con el estado del sistema z.

El filtro de Kalman de conjuntos surge de la dificultad de obtener esas ganancias, que Evensen (1994) propone calcular usando un conjunto de realizaciones que representen alternativas plausibles de la realidad (tanto de los parámetros como del estado del sistema) y que todas ellas se usan para predecir el estado del sistema en el tiempo y se van corrigiendo tras cada conjunto de observaciones es hecho usando las ecuaciones antes descritas. El disponer de una colección de realizaciones hace que el cálculo de las covarianzas antes mencionadas, que en la formulación original se hacía mediante funciones analíticas aproximadas, ahora se pueda hacer con más precisión a partir de la covarianza experimental del conjunto de realizaciones. De nuevo, no es este el lugar donde desarrollar estos cálculos.

Aplicación del filtro de Kalman de conjuntos a la identificación conjunta de la heterogeneidad del medio y los parámetros de la fuente de contaminación

Retomando los casos sintéticos de la figura 3, partiendo del desconocimiento total de la distribución espacial de las conductividades y asumiendo que la fuente puede estar dentro de los rectángulos que se ven en la figura, el filtro de Kalman de conjuntos aplicado en sesenta instantes de tiempo (es decir, las ecuaciones anteriores fueron aplicadas sesenta veces) tras sesenta medidas de piezometría y concentración en los 25 piezómetros de observación es capaz de producir una estimación de las conductividades junto con una varianza de la estimación como se observa en la figura 4. Los resultados son espectaculares, especialmente considerando que, de partida, no se tenía información sobre la distribución de las conductividades. La figura 5 muestra la evolución de los parámetros que describen la fuente de contaminación estimados en el conjunto de realizaciones usando diagramas de caja, donde se observa que tras la última observación (paso de tiempo 60) los valores estimados (líneas rojas) son próximos a los reales (línea negra discontinua) y, en general, la incertidumbre, medida por el ancho de los bigotes, disminuye con el tiempo.

Conclusiones

La hidrogeología forense persigue encontrar las causas de un problema en un contexto hidrogeológico como es el de identificar dónde y cuándo se ha producido un vertido en un acuífero. Para que el testimonio de un hidrogeólogo pueda usarse como juicio experto ante un tribunal, es necesario que las técnicas utilizadas hayan sido verificadas y puedan usarse en casos reales. Desafortunadamente, la identificación de una fuente de contaminación en un acuífero sigue siendo un tema de investigación que todavía no está maduro para su aplicación en la práctica forense.

Referencias

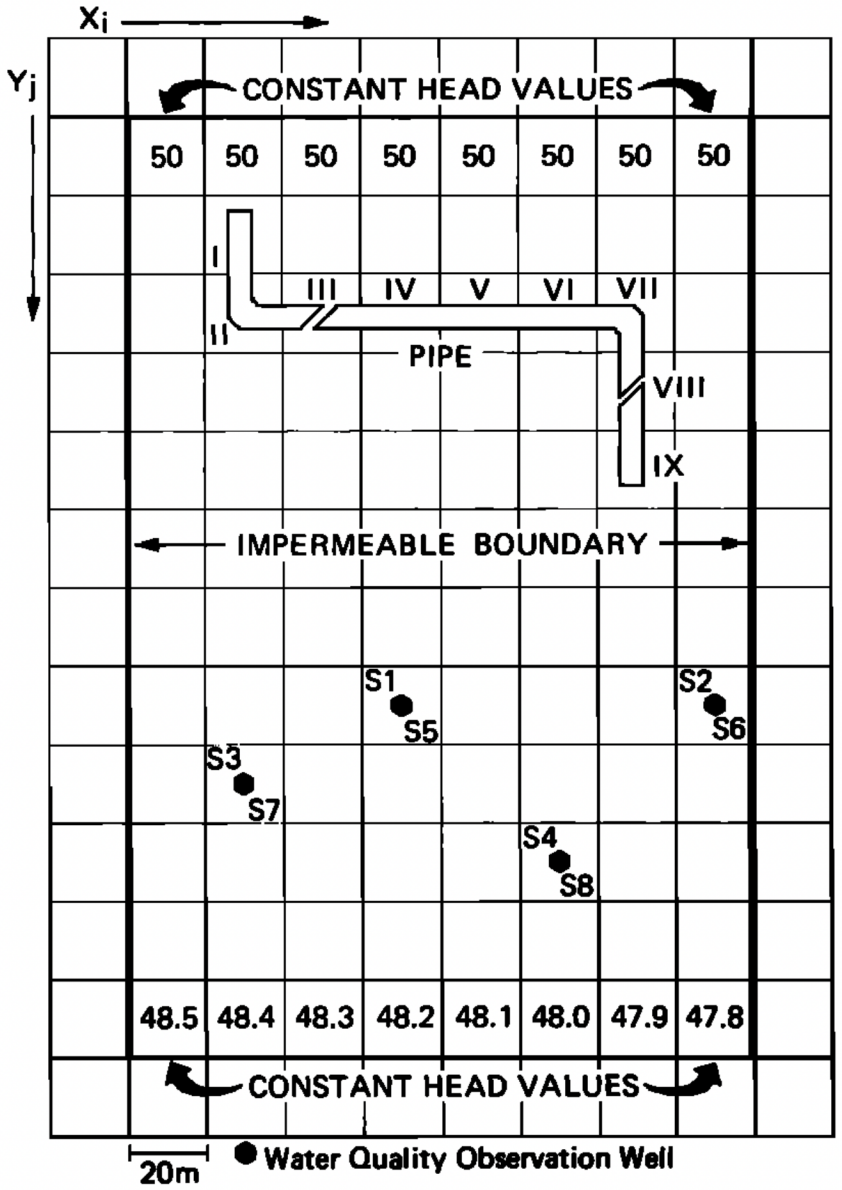
Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, *99*(C5), 10143-10162.

Gómez-Hernández, J. J., & Xu, T. (2022). Contaminant Source Identification in Aquifers: A Critical View. Mathematical Geosciences, *54*,437-458.

Gorelick S. M. (1981) Numerical Management Models of Groundwater Pollution. Ph.D., Stanford University.
Gorelick, S. M., Evans, B., & Remson, I. (1983). Identifying sources of groundwater pollution: An optimization approach. *Water Resources Research*, *19*(3), 779-790.
Leonhart, L. S., & Hargis, D. R. (2005). Forensic Hydrogeology. *Water Encyclopedia*, *3*, 168-170.
Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Trans. ASME, Series D, J. Basic Engineering, *82*(March), 35-45.
Smyth, R. C., Dutton, A. R., & Gutierrez, R. (2002). Forensic Hydrogeology Applied to a Half-Century-Old Crude Oil Seep, Colorado River, Wharton County, Texas.
Xu, T., & Gómez-Hernández, J. J. (2016). Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. *Water Resources Research*, *52*(8), 6587-6595.
Xu, T., & Gómez-Hernández, J. J. (2018). Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. Advances in Water Resources, 112, 106-123.

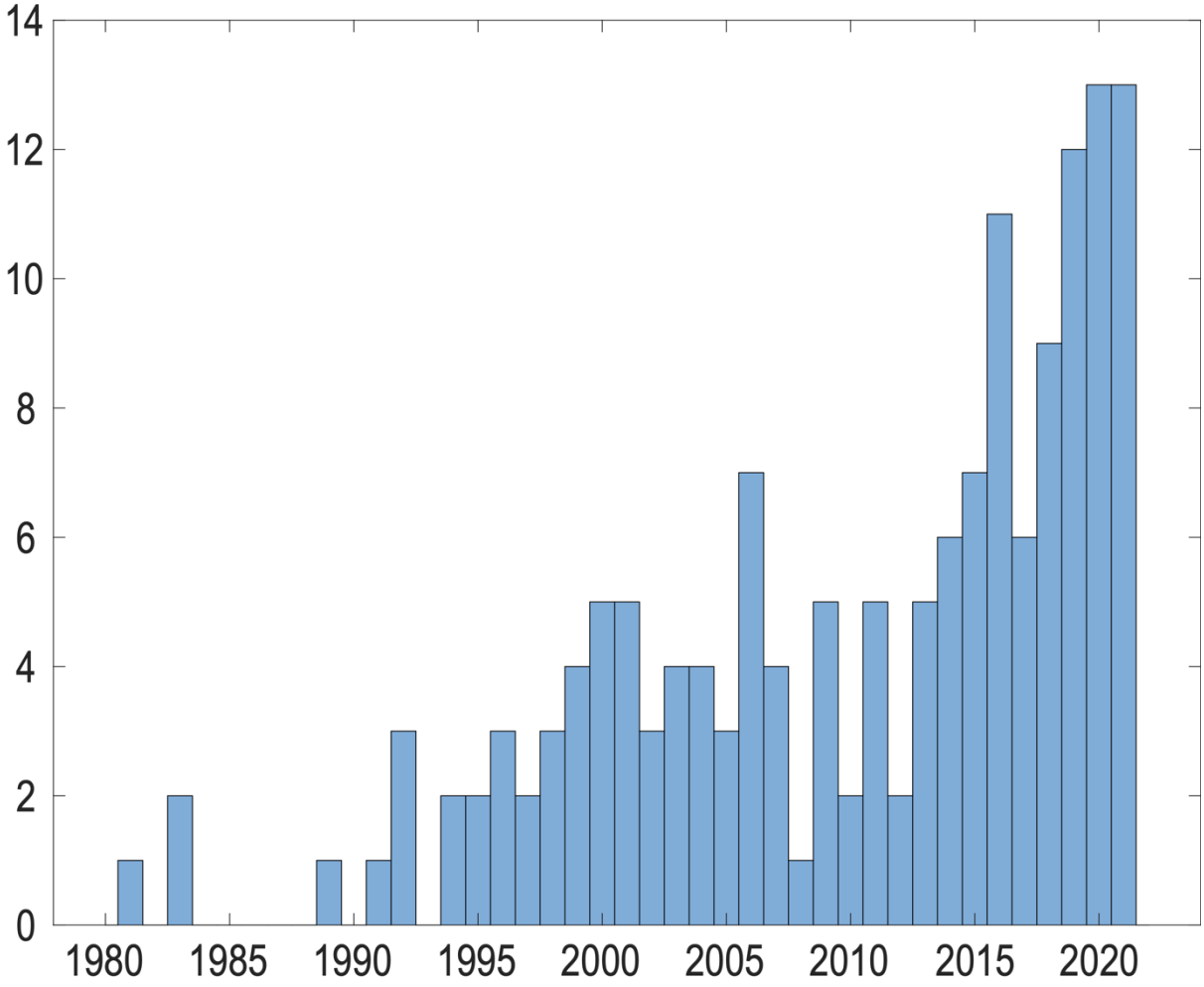
Figuras

Archivo Figura 1: Gorelickcase.png



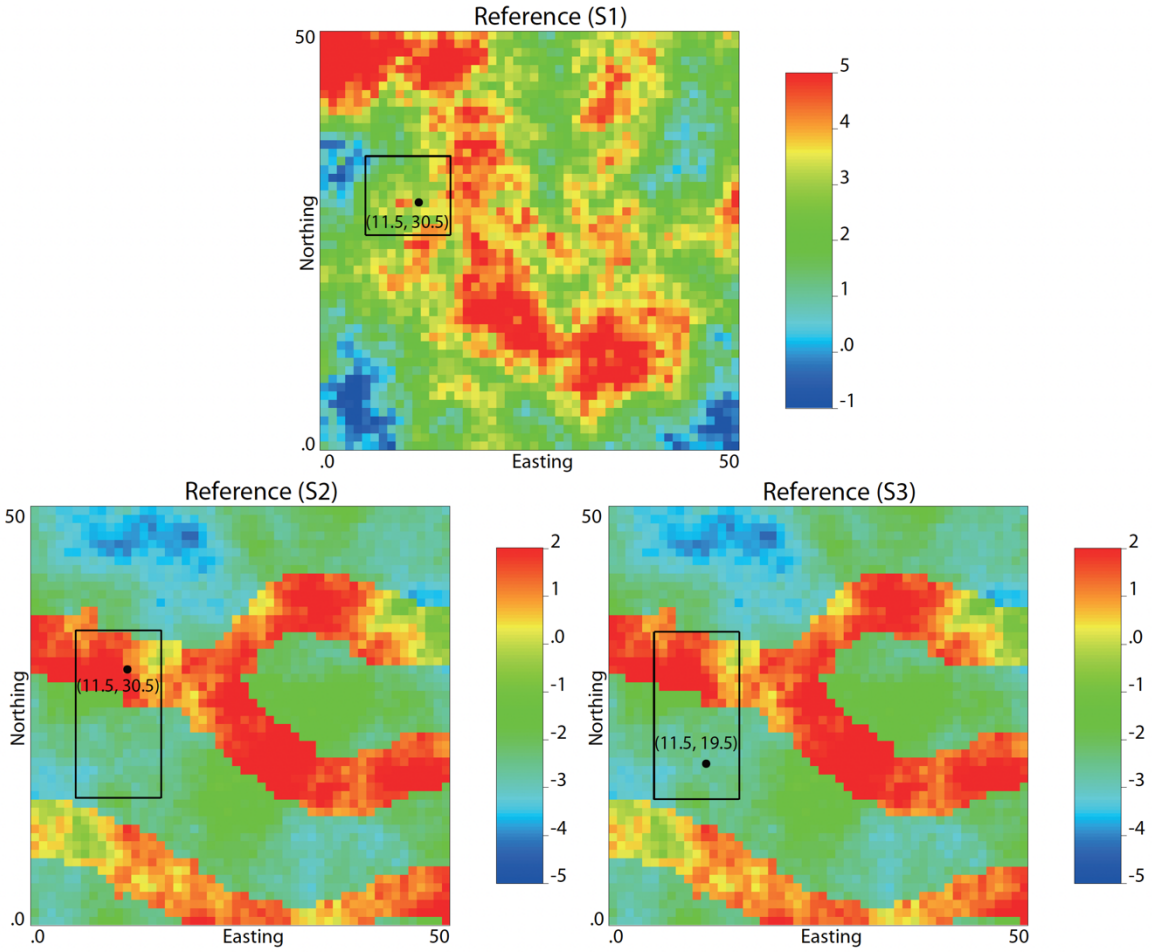
Pie Figura 1. Caso hipotético resuelto por Gorelick et al. (1983). Figura tomada, con permiso, del citado artículo, copyright American Geophysical Union.

Archivo Figura 2: HistogramNumberOfPapers.png



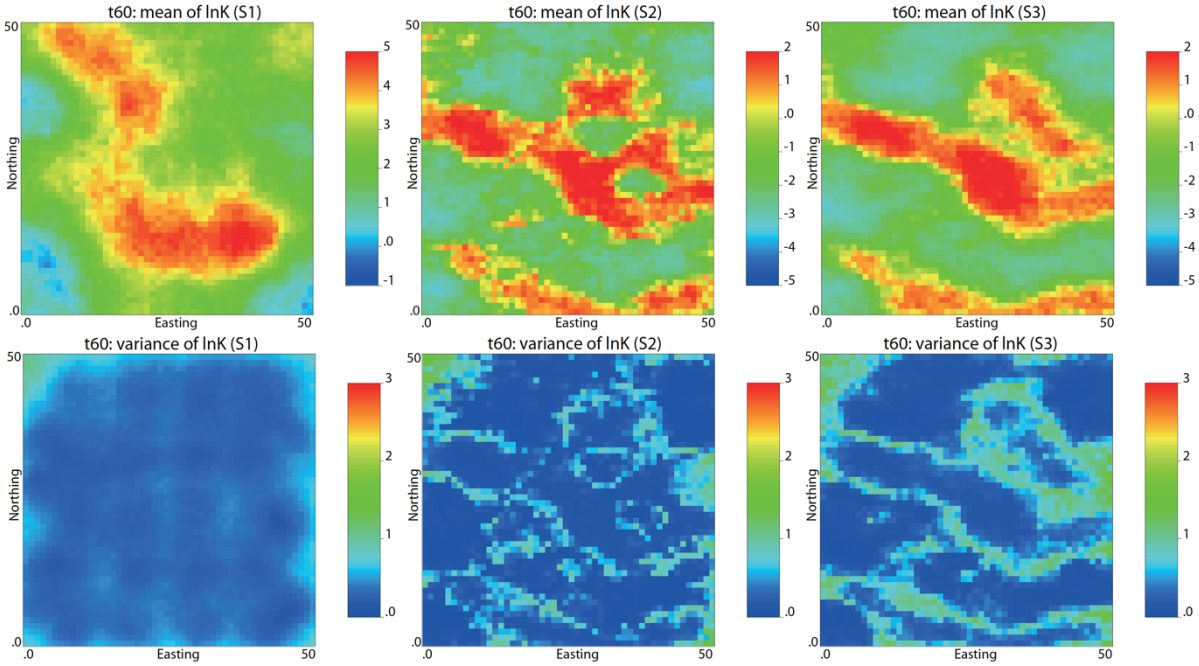
Pie Figura 2. Número de artículos publicados sobre identificación de fuentes de contaminantes en acuíferos. Figura tomada de Gómez-Hernández y Xu (2022) sujeta a la licencia Creative Commons Reconocimiento Internacional 4.0 (CC-BY-4.0).

Archivo Figura 3: Fig1XuG-H.png



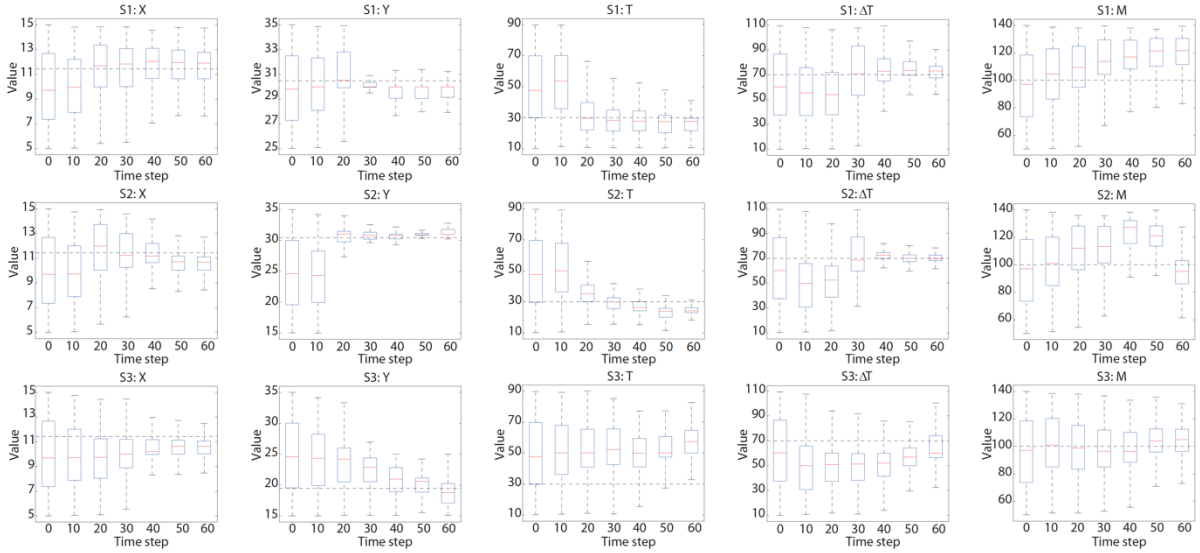
Pie Figura 3. Casos estudiados por Xu y Gómez-Hernández (2018). Distribución espacial del logaritmo de las conductividades hidráulicas y posición del punto de inyección. Figura tomada, con permiso, del citado artículo, copyright Elsevier.

Archivo Figura 4: Fig6XuG-H.png



Pie Figura 4. Fila superior: estimación de las conductividades para los tres casos de estudio de Xu y Gómez-Hernández (2018) en la fila superior. Fila inferior: varianzas de la estimación. Figura tomada, con permiso, del citado artículo, copyright Elsevier.

Archivo Figura 5: Fig11XuG-H.png



Pie Figura 5. Evolución en el tiempo de los parámetros estimados que identifican la fuente de contaminación: coordenadas X and Y, tiempo de inicio del vertido T, duración del vertido ΔT y concentración de la fuente M. Se muestran, en los pasos de tiempo indicados, diagramas de caja de cada parámetro, donde la línea roja sería el valor estimado, la caja indica el intervalo de confianza del 50% y los bigotes el intervalo de confianza del 90%. Se persigue que la línea roja coincida con la línea discontinua que es el parámetro "real" de la referencia.