

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS GEOLÓGICAS



TESIS DOCTORAL

**Interacción Datos-Modelo Matemático en el acuífero
Almonte-Marismas: un nuevo reto**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Nuria Naranjo Fernández

Directoras

**Carolina Guardiola Albert
Esperanza Montero González**

Madrid

© Nuria Naranjo Fernández, 2020



UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Geológicas

TESIS DOCTORAL

**INTERACCIÓN DATOS - MODELO MATEMÁTICO EN EL
ACUÍFERO ALMONTE-MARISMAS: UN NUEVO RETO**

NURIA NARANJO FERNÁNDEZ

Dirigida por:

Dra. Carolina Guardiola Albert
Dra. Esperanza Montero González

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS GEOLÓGICAS



TESIS DOCTORAL

**Interacción Datos-Modelo Matemático en el
acuífero Almonte-Marismas: un nuevo reto**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Nuria Naranjo Fernández

DIRECTORAS

Carolina Guardiola Albert

Esperanza Montero González

Madrid, 2020

Agradecimientos

En la realización de esta tesis doctoral he contado con el apoyo de muchas personas, sin las cuales, habría sido un camino más complicado de recorrer. Me gustaría agradecer infinitamente...

Al proyecto “El impacto del cambio climático en los recursos de agua subterránea”, CLIGRO (MINECO, CGL2016-77473-C3-1-R), del Plan Estatal de Investigación Científica y Técnica de Innovación (<http://www.igme.es/CLIGRO/>) a través del cual se ha podido realizar esta tesis doctoral.

A mis tutoras de tesis, Carolina y Esperanza, por creer en mí, por su paciencia, dedicación y apoyo. Por numerosas horas de su tiempo dedicadas a una tesis de sobresaltos, novedades, alegrías y alguna que otra pena. Por aceptar acompañarme en este reto.

A Héctor, que ha sido el padrino de esta tesis en las sombras, inspirándome con sus ideas y su forma de unir hilos, atando cabos y resolviendo miles de dudas. Por todo el trabajo que juntos hemos sacado adelante y las ideas felices que nos quedan que debatir.

A los tres pilares principales con los que cuento en mi vida, mi madre y mis hermanas, Sarín, Marta y Sofía. Especialmente a mi madre, inspiradora de mi carrera académica e investigadora, luchadora imparable y apoyo incondicional, a la que dedico cada logro de esta tesis.

A mis compañeros de doctorado, Gabriel, Toni, Carlota, Carmen y Bea que han compartido miles de desayunos, celebraciones de aceptación de artículos, tensiones, días de terraza y mucho, mucho cariño.

A mis amigos Isabel y Sergio, que me han aguantado reír y llorar durante los altibajos de esta tesis, a los que se unen los de la vida en general. Que siempre han estado, están y estarán. Ellos son “los de siempre”, esos que te acompañan para toda la vida.

A Diego, por ser siempre una sorpresa, por sacarme una sonrisa cuando ya no me quedan ganas, por ser único en tu especie y por 11 años de montañas rusas. Porque el que cayó como una bomba en mi vida, fuiste tu.

A mis antidotos de esta tesis: viajar y bailar, que me han dado energía e inspiración, siendo mi vía de escape y la apertura de la válvula de presión. Ambas cosas me han traído muchas personas nuevas a mi vida y también, nuevas experiencias. He viajado a lugares maravillosos, vivido situaciones variopintas y visto otras formas de vivir que me han hecho valorar, de otra forma, las decisiones de mi vida. No hay nada mejor que, día a día, agrandar tu caja.

A mi Sensual Team de bachata, Julián, Esther y Holly, mis duendecillos con los que he aprendido a compartir, a emocionarme y sacar lo mejor de mí. Por todos esos jueves de bachata online y por la de pistas que nos quedan por estrenar: “Bailemos, porque la vida es más bella bailando”.

A Ana Fernández Ayuso, por nuestras charlas infinitas sobre el doctorado, por nuestras marismas, congresos, póster y deadlines. Por una nueva amistad marismeña de las jóvenes investigadoras de Doñana.

A todo el personal del IGME, tanto de Alenza como de Sevilla, por la recogida de datos durante años, el mantenimiento de las redes de control y su disposición siempre a colaborar y responder a mis dudas, especialmente a Fernando Ruiz.

A todas aquellas personas que han pasado por mi vida durante este tiempo, dejando alguna semilla en mi camino, que ha florecido como aprendizaje.

¡Gracias a todos!

Resumen: Interacción Datos-Modelo Matemático en el acuífero Almonte-Marismas: un nuevo reto

El acuífero Almonte-Marismas alberga el Espacio Natural de Doñana y su entorno, en el cual se sitúa uno de los humedales más importantes de Europa. Tanto las marismas como las lagunas en el entorno de Doñana poseen un elevado valor ecológico que depende estrechamente del buen estado del acuífero. Actualmente y debido al desarrollo socioeconómico de la zona desde los años 60, el acuífero se encuentra sometido a una elevada presión antrópica procedente de la agricultura, el turismo estacional y la industrial; acusada, además, por los efectos del cambio global.

Una de las herramientas para estudiar el estado cuantitativo del acuífero y su gestión son los modelos matemáticos. El desarrollo de los modelos matemáticos aplicados a la Hidrogeología ha aumentado exponencialmente desde los años 80 hasta la actualidad. No obstante, uno de los factores limitantes de sus resultados es la calidad de los datos de entrada. Con los avances actuales en tecnología, la cantidad y calidad de las bases de datos se han incrementado notablemente, lo cual genera el reto de homogeneizar los datos recientes con los más antiguos. En el caso concreto del acuífero Almonte-Marismas, el modelo matemático de flujo utilizado tradicionalmente ha quedado obsoleto ya que no se han incorporado mejoras del conocimiento disponibles, y se desconoce el porcentaje del volumen hídrico extraído de manera irregular. Uno de los primeros pasos para la mejora del modelo requiere la explotación de los datos de tres de las variables más controladas y estudiadas en los últimos 40 años: la geología, la precipitación y la piezometría.

La estadística es hoy en día una herramienta de extendido uso en Hidrogeología para solventar el reto de alcanzar una buena calidad de datos de entrada. En esta tesis han sido aplicadas las técnicas de la simulación secuencial indicadora, krigeado espacio-temporal, y técnicas de agrupamiento estadístico para el tratamiento de grandes bases de datos.

El objetivo principal de esta tesis doctoral es el estudio de la interacción entre los datos de entrada y el modelo matemático de flujo en el acuífero Almonte-Marismas. Para ello

se han alcanzado tres objetivos parciales: se ha introducido un modelo geológico 3D en el modelo hidrogeológico de flujo y se ha evaluado la mejora sobre los resultados; se ha estudiado la variabilidad espacio-temporal de la precipitación y cómo afecta sobre el cálculo de la recarga; y, por último, se ha desarrollado una metodología para agrupar comportamientos piezométricos que faciliten el estudio de la respuesta del acuífero a la explotación antrópica.

En primer lugar, en esta tesis doctoral se ha profundizado y mejorado el conocimiento geológico del acuífero. El Instituto Geológico y Minero de España (IGME) disponía de una base de datos de 16.000 m de sondeos geológicos testificados, que nunca se había incluido al completo como variable de entrada en los modelos matemáticos de flujo existentes. Para la consecución de este objetivo, se transformaron y clasificaron las testificaciones originales a unidades hidrogeológicas, obteniendo seis hidrofacies. Seguidamente, se calcularon y modelaron los variogramas indicadores experimentales para obtener una caracterización cualitativa de la distribución espacial de las seis hidrofacies. A partir de esos variogramas se generó una simulación secuencial indicadora (SIS). Finalmente, la simulación obtenida fue introducida en el modelo matemático estacionario desarrollado con MODFLOW a través de la interfaz ModelMuse. Para cuantificar la mejora que supone la introducción de dicha simulación en el modelo hidrogeológico se comparó con el modelo anterior que está construido con zonas homogéneas de permeabilidad. En concreto, se compararon los balances hídricos y mapas piezométricos obtenidos. Variaciones de hasta un 30% entre ambos modelos corroboraron la importancia del tratamiento de los datos geológicos de entrada.

Como segundo reto de esta tesis, se ha realizado un estudio de la variabilidad espacio-temporal de la precipitación y los efectos que, a su vez, suponen en el cálculo de la recarga del acuífero. Para ello, se creó una base de datos con información recogida en 112 estaciones pluviométricas entre 1975 y 2016. Se aplicó un algoritmo de krigeado espacio-temporal y se obtuvieron resultados a diferentes resoluciones espaciales y temporales. Con el análisis de los resultados se determinó que utilizando el cálculo de la recarga mensualmente se puede sobreestimar la lluvia máxima entre un 21 y un 91%, lo cual genera una propagación del error sobre el cálculo de la recarga. Se concluyó que

cuando se trabaja en la gestión de los recursos hídricos, donde es común realizar cálculos mensuales, la escala espacio-temporal utilizada para la precipitación y la recarga puede ser crítica en los resultados obtenidos, por lo que el uso de datos diarios es imprescindible.

Por último, se han analizado las series temporales disponibles de 160 piezómetros. Se seleccionaron los 24 piezómetros más completos, con los que se realizó el trabajo de agrupamiento estadístico. La información estudiada de los niveles piezométricos corresponde a 41 años hidrológicos (1975 – 2016), durante los cuales se han producido efectos antrópicos importantes debido al incremento de la explotación del acuífero para regadío. Se realizó una primera agrupación visual como base para la familiarización con los datos. Posteriormente, se aplicaron dos métodos de agrupamiento, k-medias y un agrupamiento de series temporales (TSC). Se comprobó que el número de grupos está relacionado con las diversas afecciones producidas por la explotación de las aguas subterráneas. TSC permitió distinguir áreas con elevadas alteraciones en la hidrodinámica local y subrayar cuantitativamente los descensos de los niveles piezométricos en el periodo estudiado. Además, los resultados de la agrupación de piezómetros se compararon con las precipitaciones para determinar si su respuesta está relacionada o no con la recarga, y en qué medida. Se obtuvo que, la piezometría y la precipitación, tienen menor relación en las zonas en las que la explotación agrícola de los recursos hídricos es elevada.

Abstract: Data-Mathematical Model interaction of Almonte-Marismas aquifer: a new challenge

The Almonte-Marismas aquifer host the Doñana Natural Park and its surroundings, and is one of the most important wetlands in Europe. Both the marshland and the temporary ponds in the Doñana environment have a high ecological value that closely depends on the good condition of the aquifer. Currently and due to the socioeconomic development of the area since the 1960s, there is a high anthropic pressure derived from agriculture, seasonal and industrial tourism; worsening the effects of global change.

The development of mathematical models applied to hydrogeology has increased exponentially from the 1980s to the present day. However, one of the limiting factors in the results is the quality of the input data. The quantity and quality of databases have increased significantly due to the current advances in technology, which has created the challenge of homogenizing recent data with older data. In the specific case of the Almonte-Marismas aquifer, the mathematical flow model has become obsolete since no available knowledge improvements have been incorporated and a percentage of the irregularly extracted water volume is unknown. One of the first steps to improve the model requires the exploitation of data from three of the variables that have been most controlled and studied in the last 40 years: geology, precipitation and piezometry.

Statistics is one of the widely used tools in hydrogeology today and solves the challenge of achieving good quality input data. Sequential indicator simulation, spatio-temporal kriging, and clustering have been applied in this thesis for the treatment of long databases.

The main objective of this Doctoral Thesis was to study the interaction between the input data and the mathematical flow model in the Almonte-Marismas aquifer. To reach this aim, three partial objectives were proposed: a new 3D geological model was introduced in the hydrogeological flow model and the improvement on the results of the water balance was evaluated, effects of precipitation spatio-temporal variability on the recharge calculation, and finally, a methodology was developed to group

piezometric behaviors that facilitate the study of the response of the aquifer to anthropic exploitation.

Firstly, in this Doctoral Thesis, the geological knowledge of the aquifer has been deeply studied and improved. The Spanish Geological and Mining Institute (IGME) had a database of 16,000 m of testified geological surveys, which had never been fully included as an input variable in existing mathematical flow models. To achieve this objective, the original testifications were transformed and classified to hydrogeological units, obtaining six hydrofacies. After that, the experimental indicator variograms were calculated and modeled to obtain a qualitative characterization of the spatial distribution of the six hydrofacies. These variograms allowed to generate an indicator sequential simulation (SIS). Finally, the obtained simulation was introduced in the stationary mathematical model developed with MODFLOW through the ModelMuse interface. In order to quantify the improvement that the introduction of such simulation in the hydrogeological model, it was compared with the previous model that is built with homogeneous permeability zones. Specifically, the water balances and piezometric maps obtained were obtained. Variations of up to 30% corroborated the importance of the treatment of the input data.

The second target of this thesis was the study of the impact of precipitation spatio-temporal variability on the recharge calculation. A database was created with information collected at 112 rain gauges between 1975 and 2016. A spatio-temporal kriging algorithm was applied, and the results were obtained at different spatio-temporal resolutions. With the analysis of the results, it was determined that computing monthly recharge, the maximum rainfall can be overestimated between 21 and 91%, which generates an error propagation regarding the recharge calculation. It was concluded that when working at water resources management, where monthly calculations are common, the spatio-temporal scale used for precipitation and recharge can be critical in the results, recommending the use of daily data.

Finally, the available time series of 160 piezometers were analysed. A group of the 24 most complete piezometers, with which the statistical cluster work was carried out, was selected. The studied piezometric levels correspond to 41 hydrological years (1975-2016), during which significant anthropic effects have occurred due to the increased

exploitation of the aquifer for irrigation. A first visual classification was performed. Subsequently, two clustering methods were applied, k-means and a time series cluster (TSC). It was found that the number of groups is related to the various conditions, caused by the exploitation of groundwater. TSC made possible to distinguish areas with high alterations in local hydrodynamics and quantitatively underline the drops in piezometric levels in the period studied. Furthermore, results were compared with rainfall, to determine whether and to what extent group responses are related to recharge. It was found that piezometry and precipitation have a lower relationship in areas where the agricultural exploitation of water resources is high.

ÍNDICE

<u>RESUMEN: INTERACCIÓN DATOS-MODELO MATEMÁTICO EN EL ACUÍFERO ALMONTE-MARISMAS: UN NUEVO RETO</u>	<u>I</u>
<u>ABSTRACT: DATA-MATHEMATICAL MODEL INTERACTION OF ALMONTE-MARISMAS AQUIFER: A NEW CHALLENGE</u>	<u>IV</u>
<u>RELACIÓN DE PUBLICACIONES.....</u>	<u>1</u>
<u>CAPÍTULO 1: INTRODUCCIÓN</u>	<u>2</u>
1.1. MODELOS MATEMÁTICOS DE FLUJO DE AGUAS SUBTERRÁNEAS	4
1.1.1. CONCEPTO DE MODELO.....	4
1.1.2. MODELOS HIDROGEOLÓGICOS DE FLUJO	6
1.1.3. DATOS DE ENTRADA	7
1.1.4. DATOS DE SALIDA.....	9
1.1.5. APLICACIÓN DE LOS MODELOS MATEMÁTICOS DE FLUJO.....	10
1.2. TÉCNICAS ESTADÍSTICAS APLICADAS A MEJORAR LOS DATOS DE ENTRADA DE LOS MODELOS MATEMÁTICOS DE FLUJO.....	11
1.2.1. CARACTERIZACIÓN DE LA HETEROGENEIDAD GEOLÓGICA EN ACUÍFEROS REGIONALES	11
1.2.2. CARACTERIZACIÓN ESPACIO-TEMPORAL DE LA PRECIPITACIÓN Y LA RECARGA	12
1.2.3. USO DE TÉCNICAS DE AGRUPAMIENTO ESTADÍSTICO PARA EL ESTUDIO DE LA PIEZOMETRÍA.....	13
1.3. ZONA DE ESTUDIO	15
1.3.1. DESCRIPCIÓN GENERAL	15
1.3.2. IMPORTANCIA ECOLÓGICA DEL ÁREA DE ESTUDIO.....	19
1.3.3. PROBLEMÁTICA MEDIOAMBIENTAL Y SOCIOECONÓMICA	21
1.4. MODELO MATEMÁTICO DEL ACUÍFERO ALMONTE-MARISMAS	24
1.4.1. DESCRIPCIÓN GENERAL	24
1.4.2. CONDICIONES DE CONTORNO	25
1.4.3. PARÁMETROS HIDRÁULICOS	28
1.4.4. CALIBRACIÓN	28
<u>CAPÍTULO 2: OBJETIVOS.....</u>	<u>30</u>
2.1. OBJETIVO PRINCIPAL	30
2.2. OBJETIVOS PARCIALES.....	30
<u>CAPÍTULO 3. RESUMEN METODOLÓGICO</u>	<u>32</u>
<u>CAPÍTULO 4: ARTÍCULOS</u>	<u>34</u>
ARTÍCULO I	35
ARTÍCULO II	61
ARTÍCULO III.....	95
<u>CAPÍTULO 5: DISCUSIÓN GENERAL</u>	<u>125</u>

5.1. NUEVAS METODOLOGÍAS DE TRABAJO APLICADAS A LOS DATOS DE ENTRADA PARA LA MEJORA DEL MODELO MATEMÁTICO DEL ACUÍFERO ALMONTE-MARISMAS	125
5.1.1. SIMULACIÓN ESTOCÁSTICA CONDICIONADA PARA LA REALIZACIÓN DE UN MODELO DE HIDROFACIES EN 3D	125
5.1.2. INTERPOLACIÓN ESPACIO-TEMPORAL DE LA PRECIPITACIÓN Y SU INFLUENCIA SOBRE LA RECARGA	126
5.1.3. AGRUPAMIENTO ESTADÍSTICO DE SERIES TEMPORALES DE PIEZOMETRÍA	128
5.2. APORTACIONES A LA INVESTIGACIÓN SOBRE MODELOS MATEMÁTICOS Y SUS CONSECUENCIAS PARA LA MEJORA DE LA GESTIÓN DE LOS RECURSOS HÍDRICOS QUE SE APOYAN EN ELLOS	129
<u>CAPÍTULO 6: CONCLUSIONES Y LÍNEAS DE TRABAJO FUTURAS</u>	<u>132</u>
<u>REFERENCIAS.....</u>	<u>135</u>

Relación de figuras de la memoria de Tesis Doctoral

Capítulo 1

- Figura 1. Esquema de los elementos que conforman un modelo matemático de flujo de aguas subterráneas.
- Figura 2. Modelo digital de elevaciones y localización geográfica del acuífero Almonte-Marismas.
- Figura 3. Mapa geológico simplificado del acuífero Almonte-Marismas.
- Figura 4. Corte geológico del acuífero Almonte-Marismas dirección NO-SE cruzando por Almonte (Modificado de Martín-Machuca et al. 1992). Nivel piezométrico medio medido en 2016.
- Figura 5. Izquierda: capa superior del modelo. Derecha: capa inferior del modelo. Corte longitudinal (línea verde) representado debajo de la figura central y corte transversal (línea azul) representado a la derecha de la figura central.
- Figura 6. Zonificación espacial de la recarga.
- Figura 7. Izquierda: Situación de la condición de contorno de DRAIN en Cuesta Maneli (amarillo), la Vera-Retuerta (negro), borde norte de la marisma (verde) y marisma (rosa fuxia). Derecha: Situación de la condición de contorno STREAM.
- Figura 8. Izquierda: Distribución de espacial de las zonas de evapotranspiración freática del estado estacionario. Derecha: Localización de las celdas de potencial constante (CHB).
- Figura 9. Arriba: permeabilidad de la capa superior. Abajo: permeabilidad de la capa inferior.

Capítulo 3

- Figura 10. Esquema metodológico seguido en la presente tesis doctoral.

Capítulo 5

- Figura 11. Esquema de flujo de las ventajas que proporciona la utilización de datos de entrada mejorados en el modelo matemático de flujo.
- Figura 12. Esquema de trabajo en modelización hidrogeológica.

Relación de tablas de la memoria de Tesis Doctoral

Capítulo 1

- Tabla 1a. Recarga anual, recurso disponible y extracciones de las nuevas masas de agua subterránea (extraída de CHG 2019b).
- Tabla 1b. Recarga anual, recurso disponible y extracciones de las nuevas masas de agua subterránea (extraída de CHG 2019b).
- Tabla 2. Evaluación del estado cuantitativo y químico de las nuevas masas de agua subterránea (extraída de CHG 2019b).

Relación de publicaciones

- I. **Naranjo-Fernández, N.**, Guardiola-Albert, C., Montero-González, E. (2019) Applying 3D stochastic simulation to improve the hydrogeological modeling of heterogeneous aquifers. Case of Doñana (SW Spain). *Water*, 11, 39, doi: 10.3390/w11010039.
- II. **Naranjo-Fernández, N.**, Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Rodríguez-Rodríguez, M., Fernández-Ayuso, A., Ruiz-Bermudo, F., Montero-González, E. (2020) Relevance of spatio-temporal rainfall variability regarding groundwater management challenges under global change. Case study in Doñana (SW Spain). *Stochastic Environmental Research and Risk Assessment*, doi: 10.1007/s00477-020-01771-7.
- III. **Naranjo-Fernández, N.**, Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Montero-González, E. (2020) Clustering groundwater level time series of the exploited Almonte-Marismas aquifer in Southwest Spain. *Water*, 12, 1063, doi: 10.3390/w12041063.

Capítulo 1: Introducción

Desde hace décadas, la necesidad de gestionar las aguas subterráneas que alimentan los humedales del Parque Nacional de Doñana hizo necesario el uso de modelos matemáticos para comprender su funcionamiento, estimar sus recursos, simular los efectos de su uso y explotación, y más recientemente, estimar la influencia que el cambio global pueda tener sobre las mismas (IGME, 1976; IGME, 1982; IGME, 1994; IGME, 1999; IGME, 2005; IGME, 2009; IGME, 2011; Guardiola-Albert et al., 2016; Kohfahl et al., 2019). Del sistema acuífero Almonte-Marismas, depende estrechamente el buen estado de sus humedales, tanto en concepto de cantidad como de calidad del agua (Serrano & Serrano, 1996; Muñoz-Reinoso, 2001; Custodio et al., 2009; Manzano et al., 2009; Green et al., 2017). Por este motivo, toda mejora en el modelo matemático que represente de una forma más fiel la naturaleza de su comportamiento es de vital importancia (Silberstein, 2006; Dogrul y Brush, 2016). A través del nuevo Plan Hidrológico (CHG, 2020), la Confederación Hidrográfica del Guadalquivir (CHG) tiene encomendada la gestión del acuífero Almonte-Marismas. La CHG consulta todos los años los resultados del modelo del IGME (Guardiola-Albert et al., 2016) para evaluar el Plan de Extracciones Anual de la Masa de Agua Subterránea de La Rocina (MASb ES050MSBT000055105; CHG, 2020). Por ello, incrementar la calidad del modelo hidrogeológico a través del tratamiento de los datos representa una mejora en la gestión de los recursos del acuífero. La gestión del agua en Doñana sigue siendo un tema de continua preocupación. Ejemplo de ello es la elaboración de una ficha en la lista de Temas Importantes del Tercer Ciclo de Planificación Hidrológica de la Demarcación del Guadalquivir, dedicada en exclusiva a la gestión de las aguas de Doñana con la finalidad de describir su estado y mejorar su gestión (MITECO, 2020).

Tal y como se ha mencionado, el Instituto Geológico y Minero de España (IGME) tiene desarrollado un modelo matemático de flujo del Sistema Acuífero Almonte-Marismas (Guardiola-Albert et al. 2009). Éste ha sido construido en base a versiones anteriores que fueron evolucionando, incrementando su calidad y la representación de la dinámica real del sistema (IGME, 1976; IGME, 1982; IGME, 1994; IGME, 1999; IGME, 2005; IGME, 2009).

La dificultad fundamental de la realización de un modelo matemático siempre estriba en la disponibilidad y calidad de los datos de entrada existentes de la zona de estudio (Silberstein, 2006). De ello depende fundamentalmente la precisión de las salidas obtenidas, lo que permitirá o no depositar la confianza sobre las simulaciones a futuro que se obtienen del mismo (Du et al., 2018). Por supuesto, esto está directamente relacionado con la finalidad para la cual se construye el modelo. La creciente sofisticación de los modelos ha traído consigo el desafío de estimar más parámetros de entrada y estructuras de modelo más detalladas, a partir de datos que a menudo son imperfectos y heterogéneos (Lu et al., 2002; Nøttinger et al., 2005; Manzione et al., 2012).

Actualmente, existen numerosas tecnologías que han hecho posible la recogida de una gran cantidad y diversidad de datos hidrogeológicos (WMO, 2003; WMO, 2014). La aparición de sensores que registran datos de manera continua y no dependen de la presencia de personal humano (Xiong y Wang, 2012), la teledetección (Hoffmann y Sander, 2006) o el desarrollo de potentes herramientas matemáticas para el pretratamiento de los datos (Sahoo et al., 2017; Ahmaadali et al., 2018; Ardabili et al., 2019; Aguilera et al., 2020) son ejemplos de las mejoras que se han introducido en cuanto a las bases de datos y la interpretación automática de la información.

El enfoque tradicional para estudiar la interacción datos de entrada-modelo matemático de flujo se ha centrado principalmente en el uso de datos piezométricos para la calibración y validación del modelo (Castro et al., 1999). El número de datos adquiridos no llega al límite de los *big data* (Babovic, 2005; Chen & Han, 2014), pero si se acerca cuando hay puntos que toman datos piezométricos y de humedad del suelo cada 10 minutos (Kohfahl et al., 2019), o incluso en un lisímetro cada minuto (Kohfahl et al., 2019b; Saaltink et al., 2020). Desde una perspectiva ampliada y cada vez más común, la interacción datos-modelo se ve como un proceso bidireccional (Jones et al., 2005). Por un lado, la información contenida en los datos se puede usar para refinar los modelos, restringir las estimaciones de los parámetros y estados del modelo, e incrementar su nivel de confianza. Por otro lado, los modelos se pueden utilizar para proporcionar un marco para la validación del modelo conceptual, la extracción de información, la interpretación de las observaciones, la generalización e interpolación en zonas no

monitoreadas, y la identificación de datos de alta o baja calidad. Además, la interacción datos-modelo permite la cuantificación de la incertidumbre tanto en el modelo (incluidos los parámetros, la estructura, distribución espacial de la geología y las variables de estado), como en los propios datos (recarga, piezometría, etc). Por tanto, este binomio datos-modelos posee un elevado potencial para la modernización de los modelos matemáticos de aguas subterráneas y el análisis de datos hidrogeológicos, viendo el modelo utilizado en la gestión hídrica como un proceso dinámico, que difícilmente puede darse por bueno o concluido en una fecha determinada. Al igual que los algoritmos de inteligencia artificial (Rajaei et al., 2019) se alimentan de forma continua con datos para una mejora de las predicciones, los modelos matemáticos de agua subterránea han de alimentarse y actualizarse de una forma constante. Tanto en cuanto evolucionen los cambios naturales, climáticos y antrópicos, debe hacerlo la mejora de los modelos para poder avanzar a la vez en la buena gestión de los acuíferos.

1.1. Modelos matemáticos de flujo de aguas subterráneas

1.1.1. Concepto de modelo

Un modelo numérico es una reproducción equivalente al medio físico que permite simular los procesos más importantes que ocurren en el mismo (McDonald & Harbaugh, 1988). Es importante tener en cuenta que, dependiendo de su finalidad, se simplificarán en mayor o menor medida algunas de sus características y procesos físico-químicos.

En el caso de los modelos matemáticos hidrogeológicos de flujo, el objetivo es crear un prototipo virtual que calcule las cantidades y direcciones de flujo subterráneo, la evolución de los niveles piezométricos y el balance hídrico total o parcial (Rushton, 1998), aplicado por primera vez en modelos transitorios por Freeze (1971). Una vez calibrado y validado, el modelo matemático hidrogeológico se convierte en una herramienta predictiva, es decir, se puede utilizar para conocer cuál será su evolución en el futuro, tanto si el régimen se conserva, como si se introducen modificaciones (cambio climático, planes de gestión, escenarios de introducción de usos del agua diferentes, perímetros de protección, etc; Núñez Muñoz, 2001).

La complejidad del funcionamiento de los acuíferos ha obligado al desarrollo de este tipo de herramientas de modelación, que ayudan a gestionar la extracción y utilización de uno de los recursos naturales básicos necesarios para la vida, como es el agua. Los elementos básicos de una modelación hidrogeológica son el modelo conceptual, los datos de entrada, el propio modelo y los datos de salida (Figura 1).

Además de los modelos de flujo, es muy común el uso de modelos de transporte y modelos hidroquímicos (Freeze y Cherry, 1979). Los modelos de transporte reproducen el movimiento de solutos que se introducen puntualmente o de manera difusa con la finalidad de definir plumas de contaminación, posibles áreas de riesgo o perímetros de protección (Zeng 1990; Cazenave y Vives, 2004). Los modelos hidroquímicos se utilizan generalmente para interpretar análisis iónicos de muestras de agua detectando su origen o las mezclas de aguas de las que procede. El software más utilizado es PHREEQC (Parkhurst, 1995).

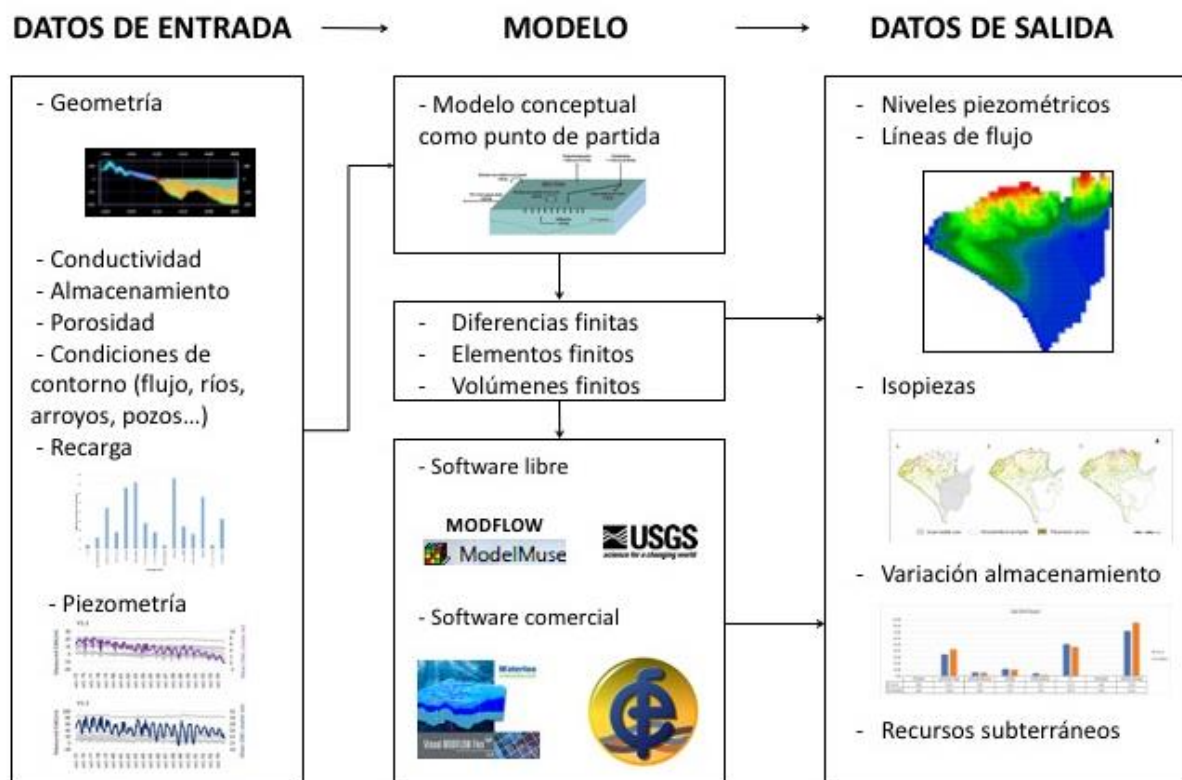


Figura 1. Esquema de los elementos que conforman un modelo matemático de flujo de aguas subterráneas.

1.1.2. Modelos hidrogeológicos de flujo

El código numérico es la representación matemática de los procesos físico-químicos que suceden en el acuífero. Particularmente, esta tesis doctoral se centra en los procesos de flujo.

La construcción de un modelo numérico de flujo requiere, además, de una discretización temporal y espacial. Una vez dividido en unidades o celdas, el programa elegido resuelve la ecuación diferencial de la continuidad mediante matrices en cada una de las celdas en que ha sido discretizado (IHOBE, 2006). La ecuación de la continuidad suele resolverse mediante diferencias finitas, elementos finitos o más recientemente con volúmenes finitos.

En el método de diferencias finitas (Pinder y Gray, 1977) se calcula la carga hidráulica para cada nodo, representando el promedio de las cargas que comprende el nodo. Los parámetros asignados a un nodo se consideran constantes para la celda que representa, y lo más común es que sean celdas cuadradas o rectangulares. El método resuelve de forma iterativa la ecuación tridimensional de flujo o de transporte de masa. En la resolución del flujo, el código MODFLOW (Harbaugh, 2005) es el más utilizado y por tanto el más validado. Una de sus principales ventajas es la numerosa oferta de interfaces gráficas disponibles (Visual Modflow, Model Muse, GW Vistas, GMS). Su desventaja más notable es la limitación para representar geologías complejas, como por ejemplo fallas y fracturas (Kumar, 2019). En el método de elementos finitos se aplican funciones de interpolación para definir la variación de las cargas hidráulicas dentro de cada elemento de malla. Éste es más recomendado para representar geologías complejas ya que los elementos pueden tener cualquier geometría y facilita la obtención de información detallada en elementos críticos como fallas, fracturas, túneles, conductos kársticos o pozos. La principal ventaja que ofrece es la generación automática y adaptación de mallas refinables en zonas que requieran mayor precisión. La heterogeneidad de la malla mejora la optimización computacional. El software de uso más extendido en elementos finitos es FEEFLOW (Diersch, 2013), capaz de simular flujos en medios saturados o variablemente saturados. Por otro lado, debido a su complejidad, el aprendizaje para el uso de este tipo de modelos es mucho menos intuitivo y complicado de dominar (Kumar, 2019). Por último, el método de los volúmenes finitos

permite discretizar y resolver numéricamente ecuaciones diferenciales (McDonald 1971; Mac-Cormack and Paullay 1972). Es una metodología alternativa a las dos anteriores que consiste en que el volumen total de fluido es igual a la suma de los volúmenes de control considerados.

1.1.3. Datos de entrada

En modelación matemática de flujo de aguas subterráneas los datos de entrada son todos aquellos requeridos por un determinado software para su posterior procesamiento conforme a las ecuaciones de flujo. Los datos de entrada, junto con el modelo conceptual y el código numérico a aplicar, conforman las tres partes fundamentales de un modelo hidrogeológico (Figura 1). El modelo conceptual es una representación esquemática del funcionamiento del acuífero, que se realiza con criterio hidrogeológico para definir los procesos más relevantes, las entradas y salidas del sistema, las propiedades físicas más importantes y su distribución espacial (Pholkern et al., 2019).

Existen dos tipos principales de datos que se introducen en el modelo: los espaciales y las series temporales. Los espaciales, se componen de datos topográficos, como el modelo digital de elevaciones; datos geológicos, que se pueden introducir a través de alguna de las variables asociadas a cada litología como, por ejemplo, la conductividad hidráulica; datos hidrológicos, como la definición de las cuencas, ríos, lagunas, manantiales o localización de las estaciones climáticas; y, finalmente, datos hidrogeológicos, como las redes piezométricas o de explotación de aguas subterráneas. Los datos de series temporales van asociados a alguno de los datos espaciales introducidos, a los que se añade información sobre sus variables, por ejemplo, los datos piezométricos, el registro de flujo de una estación de aforos o la recarga.

En cuanto a la relación de los datos de entrada con el acuífero, éstos pueden definir propiedades intrínsecas del mismo como, por ejemplo, la geología y las propiedades que dependen de ella (porosidad, conductividad hidráulica o coeficiente de almacenamiento). Se trata, en general, de propiedades estáticas o constantes en todo el periodo de ejecución del modelo. Existen también propiedades extrínsecas al acuífero como las variables climáticas. La precipitación, por ejemplo, no depende del tipo de

acuífero que se estudie, sino de su localización, altitud, etc. Por último, se pueden definir aquellas variables mixtas, que dependen tanto de factores extrínsecos como intrínsecos al acuífero. Tanto la recarga como la piezometría son buenos ejemplos de este tipo de variables, ya que dependen de factores externos como la precipitación y responden influenciados por propiedades intrínsecas como la geología y su distribución. Toda esta información procede tanto de instrumental hidrogeológico, instalado de forma permanente o medido puntualmente, como de ensayos de laboratorio y ensayos de campo (bombeo, trazadores, etc).

Las propiedades extrínsecas y las mixtas suelen introducirse como datos espaciales a los que se asocia una tabla con las variables que la definen y su valor en cada paso de tiempo que se evaluará para obtener los resultados.

Cabe destacar la dificultad que existe en el cálculo de alguna de estas propiedades mixtas, como por ejemplo la recarga, cuya variabilidad temporal (Balek, 1988) y espacial (Balek, 1988; Custodio, 1997) es muy elevada. Lerner (1990) y Samper (1997) establecen cinco métodos de estimación de la recarga: medidas directas, balance de agua (Samper 1997), técnicas de Darcy (Hendrickx y Walcker, 1997), técnicas de trazadores (Lerner et al. 1990) y métodos empíricos (Lerner et al., 1990). Muchas de las técnicas indirectas calculan la recarga potencial (Rushton, 1988), que es el agua que se infiltra pero puede alcanzar o no el nivel freático (Scanlon et al., 2002). Concretamente para el territorio español, Custodio realiza una revisión conceptual y metodológica para la evolución en la recarga natural de los acuíferos (Custodio 2020).

El primer reto para la construcción de un modelo hidrogeológico comienza con la recopilación y tratamiento de los datos. Las variables estáticas, se obtienen tanto de mapas como de trabajos y ensayos de campo, lo que requiere elevados costes de personal, de tiempo y económicos. En el caso de las series de datos temporales, éstas requieren de homogeneidad, calidad y un número mínimo de datos para que puedan ser de utilidad en la elaboración de modelos que reproduzcan el funcionamiento que se tenga como objetivo. La calidad de los datos de entrada tiene un fuerte impacto en los resultados obtenidos (Baalousha, 2009).

Asociado a este reto, surge el uso de todo un conjunto de metodologías, con una fuerte base matemática, capaces de mejorar la calidad de los datos de entrada, eliminar errores, hacer una estimación de los datos faltantes o interpolar/extrapolar y simular datos espaciales/temporales. La aplicación de este tipo de técnicas, de manera previa a la introducción de los datos de entrada, puede mejorar de manera considerable los resultados obtenidos. Dichas metodologías han sido aplicadas en la presente tesis doctoral para la mejora de los datos de entrada del acuífero Almonte-Marismas. Los resultados obtenidos ponen de manifiesto la necesidad de su aplicación por parte de los organismos públicos que realizan la gestión de los recursos hídricos, sobre todo, en aquellas zonas de elevada vulnerabilidad.

1.1.4. Datos de salida

El resultado de la ejecución de un modelo matemático de flujo construido correctamente debe reproducir de forma bastante aproximada el comportamiento hidrogeológico a la escala elegida. Los datos de salida que se pueden obtener son niveles piezométricos, mapas de isopiezas, líneas de flujo, variación en el almacenamiento, balances hídricos, entradas y salidas de caudales e información sobre los recursos subterráneos disponibles.

El primer paso para comprobar que los resultados obtenidos son correctos es la comprobación de que el balance de masas obtenido es coherente con el modelo conceptual y que la evolución de la variable de estado (p.e. niveles piezométricos) calculada por el modelo numérico reproduce las observaciones de esta variable en los puntos de control (p.e. piezómetros). Seguidamente, la calibración se realiza modificando los parámetros del modelo para que los resultados obtenidos (simulados) sean lo más parecido a los datos observados y medidos (Harbaugh et al., 2000). De esta manera, se comprueba el funcionamiento del modelo. En muchas ocasiones se divide la serie temporal disponible en dos periodos, utilizando el primero para calibrar los parámetros del modelo y el segundo para validar la bondad del modelo fuera del intervalo temporal utilizado en la calibración de los parámetros del modelo (Konikow y Bredehoeft, 1992).

En general, los archivos que se producen tras la ejecución del modelo no son amigables para su uso ya que, en general, incluso utilizando las interfaces disponibles hoy en día, se trata de ficheros de texto que requieren de un post-procesado. Es un proceso muy laborioso, al igual que la calibración, pero necesario para presentar los resultados obtenidos de una manera legible.

Los datos de salida suelen traducirse tanto en tablas como en mapas, a través de los cuales se puede visualizar la información. Su finalidad es poder interpretar los resultados, entender el funcionamiento del acuífero, sacar conclusiones y poder tomar decisiones. Los mapas de uso más común son los de isopiezas, direcciones de flujo, intercambio de flujo entre diferentes cuencas o cuerpos/masas de agua y variación del almacenamiento. Las tablas más utilizadas están asociadas a los balances hídricos y a las series temporales de piezometría simulada. La representación de los resultados que se elija dependerá del objetivo para el cual se construye el modelo.

1.1.5. Aplicación de los modelos matemáticos de flujo

Los modelos matemáticos en Hidrogeología tienen muchas aplicaciones. En primer lugar, sirven para caracterizar el acuífero, definiendo la cantidad de agua disponible y la dirección de flujo (Illman et al., 2010). Se pueden utilizar también para calcular el balance hídrico, de todo el acuífero o de alguna parte del mismo de especial interés, así como para estimar el intercambio de agua entre distintas zonas (Amini y Homayounfar, 2017). Además, también se aplican en el estudio de la relación río-acuífero (Kumar & Nagaraj, 2018) o río-humedal (Rahman et al., 2016) para evaluar el intercambio hídrico producido entre ambos. Otro aspecto importante en donde los modelos matemáticos de flujo presentan gran interés es en la gestión de los recursos hídricos y en su simulación con fines predictivos, especialmente en el contexto de cambio climático (Ghazavi y Ebrahimi, 2019). Sectores como la geología-hidrogeología, la política de agua, la ingeniería civil o la ingeniería agronómica, entre otros, requieren del uso de esta potente herramienta.

1.2. Técnicas estadísticas aplicadas a mejorar los datos de entrada de los modelos matemáticos de flujo

1.2.1. Caracterización de la heterogeneidad geológica en acuíferos regionales

La heterogeneidad geológica de acuíferos de origen sedimentario es una característica que aporta complejidad al comportamiento hidrogeológico, siendo muy importante su conocimiento a escala regional o local para determinar el modelo conceptual. Los límites horizontales y verticales de las litologías no siempre son claros, lo cual dificulta el trabajo de precisión en este tipo de acuíferos. La caracterización espacial de las unidades geológicas tiene como objetivo principal en Hidrogeología conocer mejor los parámetros hidráulicos de cada unidad, como son, por ejemplo, la conductividad hidráulica, la porosidad o la transmisividad (Saha et al., 2017).

Por este motivo, surgen gran cantidad de trabajos en los que mediante técnicas estadísticas se pretende caracterizar mejor la distribución espacial de la geología, tratando datos puntuales de campo provenientes tanto de sondeos, como de geofísica o teledetección (Priyanka y Mohan Kumar, 2019; Shan et al., 2019). A pesar de la disponibilidad de estos datos, es importante tener en cuenta la resolución que ofrecen las técnicas utilizadas en campo para tratar y extrapolar estos datos a otras escalas de trabajo (Vistrand, 2005). En 1978, Gutjahr y colaboradores ya aplicaron un análisis estocástico para estudiar la variabilidad espacial de la conductividad eléctrica efectiva en medios porosos. En las dos últimas décadas ha crecido la cantidad de trabajos realizados para paliar la incertidumbre que genera el desconocimiento de algunos parámetros hidrogeológicos a la hora de realizar modelos matemáticos de flujo (Jankowski y Beck., 2000; Noetiger et al., 2005). Específicamente, las técnicas geoestadísticas han sido utilizadas para estimar variables de interés en una localización donde no han sido medidas, conservando las propiedades de su distribución espacial y cuantificando la incertidumbre de estas estimaciones (Seifert and Jensen, 1991; Renard, 2007). Estas técnicas son comúnmente utilizadas en la modelación de facies en el ámbito del petróleo (Pyrzcz y Deutsch, 2014), pero no tanto en Hidrogeología (Fiori et al., 2016). Existen numerosas técnicas geoestadísticas de estimación y simulación: kriging simple y ordinario, simulación secuencial indicadora (SIS; Alabert, 1987), simulación gaussiana

truncada (TGS; Beucher y Renard, 2016), simulación basada en objetos (OBS; Beursch y Tran, 2002), estadística de puntos múltiples (MPS; Hu y Chugunova, 2008), etc. Cuando se aplica SIS, las hidrofacies previamente definidas se transforman en variables indicadoras (se asigna 1 si está presente y 0 si no lo está), se utilizan los variogramas individuales para el krigado, y por último, se genera una malla 3D en función de la probabilidad asociada a la existencia de esa facies en ese punto (Gómez-Hernández y Srivastava, 1990). En los casos en los que el objetivo es resolver geometrías complejas y se posee una densidad de datos adecuadas, está recomendado el uso del SIS (Deutsch, 2006).

En los últimos modelos del acuífero Almonte-Marismas se representaban los parámetros hidrogeológicos con zonas homogéneas basadas en el mapa geológico de superficie y no en una distribución geológica tridimensional (IGME, 2005; IGME, 2009). En la presente tesis se propone incorporar la variabilidad espacial de la geología en el modelo matemático de flujo del acuífero Almonte-Marismas para evaluar la mejora que produce esta heterogeneidad caracterizada con detalle.

1.2.2. Caracterización espacio-temporal de la precipitación y la recarga

La caracterización espacio-temporal de la precipitación en zonas mediterráneas con elevada influencia atlántica es esencial. En estas zonas, el margen entre el volumen de precipitación y la evapotranspiración es muy reducido, e incluso negativo, lo que genera que las entradas principales del balance hídrico sean aportes mayormente debidos a la recarga. Son muchos los estudios sobre la influencia de la escala espacial y temporal de la lluvia (Buytaert et al., 2006; Nandargi y Mulye, 2012; Moschou et al., 2013; Sapriza et al., 2013; Shi et al., 2014; Zurbieta et al., 2017; Deng et al., 2018). Sin embargo, solo algunos de ellos analizan la influencia de su variabilidad en la evaluación de los recursos hídricos (Narjary et al., 2014; Sapriza et al., 2015; Mumtaz et al., 2018).

Una de las herramientas más utilizadas para el estudio de los datos hidrometeorológicos es la interpolación geoestadística (Carrera-Hernández y Gaskin, 2007; Ly et al., 2011; Chappell et al., 2013; Yang et al., 2015). En España, también ha sido aplicada anteriormente para el cálculo de la precipitación diaria con krigado (Herrera et al., 2012).

El kriging es un procedimiento geoestadístico de estimación de una variable aleatoria con una estructura de correlación espacial, en puntos donde no existen datos disponibles de esa variable. Proporciona una superficie de predicción óptima con una medida de confianza sobre la probabilidad de que la predicción sea cierta. Para ello se caracteriza la estructura de correlación espacial de la variable estudiada a través de su variograma y parte de la premisa de que la variación espacial tiene un patrón homogéneo (Krige, 1951).

En 2012, Pebesma publicó una nueva herramienta geoestadística con la que se podía analizar la variabilidad a través de la modelación del variograma espacio-temporal y un posterior kriging. Desde entonces, ya ha sido aplicado en algunos trabajos hidrogeológicos (De Iaco y Posa, 2012; De Iaco et al., 2012; Mazione et al., 2019; Ruybal et al., 2019; Varouchakis y Hristopulos, 2019). En la presente tesis se ha podido evaluar la importancia que tiene la variabilidad espacio-temporal de la precipitación en el cálculo de la recarga como fuente principal de entradas al acuífero. Es importante cuantificar cómo influyen la variabilidad y escalas temporales y espaciales en los resultados obtenidos en el balance hídrico. Además, este tipo de análisis pueden ayudar a detectar zonas de recarga preferencial y resaltar la adecuada discretización espacio-temporal en los estudios del impacto del cambio climático.

1.2.3. Uso de técnicas de agrupamiento estadístico para el estudio de la piezometría

Las técnicas de agrupamiento o cluster posibilitan generar grupos de datos con las mismas características. Existen numerosos métodos de agrupamiento estadístico para agrupar las medidas de una variable determinada, pero uno de los más utilizados en todos los ámbitos científicos es el k-medias (Hartigan et al., 1979). La particularidad de este método reside en la elección del número de grupos, que se realiza previamente a la clasificación y depende de la media de los datos a agrupar. Es una técnica fácil de implementar y garantiza la existencia de convergencia (Bottou y Bengio, 1995).

En el caso de las series de datos temporales, como es el caso de los datos de piezometría, puede aplicarse un cluster de series temporales. Tiene la desventaja de que la elección del número de grupos es una ardua tarea y se realiza a través de los Índices de Validez

del Cluster (CVI; Maulik et al., 2002). Sin embargo, esta técnica tiene en cuenta la secuencia de la distribución temporal de los datos, por lo que es una técnica más adecuada para la clasificación de datos piezométricos (Sardá-Espinosa, 2019).

Las técnicas de agrupamiento han sido aplicadas en trabajos de Hidrogeología e Hidrogeoquímica (Irawan et al., 2009; Bloomfield et al., 2015; Nakagawa et al., 2019; Yuan et al., 2020; Ashgarinia y Petroselli, 2020), pero únicamente existe un ejemplo del uso de cluster de series temporales de nivel piezométrico (Rinderer et al., 2019).

Una de las aplicaciones de los agrupamientos piezométricos obtenidos con un agrupamiento en la modelación numérica de flujo es la posibilidad de reducir tanto los tiempos de computación como de calibración al poder calibrar un representante de cada grupo en el modelo y no todos los piezómetros a la vez. En modelos de elevada complejidad, con un gran número de datos y gran heterogeneidad en la hidrodinámica del acuífero, es muy importante la optimización tanto del tamaño de los archivos como de la velocidad de ejecución para que sean modelos ágiles en su ejecución y en los que la calibración automática sea factible.

En el caso del acuífero Almonte-Marismas, la piezometría histórica ha sido estudiada en algunos trabajos anteriores (UPC, 1999; Olías Álvarez y Rodríguez Rodríguez, 2013; CHG, 2019) a través de técnicas visuales pero ninguno de ellos aportan una clasificación sobre el comportamiento piezométrico. En general, analizar las características y los patrones de la compleja fluctuación de los niveles piezométricos (GWL) es importante para conseguir una clasificación objetiva (Nguyen et al., 2015), que puede aportar una mejora en el conocimiento del comportamiento del acuífero, cuyo estado cualitativo y cuantitativo se encuentra en riesgo.

Las consecuencias del cambio climático que conllevan el incremento de la evaporación, y los eventos extremos, incrementados por la presión antrópica representan una seria amenaza para el estado hidrológico del acuífero Almonte-Marismas (Guardiola-Albert y Jackson, 2011), que se encuentra sobreexplotado, principalmente, por la actividad agrícola. A través de la clasificación piezométrica con técnicas de agrupamiento, es posible analizar los efectos de la expansión agrícola actual. Para evaluar el comportamiento de las técnicas disponibles, se han aplicado tres técnicas diferentes:

clasificación visual como análisis cuantitativo de base, k-medias (MacQueen, 1967) y agrupamiento estadístico de series temporales (TSC; Sardá-Espinosa, 2019). Los resultados obtenidos han ayudado a comprender mejor la hidrodinámica compleja y antropizada del acuífero estudiado.

1.3. Zona de estudio

1.3.1. Descripción general

El sistema acuífero Almonte-Marismas está situado entre las provincias de Huelva y Sevilla, al suroeste español. En el territorio que ocupa se encuentran el Parque Nacional y el Parque Natural de Doñana (Figura 2), que conforman actualmente la unidad de gestión medioambiental denominada Espacio Natural de Doñana. El acuífero está dentro de los términos municipales de Almonte, El Rocío, Matalascañas y Mazagón, entre otros, y está cercano a las ciudades de Sevilla y Huelva (Figura 2). Se halla en las cuencas bajas de los ríos Guadalquivir y Tinto y presenta un área total de 2.409 km². Se encuentra conectado lateralmente por el noreste con el acuífero de El Aljarafe. Por el este, limita con el río Guadalquivir, por el suroeste con el océano Atlántico y por el norte con los materiales impermeables de la Sierra Morena y el río Tinto (Figura 2).

En el año 2016, según el segundo ciclo de planificación hidrológica de la CHG, el acuífero se dividió administrativamente en 5 Masas de Agua Subterránea (BOE, Real Decreto 1, 2016): "Almonte", "Marismas", "Marismas de Doñana", "Manto Eólico de Doñana", y "La Rocina". Tres de estas Masas de Agua Subterránea, "La Rocina", "Almonte" y "Marismas", están en trámites de ser declaradas en riesgo de no alcanzar el buen estado cuantitativo, y una de ellas, La Rocina, en riesgo de no alcanzar el buen estado químico desde mayo de 2019 (BOE, Anuncio 39064, 2019) y que se encuentra en exposición pública actualmente.

Las cotas topográficas máximas y mínimas en el área del acuífero son 190 m s.n.m. y 0 m s.n.m., respectivamente (Figura 2). Es decir, es una zona con desniveles muy suaves y paulatinos, siguiendo la configuración típica de una zona de tramo bajo de río y desembocadura fluvial. El acuífero tiene un espesor total máximo de 425 m, variando su cota entre 190 y -235 m s.n.m.

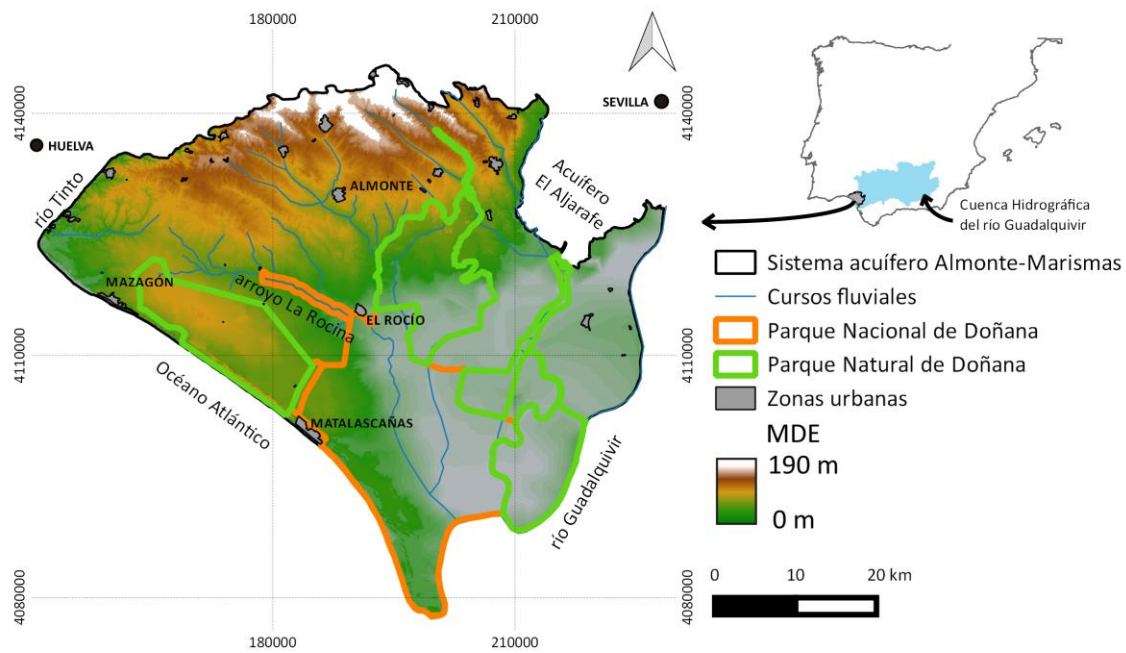


Figura 2. Modelo digital de elevaciones y localización geográfica del acuífero Almonte-Marismas.

La zona se caracteriza por un clima mediterráneo suave, con veranos secos e inviernos húmedos y una temperatura media de unos 24 °C (IGME, 2005). La lluvia anual media es de 550 mm para el periodo de estudio de esta tesis (1975-2016). La mayor intensidad de precipitación registrada se produjo en la estación meteorológica de Niebla y fue de 168 mm/día en el mes de noviembre de 2001.

Geológicamente el acuífero Almonte-Marismas ha sido descrito extensamente por Salvany y Custodio en numerosos trabajos (Salvany y Custodio, 1995; Salvany y Custodio, 2001; Salvany et al., 2010; Salvany et al., 2011). Los materiales que componen el acuífero fueron depositados entre el Mioceno y el Cuaternario, formando una serie monoclinial que buza suavemente hacia el sur-suroeste. De más antiguo a más moderno, la serie comienza con las margas azules del Mioceno, que conforman el sustrato del acuífero sobre el que se depositan las unidades limo-arenosas del Plioceno, con espesores de entre 10 y 150 m, crecientes hacia el sur, y el manto eólico y dunas antiguas sobre éstas. Por encima, se encuentran las unidades aluviales del Pleistoceno, que no afloran en superficie y están formadas por gravas, arenas y lentejones de arcilla. Sobre las unidades pleistocenas se sitúan los materiales holocenos de marisma, formados por arcillas que colmatan la desembocadura del río Guadalquivir y poseen espesores de hasta 200 m.

Por último, los materiales más recientes son los aluviales de los ríos y arroyos y las arenas de las playas y dunas costeras actuales. Estas últimas, formadas por arenas silíceas extendidas paralelamente a la costa del océano Atlántico (Figuras 3 y 4) originan voluminosas dunas móviles con espesores de hasta 150 m (Vela, 1984). En general, los materiales descritos anteriormente no han sufrido procesos de deformación severos y están dominados únicamente por procesos sedimentarios (Lario et al., 2001; Ruiz et al., 2004; Cáceres Puro y Rodríguez Vidal, 2008). Tan solo se han descrito un conjunto de fallas de dirección E-S que afectan a capas más antiguas y no al propio sistema acuífero (Salvany et al., 2011).

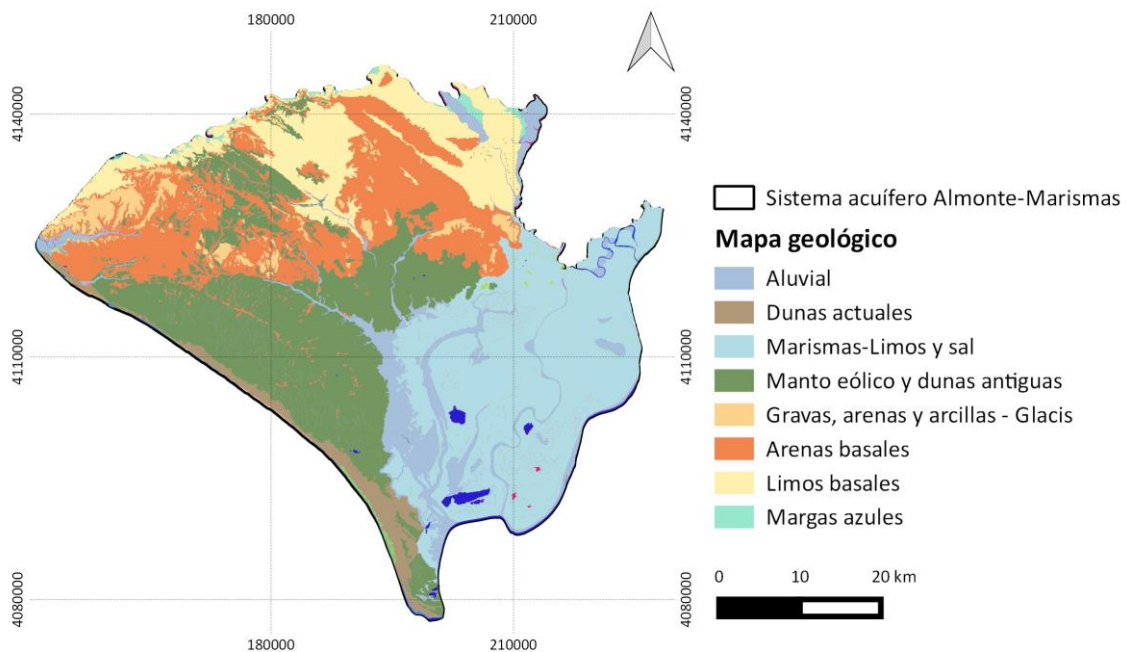


Figura 3. Mapa geológico simplificado del acuífero Almonte-Marismas.

La complejidad sedimentaria que caracteriza al sistema acuífero Almonte-Marismas se traduce a su vez en una elevada complejidad hidrogeológica. Se distinguen dos comportamientos hidrogeológicos claros, la zona norte y central, que es un acuífero libre, y la zona situada bajo la marisma, que presenta un comportamiento de acuífero confinado, con una franja semiconfinada en las dos zonas (Custodio et al., 2008). Tanto la zona libre como la confinada se encuentran limitadas a muro por las margas azules del Mioceno (Salvany, 2010). En cuanto al techo, la zona confinada limita por las arcillas de marisma. Las unidades acuíferas principales están formadas por los materiales aluviales del Pleistoceno. Los estudios hidrogeológicos de detalle muestran la presencia

de numerosas intercalaciones y lentejones de arcilla, lo que, en algunas zonas concretas, transforma el acuífero en un modelo multicapa (Figura 4; Martín-Machuca et al., 1992; Salvany y Custodio, 1995). Esta heterogeneidad lateral de los depósitos complica la posibilidad de describir las direcciones de flujo a escala local. El flujo general presenta direcciones N-S o NO-SE.

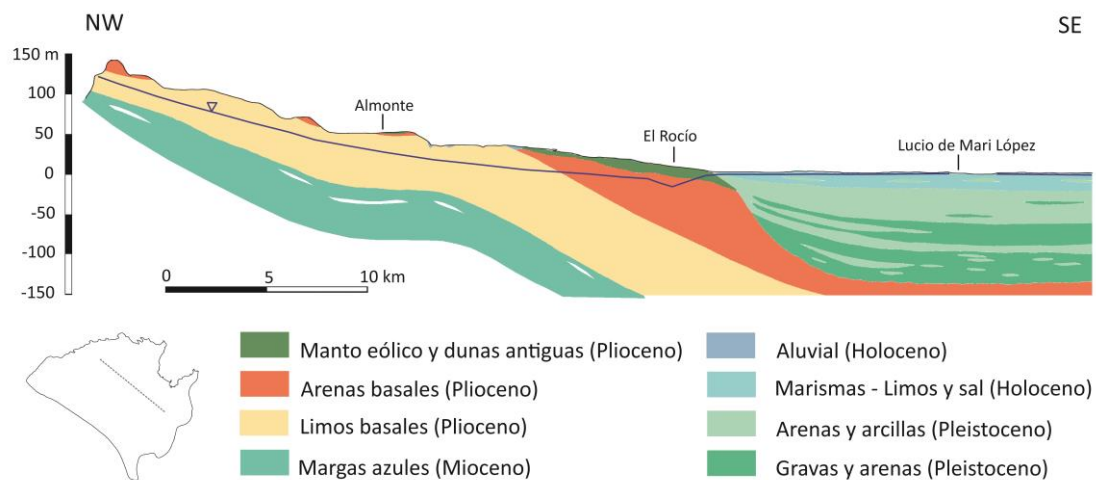


Figura 4. Corte geológico del acuífero Almonte-Marismas dirección NO-SE cruzando por Almonte (Modificado de Martín-Machuca et al. 1992). Nivel piezométrico medio medido en 2016.

Los parámetros hidrogeológicos tienen gran variabilidad espacial, dependiendo sobre todo de la granulometría y el espesor de los depósitos. La transmisividad varía entre los órdenes de magnitud de 10^2 y 10^3 m²/d (Martín-Machuca et al., 1992) y la permeabilidad tiene una gran variabilidad, entre 0.001 y 30 m/día (Guardiola-Albert et al., 2009). La recarga se produce principalmente por infiltración de las precipitaciones, siendo la media anual de 5.8×10^5 m³/d (Guardiola-Albert et al., 2009). La descarga principal tiene lugar por el drenaje de los cursos de agua superficiales (arroyos y ríos), las salidas al mar, las descargas en el límite entre las arenas eólicas y la marisma, los intercambios de flujo con las lagunas permanentes y temporales y, de una manera notable, por la extracción de agua mediante pozos para regadío, uso industrial, turístico y abastecimiento de la población.

En la figura 4 se ha representado el nivel piezométrico medio medido en 2016. Se observa un cono de depresión piezométrica en los alrededores de la aldea de El Rocío,

asociado a los bombeos de agua subterránea en las zonas agrícolas que bordean Doñana.

Además de la relación del acuífero con la marisma, existe una importante relación aguas superficiales – aguas subterráneas localizada principalmente en las lagunas temporales situadas sobre las arenas de las dunas móviles, cercanas a la costa del Océano Atlántico. La elevada explotación de las aguas subterráneas, tanto en Matalascañas como aguas arriba de esta población, ha originado la desecación total de algunas de las lagunas más cercanas al complejo turístico, como por ejemplo las lagunas de Brezo o Charco del Toro (Serrano y Zunzunegui, 2008; Dimitriou et al., 2017). También se ha producido la modificación de los hidroperiodos de algunas de las lagunas (Díaz-Paniagua y Aragonés, 2015; Dimitriou et al., 2017). Una excepción a esta disminución del área encharcada lo constituye la única laguna permanente, la Laguna de Santa Olalla (Fernández- Ayuso et al., 2019).

1.3.2. Importancia ecológica del área de estudio

El Parque Nacional de Doñana, declarado como tal en 1969, es una de las mayores reservas ecológicas de Europa, con más de 50.000 ha. Fue incluido como Reserva de la Biosfera en 1980 y en la Lista de Humedales de Importancia Internacional en el convenio de Ramsar en 1982. En 1989 fue constituido por la Junta de Andalucía como Parque Natural. Años más tarde, en 1995, y tras la evaluación de su estado, se inscribe en la lista de Patrimonio Mundial por la UNESCO. Doñana también es espacio protegido de la Red Natura 2000, Zona Especial de Protección para las Aves (ZEPA) y Zona Especial de Conservación (ZEC). Actualmente, y desde 2006, está integrado en la Red de Parques Nacionales y su gestión corresponde, en exclusividad, a la Junta de Andalucía. Las competencias de cuidado del parque fueron transferidas por la Administración General del Estado a la Junta de Andalucía mediante el Real Decreto 712/2006, de 9 de junio (BOE, 2006).

El Parque Nacional de Doñana se considera un ejemplo icónico de importancia ecológica, cuyo equilibrio es muy delicado (Scheffer et al., 2015). Existe una gran cantidad y diversidad de especies que forman parte de su biota. Muchas de ellas son dependientes de las condiciones hidrogeológicas de los encharcamientos (Llamas, 2005) de las lagunas

temporales, permanentes y de la marisma (Díaz-Paniagua y Aragonés, 2015). Por tanto, existe una relación muy estrecha entre el buen estado ecológico del medio y el binomio cantidad-calidad de las aguas que alimentan las marismas de Doñana (García-Murillo y Sousa, 1999; Muñoz- Reinoso, 2001; Sousa y García-Murillo, 2003).

Existen tres grandes sistemas ecológicos en el Parque Nacional de Doñana: marismas, dunas móviles y cotos, acompañadas de una gran zona de contacto o ecotono. La marisma es el más extenso, ocupando 27.000 ha de zona húmeda para el paso, cría e invernada de aves, tanto europeas como africanas (Fotografía 1; IGME, 1983). Se inunda a través de los aportes de agua de los arroyos y las lluvias, y presenta una estacionalidad muy marcada (Díaz-Delgado et al., 2016). Las playas tienen un perfil que se modifica incesantemente formando bajos y flechas a lo largo de todo el litoral. Asociadas a éstas se encuentran las dunas móviles, cuya vegetación forma obstáculos que serán el origen de las dunas vivas (Vela, 1984). En su movimiento, éstas forman zonas de corral colonizadas principalmente por los característicos pinos de Doñana. Los cotos representan una etapa intermedia del bosque mediterráneo caracterizada por el matorral. Su aspecto varía poco durante el año y alberga en Doñana al lince ibérico y otros mamíferos de gran tamaño. Finalmente, la transición entre la marisma y los cotos está representada por una franja de entre 200 y 1500 m, el ecotono o la vera, de una elevada riqueza ecológica (Díaz-Paniagua y Aragonés, 2015). Coincide con la zona de contacto entre las arenas y las arcillas, donde rezuman pequeñas zonas húmedas y crecen junqueras y pastizales (Llamas, 1988). En la vera confluyen las especies de ambos ecosistemas (Virgos et al., 1983; MITECO, 2020b).



Fotografía 1. Marismas de Doñana en El Rocío 10/03/2020. Fuente: Ana Fernández Ayuso.

1.3.3. Problemática medioambiental y socioeconómica

Desde inicios de los años 70, la actividad socioeconómica principal de la zona de Doñana es la agricultura, tanto en términos de producción como de empleo. Esto ha provocado una intensa y creciente explotación de las aguas subterráneas. Aunque el agua extraída tiene un uso agrícola principalmente, también se utiliza para abastecimiento de la población, turismo estacional o la industria. Los cultivos principales son el fresón, los frutos rojos, el arroz, el olivar y el viñedo. En las encuestas realizadas por la Confederación Hidrográfica del Guadalquivir y la Cámara Agraria en 1992 ya estimaron una explotación anual promedio de 73 hm³/a. En 2003, la superficie total cultivada de fresón era de 4.536 ha (Fundación Doñana 21, 2003). Actualmente, el cultivo bajo plástico de frutos rojos (arándano, fresa y frambuesa) es de 7.995 ha. Se estima que 1.653 ha corresponden con zonas de regadío fuera de las zonas de “suelos agrícolas regables”, es decir, en situación irregular (WWF, 2019).

La evolución de los sistemas agrícolas hacia formas de cultivo más intensivas y el crecimiento de las zonas de cultivo, añadido a que existen más de 1.000 pozos en situación de ilegalidad en la zona, provoca que la cantidad de recursos hídricos que se extraen anualmente carezca de control (WWF, 2016). Únicamente se estiman las cantidades extraídas por los pozos legales bajo el dominio de la Confederación Hidrográfica del Guadalquivir (Tablas 1a y 1b), en la que ya destaca el elevado índice de

explotación al que está sometido el acuífero en alguna de las masas de agua subterránea. En las Tablas 1a y 1b se utiliza el término Recurso disponible que según la Instrucción de Planificación Hidrológica se calcula como “El Recurso disponible se obtendrá como diferencia entre los recursos renovables (recarga por la infiltración de la lluvia, recarga por retorno de regadío, pérdidas en el cauce y transferencias desde otras masas de agua subterránea) y los flujos medioambientales requeridos para cumplir con el régimen de caudales ecológicos y para prevenir los efectos negativos causados por la intrusión marina”. El Índice de Explotación se calcula dividiendo las extracciones entre el recurso disponible (CHG 2019b).

Tabla 1a. Recarga anual, recurso disponible y extracciones de las nuevas masas de agua subterránea (extraída de CHG 2019b).

Masa	Recarga [m ³ /d]	Recurso disponible [m ³ /d]	Extracciones [m ³ /d]	Índice de Explotación [m ³ /d]
Almonte	1.9x10 ⁵	1.5x10 ⁵	7.7x10 ⁵	50.70%
Marismas	9.3x10 ⁵	7.4x10 ⁴	9.3x10 ⁵	125.00%
Marismas de Doñana	3.4x10 ⁵	8.8x10 ³	0	0%
Manto Eólico Litoral	2.0x10 ⁵	8.0x10 ⁵	1.2x10 ⁵	14.88%
La Rocina	2.2x10 ⁵	1.0x10 ⁶	1.1x10 ⁶	106.69%

Tabla 1b. Recarga anual, recurso disponible y extracciones de las nuevas masas de agua subterránea (extraída de CHG 2019b).

Masa	Recarga [hm ³ /año]	Recurso disponible [hm ³ /año]	Extracciones [hm ³ /año]	Índice de Explotación [hm ³ /año]
Almonte	69.16	55.32	28.05	50.70%
Marismas	33.78	27.02	33.78	125.00%
Marismas de Doñana	12.32	0.32	0	0%
Manto Eólico Litoral	73.29	29.29	4.36	14.88%
La Rocina	80.44	36.44	38.88	106.69%

En este marco de sobreexplotación del acuífero, y teniendo en cuenta las características del flujo descritas en el apartado 1.2.1. de esta tesis, las aguas que alimentan la marisma de Doñana se encuentran afectadas tanto en términos de cantidad como de calidad (Kohfahl et al., 2019). La Confederación Hidrográfica del Guadalquivir, en el Informe de declaración de La Rocina como masa en riesgo de no alcanzar el buen estado cuantitativo y químico, publica una tabla resumen (Tabla 2) que proporciona una visión global del estado actual.

Tabla 2. Evaluación del estado cuantitativo y químico de las nuevas masas de agua subterránea (extraída de CHG 2019b).

Masa	Estado cuantitativo	Estado químico	Estado Global
Almonte	Malo	Malo	Mal estado
Marismas	Malo	Malo	Mal estado
Marismas de Doñana	Bueno	Bueno	Buen estado
Manto Eólico Litoral	Bueno	Bueno	Buen estado
La Rocina	Malo	Malo	Mal estado

Las circunstancias, tanto naturales como socioeconómicas, de la zona de estudio implican la necesidad de una gestión optimizada y sostenible de los recursos hídricos que permita establecer un equilibrio entre la actividad socioeconómica y la conservación ecológica en el acuífero Almonte-Marismas. Esta necesidad ha sido reflejada en los numerosos litigios que ha ocasionado. En 1990, la Comisión Europea inició un trámite para denunciar el descuido sobre el Parque Nacional de Doñana (IP/90/571), destacando las elevadas extracciones ocasionadas por la agricultura y el turismo. De nuevo, en 2019, la Comisión Europea denunció al Gobierno de España ante el Tribunal de Justicia de la Unión Europea (Oficial Journal of the European Union. Case C559/19) por no proteger suficientemente las aguas subterráneas que alimentan Doñana considerando que no se han tomado las medidas adecuadas. En la denuncia se detallan los artículos incumplidos de la Directiva Marco del Agua (Directiva 2000/60/EC) y se destacó que no se estaban llevando a cabo acciones para evitar el deterioro de los hábitats dependientes de este humedal. Este hecho no es más que un reflejo de los numerosos avisos que científicos y

conocidos grupos ecologistas (especialmente WWF) han puesto en valor, y que dista de las peticiones realizadas por los agricultores de la zona, que reivindican estar trabajando bajo presión y con una gran escasez de agua.

1.4. Modelo matemático del acuífero Almonte-Marismas

1.4.1. Descripción general

El modelo matemático del acuífero Almonte-Marismas utiliza MODFLOW 2005 (Harbaugh 2005) a través de la interfaz gráfica Model Muse (Winston 1999).

El modelo se sustenta sobre una malla de 500x500 m de celdas orientadas N-S, con origen (límite inferior izquierdo) en las coordenadas UTM X= 150500 Y=407600.

El límite superior del acuífero se ha establecido a partir del modelo digital de elevaciones con resolución de 20 m. El límite vertical inferior ha sido definido mediante la realización de un krigado con deriva externa del espesor del acuífero (Wackernagel, 1998; Chiles & Delfiner, 1999) obtenido de los datos geológicos y geofísicos existentes. El krigado se ha realizado sobre los valores de espesor y no sobre la cota absoluta del muro impermeable con la finalidad de que el resultado obtenido no superase en cota al techo del acuífero, asegurando así la coherencia de los límites del mismo.

Los límites de su extensión horizontal han sido definidos atendiendo a las características geológicas, topográficas y de disponibilidad de datos. Al norte queda limitado por el contacto entre las margas azules y las arenas basales. Al sur-este la existencia del río Guadalquivir se ha considerado como un divisorio hidrogeológica (límite impermeable). Al sur-oeste se encuentra el Océano Atlántico, cuyo contacto con el acuífero ha sido definido mediante celdas de nivel constante sin tener en cuenta el efecto de las mareas y definida a 1 m de distancia de la línea de costa.

En cuanto a la discretización vertical, el modelo se divide en dos capas (Figura 5). La superior, activa en la zona suroeste de la zona de estudio, definida como acuífero libre. La inferior, activa en toda el área del acuífero, ha sido modelada como acuífero confinado bajo el acuífero libre de la capa superior y bajo las marismas.

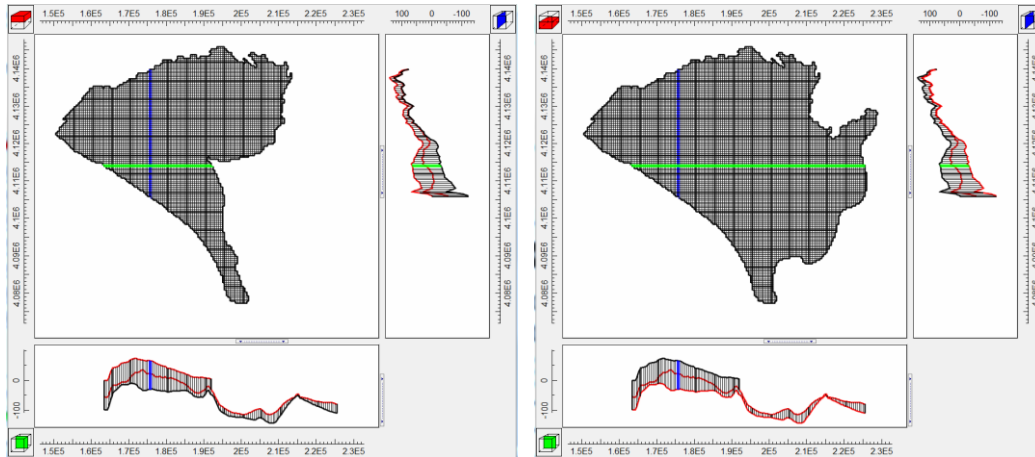


Figura 5. *Izquierda: capa superior del modelo. Derecha: capa inferior del modelo. Corte longitudinal (línea verde) representado debajo de la figura central y corte transversal (línea azul) representado a la derecha de la figura central.*

Se ha ejecutado el modelo estacionario y el transitorio de 1975 a 2016 a escala mensual.

El modelo estacionario ha sido calibrado con el mapa piezométrico obtenido de los datos de piezometría de los años 70, asumiendo que en esa época el sistema se encontraba en un estado no influenciado. Para la realización de esta tesis doctoral se ha ejecutado el modelo en estado estacionario con la finalidad de valorar cómo influye la mejora de los datos de entrada en los resultados del modelo matemático y establecer una metodología que se pueda aplicar en el estado transitorio.

1.4.2. Condiciones de contorno

Con respecto a los parámetros hidráulicos, la recarga ha sido introducida mediante la definición de 15 zonas homogéneas (Custodio et al., 2009) que siguen criterios geológicos e hidrogeológicos. La recarga se calculó con el software TRASERO (Padilla y Delgado, 2013) mediante el análisis de series temporales de datos hidrológicos eligiendo una estación representativa para cada una de las zonas (Figura 6). Previamente, los valores faltantes de datos de temperatura y precipitación se rellenaron mediante el algoritmo missforest en R (Steckhoven y Bühlmann, 2012; Steckhoven, 2013). Los valores de recarga máximos obtenidos para cada zona oscilan entre 67 y 488 mm al año y entre 23 y 78 mm al año para la mediana.

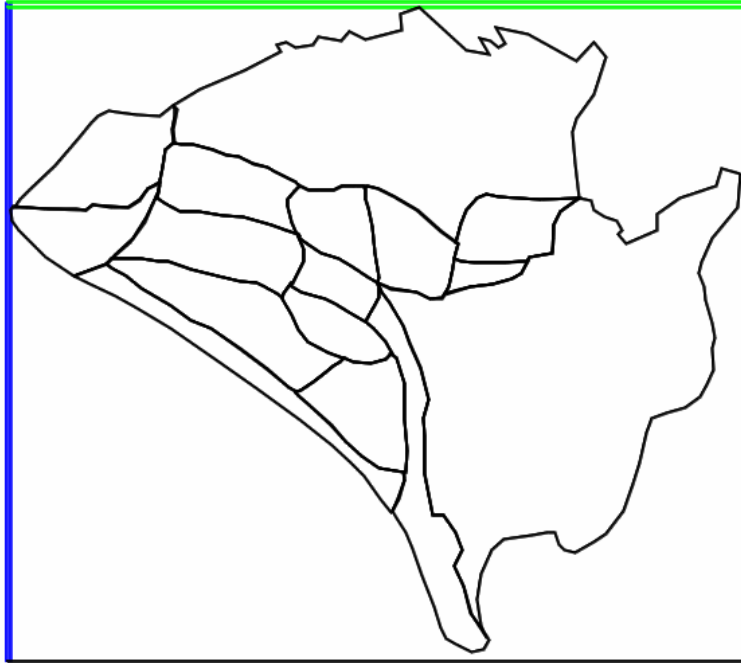


Figura 6. Zonificación espacial de la recarga.

Los ríos Guadamar y Tinto se han simulado como celdas de tipo RIVER (McDonald & Harbaugh, 1988) y los arroyos superficiales como celdas de tipo STREAM (McDonald & Harbaugh, 1988; Figura 7), simplificando los principales cursos de agua. Se definieron varias zonas que aproximan los distintos tipos de descarga con la condición de contorno DRAIN para resolver el contacto con la zona de marisma (McDonald & Harbaugh, 1988):

- Dren Marisma note: situado en el borde norte de la marisma (línea verde en la Figura 7), se le atribuye una conductancia de $350 \text{ m}^2/\text{día}$ y el nivel del dren está situado un metro por debajo de la superficie del terreno.
- Drenes manantiales Cuesta Maneli: están representados en el borde cercano a la costa (línea amarilla en la Figura 7), se le atribuye conductancia de $650 \text{ m}^2/\text{día}$, y el nivel del dren está situado 0.5 m por debajo de la superficie del terreno.
- Drenes ecotono de la vera: están representados en el borde oeste de la marisma (línea negra en la Figura 7), se le atribuye una conductancia de $650 \text{ m}^2/\text{día}$, el nivel del dren está situado 0.5 m por debajo de la superficie del terreno.
- Drenes de la propia marisma: cuya conductancia tiene 1×10^{-5} , el nivel de dren está situado en la superficie del terreno y está situado en la capa 2 al ser la capa 1 impermeable (línea rosa en la Figura 7).

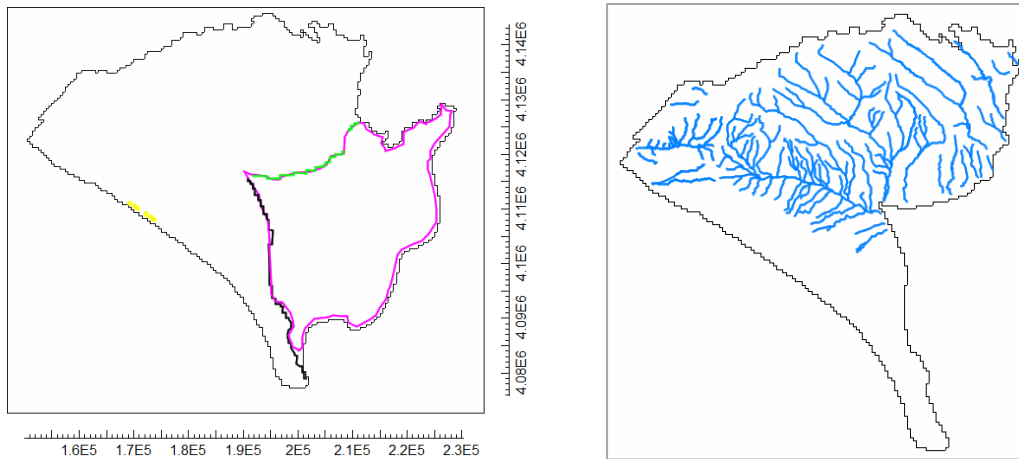


Figura 7. Izquierda: Situación de la condición de contorno de DRAIN en Cuesta Maneli (amarillo), la Vera-Retuerta (negro), borde norte de la marisma (verde) y marisma (rosa fuxia).
Derecha: Situación de la condición de contorno STREAM.

Para la simulación de los límites de intercambio con otras masas de agua, en concreto en la zona norte (línea naranja en la Figura 8) y en el contacto con la masa de El Aljarafe (línea roja en la Figura 8) se han introducido celdas de potencial constante (CHB; McDonald & Harbaugh, 1988) sin definir valores de flujo límite. El flujo varía según la diferencia de potencial.

La evapotranspiración freática introducida simula la masa forestal de la zona (eucaliptos y pinos) estudiada en la construcción del modelo de 2011 (IGME, 2011). La evapotranspiración introducida en las zonas con masa forestal es 1,51 mm/día y es máxima a una cota de 0.2 m por debajo de la superficie del terreno (zonas rojas en la Figura 8).

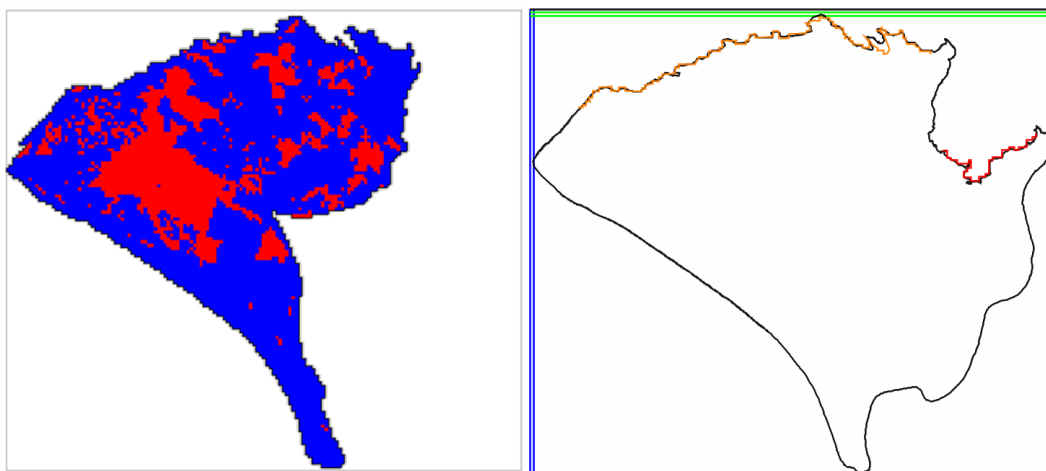


Figura 8. Izquierda: Distribución de espacial de las zonas de evapotranspiración freática del estado estacionario. **Derecha:** Localización de las celdas de potencial constante (CHB).

1.4.3. Parámetros hidráulicos

Los valores de permeabilidad y coeficiente de almacenamiento siguen la distribución espacial de la geología según la serie de mapas geológicos MAGNA 1:50.000 (Figura 9). Concretamente, los valores de permeabilidad varían entre 0.001 y 50 m/día.

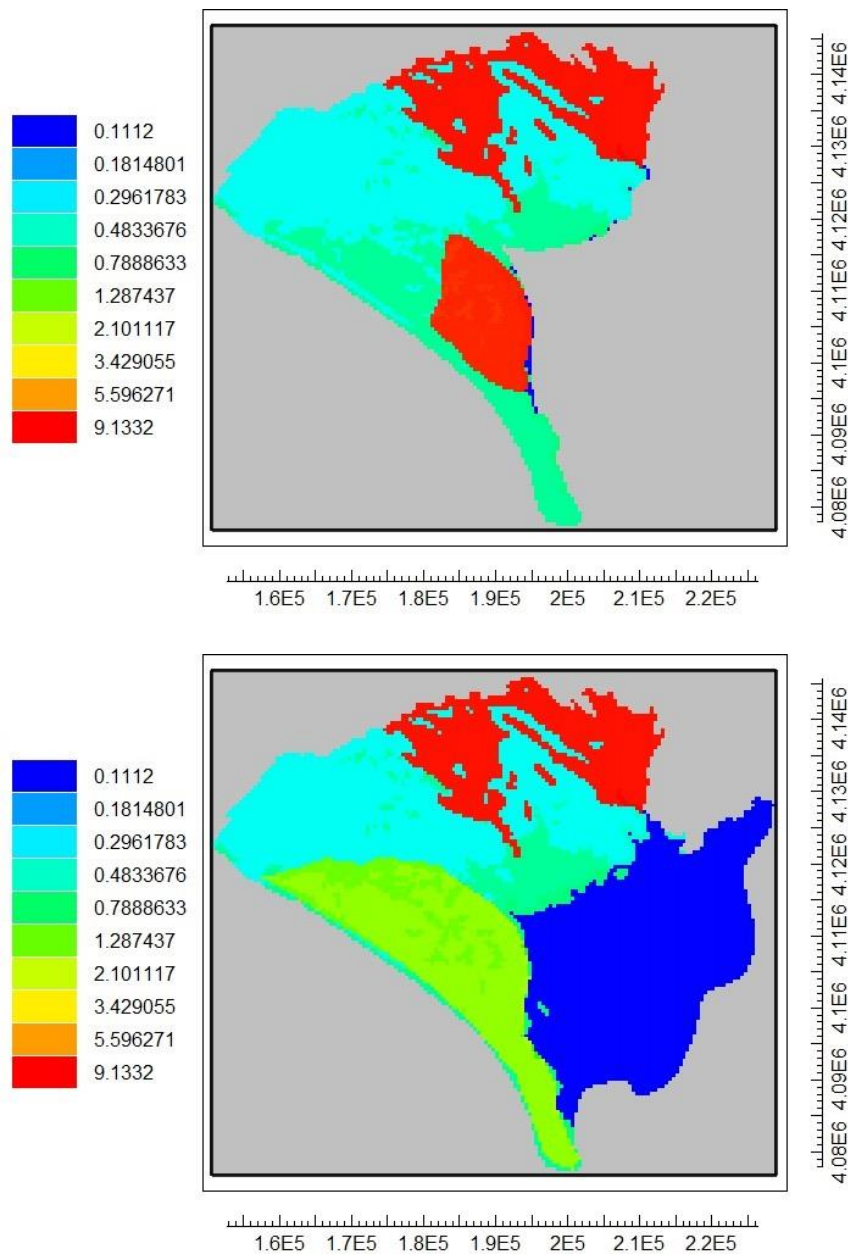


Figura 9. Arriba: permeabilidad de la capa superior. **Abajo:** permeabilidad de la capa inferior.

1.4.4. Calibración

Se utilizó el paquete LPF (Harbaugh, 2005) por ser el más recomendado por cómo calcula las conductividades entre celdas y, además, permite realizar una calibración automática.

Se definieron parámetros en cada zona de permeabilidad para poder hacer un análisis de sensibilidad y la calibración automática.

La piezometría resultante y los errores en los puntos considerados como representantes del estado estacionario están dentro del orden de magnitud asumible para un modelo regional de estas dimensiones. Se obtuvo un RMSR (Root Mean Square Residual) de 3.62.

Capítulo 2: Objetivos

La presente tesis aporta aspectos novedosos relacionados con la mejora en el análisis, interpretación e introducción de los datos de entrada en los modelos hidrogeológicos de flujo y, en particular, hace un estudio estadístico de los datos de geología, precipitación-recarga y piezometría. Estas tres propiedades representan variables intrínsecas, extrínsecas, dinámicas y estáticas del acuífero. Se ha demostrado que, mejorando la calidad de los datos, se mejora también la calidad del modelo, y por tanto de los resultados obtenidos con el mismo.

La discusión y resultados de esta tesis pueden ser de gran interés en general para la investigación de modelos matemáticos de flujo de aguas subterráneas, y en concreto en el ámbito de la gestión de los recursos hídricos, especialmente en zonas de gran estrés hídrico o áreas que merecen un especial cuidado y conservación debido a sus características ecológicas.

2.1. Objetivo principal

El objetivo principal de esta tesis doctoral es evaluar las interacciones entre los datos de entrada y el modelo matemático de flujo en el acuífero Almonte-Marismas (Huelva, España). Los resultados obtenidos se pueden incluir como buenas prácticas en la elaboración de modelos que se utilicen para la gestión de las aguas subterráneas de la zona de estudio, además de ser aplicables a cualquier otra zona geográfica.

2.2. Objetivos parciales

Para alcanzar el objetivo principal, se han fijado los siguientes objetivos parciales:

- El primer objetivo parcial es obtener un modelo geológico digital 3D mediante la recopilación de todos los datos geológicos y geofísicos existentes en el área de estudio y la inclusión de este modelo geológico en el modelo de flujo subterráneo desarrollado con ModelMuse (USGS). Para ello, las diferentes litologías se han traducido a hidrofacies, unidades hidrogeológicas que presentan el mismo comportamiento, cuya distribución espacial ha sido

interpolada gracias a una simulación geoestadística. Este objetivo trata de mejorar la definición de variabilidad espacial de la conductividad hidráulica (propiedad intrínseca), en un acuífero con una alta heterogeneidad.

- El segundo objetivo es estudiar y cuantificar geoestadísticamente la distribución espacio-temporal de la precipitación en todo el área del acuífero a partir de 113 estaciones climáticas a través de herramientas estadísticas. También, analizar las consecuencias que esta variabilidad tiene sobre el cálculo de la recarga en un acuífero tan llano, y los cambios en las tendencias durante los 41 años de las series pluviométricas (1975-2016). Todo ello enfocado a la búsqueda de estrategias para la adaptación de los planes de gestión de los recursos hídricos al cambio climático. En este segundo objetivo, la mejora de la entrada de datos del modelo se aplica sobre una propiedad extrínseca y dinámica del propio acuífero, que depende de los registros climáticos.
- El tercer y último de los objetivos parciales de esta tesis doctoral es desarrollar una metodología para agrupar estadísticamente los piezómetros según su comportamiento hidrodinámico mediante el análisis de las series piezométricas históricas, determinando los distintos grupos de respuestas ante los diferentes regímenes de explotación del sistema. En este caso, la implementación de la mejora se aplica sobre una propiedad intrínseca y dinámica del acuífero. Para la consecución de este último objetivo se utiliza una herramienta estadística de agrupamiento de series temporales que hasta ahora ha sido escasamente empleada para el análisis del comportamiento piezométrico, siendo el primer caso conocido a nivel nacional.

Capítulo 3. Resumen metodológico

La figura 10 muestra el esquema metodológico que se ha seguido para su desarrollo. Para la consecución de los objetivos de esta tesis se han aplicado tres técnicas estadísticas diferentes:

1. La Simulación Secuencial Indicadora (Gómez-Hernández y Srivastava, 1990) es una técnica que permite modelizar en 3D la información geológica y geofísica. Las variables se definen a través de la modelación de los variogramas indicadores. La simulación condicionada que se obtiene conserva los porcentajes de existencia de cada una de las variables de la muestra (Soares, 1998).
2. El krigado espacio-temporal (Pebesma, 2012) se aplica para interpolar los datos de precipitación. Es una técnica de estimación a distintas escalas espaciales y temporales en periodos de tiempo o lugares geográficos donde no existe información empírica. La aplicación de esta técnica aporta la posibilidad de tener en cuenta la correlación espacio-temporal de la precipitación.
3. Finalmente, con la implementación del argumentación estadístico de series temporales (Sardá-Espinosa, 2019) a los datos de piezometría se puede obtener una clasificación de comportamientos hidrogeológicos de los piezómetros para diferenciar áreas con distinta influencia antrópica. Además, se puede observar la presencia de comportamientos anómalos que pueden ser debidos a errores en la toma de datos, heterogeneidades geológicas, cambios climáticos o variaciones en los usos del suelo a lo largo del periodo de estudio.

La metodología de estudio se detalla de manera más completa en cada uno de los artículos que componen la tesis doctoral (Capítulo 4).

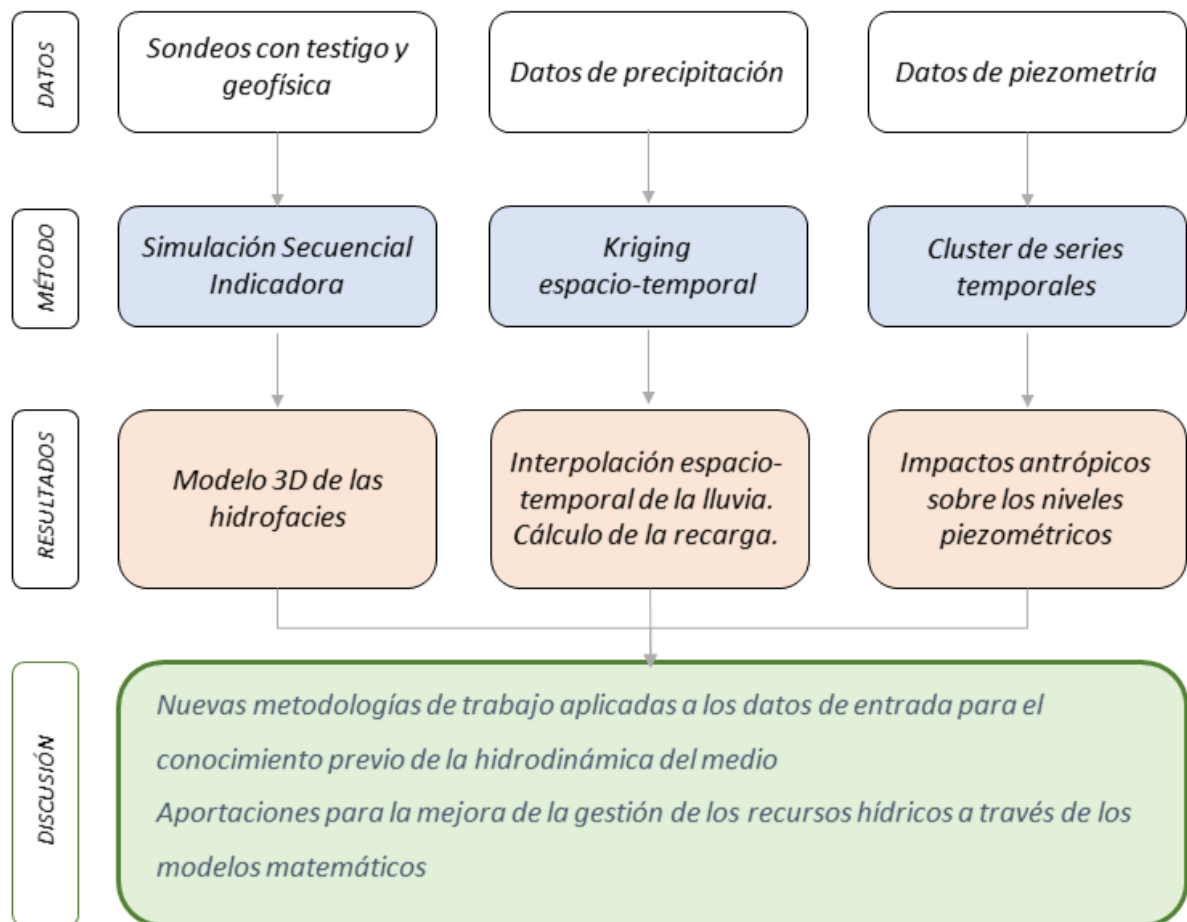


Figura 10. Esquema metodológico seguido en la presente tesis doctoral.

Capítulo 4: Artículos

Esta tesis doctoral consta de tres publicaciones científicas realizadas en base a los objetivos indicados en el apartado anterior.

El formato de presentación elegido para la tesis doctoral es el de compendio de artículos científicos. Cada uno de los artículos presenta su resumen, introducción, metodología, resultados, discusión, conclusiones y referencias. Ninguno de los artículos incluidos en esta tesis doctoral ha sido presentado previamente por otra persona en otra tesis.

- Artículo I.** Applying 3D stochastic simulation to improve the hydrogeological modeling of heterogeneous aquifers. Case of Doñana (SW Spain).

- Artículo II.** Relevance of spatio-temporal rainfall variability regarding groundwater management challenges under global change. Case study in Doñana (SW Spain).

- Artículo III.** Clustering groundwater level time series of the exploited Almonte-Marismas aquifer in Southwest Spain.

Artículo I

Con el permiso de los coautores, esta sección reproduce el contenido del artículo:

Naranjo-Fernández, N., Guardiola-Albert, C., Montero-González, E. (2019) Applying 3D stochastic simulation to improve the hydrogeological modeling of heterogeneous aquifers. Case of Doñana (SW Spain). *Water*, 11, 39; DOI: 10.3390/w11010039.

Applying 3D stochastic simulation to improve the hydrogeological modeling of heterogeneous aquifers. Case of Doñana (SW Spain)

Simulación estocástica 3D para la mejora de la modelización hidrogeológica de acuíferos heterogéneos. El caso de Doñana (SW España)

Nuria Naranjo Fernández¹, Carolina Guardiola Albert² y Esperanza Montero González¹

1 Facultad de Ciencias Geológicas. Universidad Complutense de Madrid. Calle José Antonio Novais 12, 28040, Madrid; nuria.naranjo.phd@gmail.com; emontero@ucm.es;

2 Instituto Geológico y Minero de España. Calle Ríos Rosas - 23 - 28003 Madrid – España. c.guardiola@igme.es

* Corresponding author: c.guardiola@igme.es

Abstract

Mathematical groundwater modelling with homogeneous permeability zones has been used for decades to manage water resources in the Almonte-Marismas aquifer (southwest Spain). This is a highly heterogeneous detrital aquifer which supports valuable ecological systems in the Doñana National Park. The present study demonstrates that it is possible to better characterize this heterogeneity by numerical discretization of the geophysical and lithological data available. We identified six hydrofacies whose spatial characteristics were quantified with indicator variogram modelling. Sequential Indicator Simulation then made it possible to construct a 3D geological model. Finally, this detailed model was included in MODFLOW through the Model Muse interface. This final process is still a challenge due to the difficulty of downscaling to a handy numerical modelling scale. New piezometric surfaces and water budgets were obtained. The classical model with zones and the model with 3D simulation were compared to confirm that, for management purposes, the effort of improving the geological heterogeneities is worthwhile. This paper also highlights the relevance of including subsurface heterogeneities within a real groundwater management model in the present global change scenario.

Keywords: geostatistics; hydrofacies; hydrogeological model; stochastic simulation; water balance.

Resumen

Los modelos matemáticos hidrogeológicos con zonas de permeabilidad homogénea han sido utilizados por muchas décadas para la gestión de las aguas del acuífero Almonte-Marismas (suroeste de España). Es un acuífero detrítico con elevada heterogeneidad que mantiene sistemas ecológicos de elevado valor del Parque Nacional de Doñana. El presente estudio demuestra que es posible caracterizar la heterogeneidad mediante la discretización numérica de los datos geológicos y geofísicos existentes. Se han identificado seis hidrofacies cuyas características espaciales han sido cuantificadas mediante la modelación de variogramas indicadores. La aplicación de la Simulación Secuencial Indicadora ha hecho posible construir un modelo geológico 3D. Finalmente, el modelo ha sido introducido en MODFLOW a través de la interfaz Model Muse. Este proceso final es todavía un reto debido a la dificultad de reducir el modelo a una escala numérica manejable. Han sido obtenidas nuevas superficies piezométricas y nuevos balances hidrológicos. El modelo clásico por zonas y el modelo con la simulación 3D han sido comparados para confirmar que, en términos de gestión, el esfuerzo de mejorar el conocimiento de las heterogeneidades geológicas merece la pena. En este artículo, además, se destaca la importancia de incluir las heterogeneidades subsuperficiales en los modelos hidrogeológicos de gestión bajo el escenario de cambio global.

Palabras clave: geoestadística; hidrofacies; modelos hidrogeológicos; simulación estocástica; balance hidrológico.

1. Introduction

The Almonte-Marismas sedimentary aquifer supports both Doñana National and Natural Parks. It is a site of great ecological wealth, thanks to the quantity and quality of available water. This fact makes it a vulnerable area sensitive to any climatic change (Guardiola-Albert and Jackson 2011). In addition, it is subject to great environmental pressure due to the elevated groundwater extraction necessary for agriculture and tourism. These withdrawals are causing a decrease in the quality and quantity of groundwater that is incompatible with the health of the ecological systems that characterize this park (Green et al. 2017).

Hydrogeological models enable the evaluation of groundwater changes under future scenarios, so they are a basic support tool for decisions concerning the management of water resources (Sánchez-Vila and Fernández-García 2007). Doñana is not an exception, and groundwater resource management in the Almonte-Marismas aquifer requires the use of numerical models in order to understand how the system works, to estimate available resources, and to simulate the effects of its exploitation. Numerous mathematical models have been built by IGME (the Spanish Geological Survey) since 1975 (IGME 1975; IGME 1982; IGME 1994; IGME 1999; UPC 1999; IGME 2005; IGME 2009; IGME 2011). Throughout the different versions, the model has improved. However, it is difficult to compare the different versions due to their diverse characteristics and results (time steps, time period, study area, initial data, boundary conditions, purpose, etc.).

All the above Almonte-Marismas numerical models represented the spatial variability of hydrogeological parameters (i.e., hydraulic permeability and storage coefficient) through homogeneous zones. The delimitation of these areas was based more on similar piezometric behavior and geological outcrops than on 3D lithological information. However, the spatial variability of subsurface geological material is responsible for flow and solute transport patterns (Jankowski and Beck 2000; Noetinger et al. 2005), and, consequently, more realistic models are required (De Marsily et al. 2005). This is not an easy task because of the laboriousness of getting a digitalized 3D geological model and integrating it into the numerical groundwater model. The degree of difficulty of

constructing a real 3D groundwater model also depends on the spatial scale of work and the specific objective of the study (Watson et al. 2015). An additional challenge in groundwater flow modelling is to build a congruent model linking geological properties with hydrogeological dynamics. To mitigate this difficulty, several authors, including Klingbeil et al. (Klingbeil et al. 1991) and Anderson (Anderson 1989), propose to regroup lithologies into units called hydrofacies, characterized by similar hydrogeological properties (Comunian et al, 2016).

Geostatistics is used extensively to estimate the value of a given attribute at a specific location by preserving the spatial distribution and quantifying the uncertainty of the estimates (Seifert and Jensen 1991; Renard 2007). Although geostatistical facies modeling for petroleum applications is a widely used practice (Pyrzcz and Deutsch 2014), it is not so common in groundwater applications (Fiori et al. 2016). Deterministic or smooth interpolated models of alluvial lithofacies often preclude the scale of heterogeneity, which can be much finer than characteristic inter-borehole spacing. Geostatistical tools, like variogram models and simulations, allow the quantification and modeling of the geological heterogeneity in depositional systems (Dell’Arciprete et al. 2012). There are abundant geostatistical estimation and simulation methods: simple and ordinary kriging, sequential indicator simulation (SIS; Alabert 1987), truncated Gaussian simulation (TGS; Beucher and Renard 2016), object-based simulation (OBS; Deutsch and Tran 2002), multiple point statistics (MPS; Hu and Chugunova 2008) etc. In the SIS method, hydrofacies are transformed as indicator variables (1 if present and 0 if not present), kriged individually using variograms, and assigned to cells weighting kriging results; finally, a simulated 3D distribution is created with a specific probability distribution function (Gómez-Hernández and Srivastava 1990). SIS is preferred in situations where the geometry of geobodies is complex and when the density of data allows for accurate variogram analysis and comes from geological interpretation or geophysical measurements (Deutsch 2006). Also, research by Marini et al. (Marini et al. 2018) found that SIS hydrofacies prediction was better compared with TGS and OBS.

The main goal of this study is to incorporate the spatial variability of geology into the numerical flow model of the Almonte-Marismas aquifer. The objective is twofold: on the one hand we perform a 3D geological characterization of heterogeneity in the Almonte-

Marismas aquifer, using all available geological and geophysical information and applying geostatistical SIS, and, on the other hand, this 3D geological structure and the hydrogeology properties are integrated into the existing numerical flow model. The improvements in the Almonte-Marismas groundwater numerical model, after adding the 3D geological simulation, are analyzed by comparing the results with those of the former model which employs zonification. The evaluation focuses on piezometry modification and water balance differences, paying special attention to groundwater contributions to ecosystems of vital importance in Doñana (e.g., La Rocina Stream, the ecotones, and the marshlands).

2. Materials and Methods

2.1. Study Site

The Almonte-Marismas groundwater system, in southwest Spain, is located in the mouth of the River Guadalquivir (Figure 1). It covers 2409 km² of a sedimentary basin composed of basal sands, silts, Aeolian sands, and marshlands materials at the surface. The principal hydrogeological limits are the Guadalquivir to the east, the Atlantic Ocean to the south, the River Tinto to the west, and impervious geological materials to the north. The north, west, and southwest sectors of the aquifer have unconfined behavior that becomes confined from west to southeast by the Doñana marshlands (Figure 1; Salvany and Custodio 1995). The principal groundwater flow runs from northwest to southeast, recharging mainly in northern and southwestern areas. Groundwater and surface water in the Almonte-Marismas aquifer flood some ponds and marshlands that represent one of the most important ecological areas in Spain today (Custodio et al. 2009).

Since the end of the 19th century, the Doñana marshlands have suffered a series of anthropogenic actions that have significantly changed the natural environment (Green et al. 2017). The extent of the marshland decreased from 140,000 ha at the end of the 19th century to 30,000 ha that remain semi-virgin today (Rodríguez-Ramírez et al. 2005).

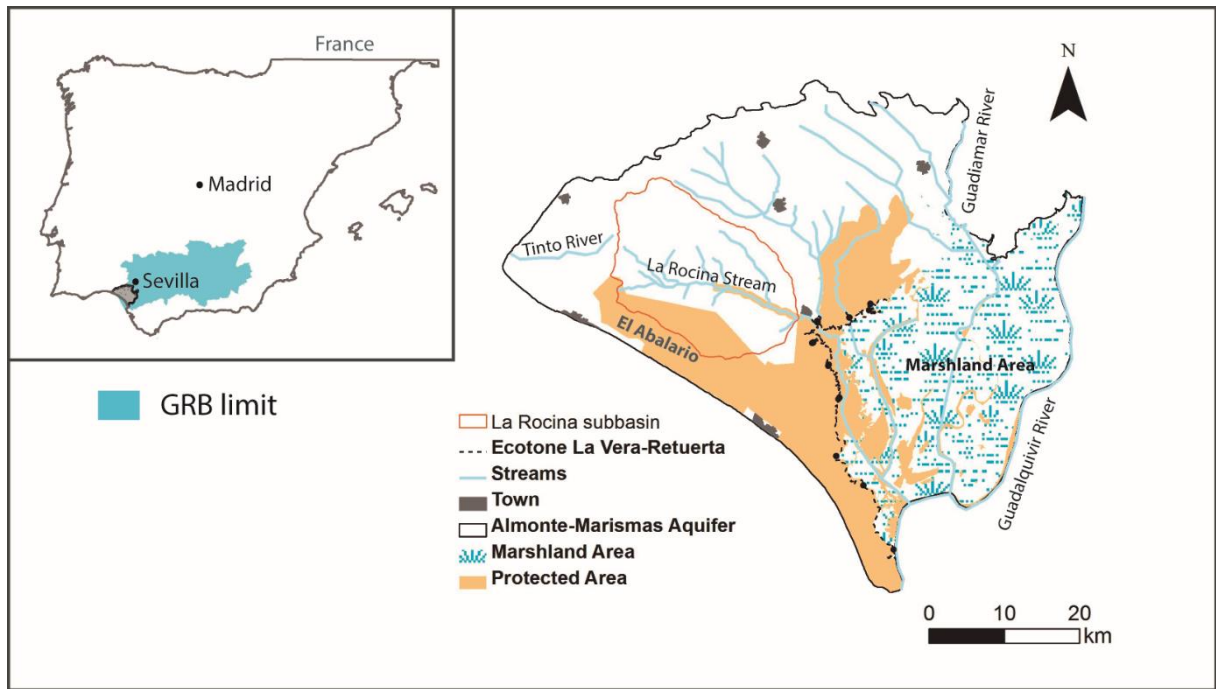


Figure 1. Location of the Almonte-Marismas Aquifer and locations studied. The Guadalquivir River Basin (GRB) limit is also shown.

2.2. Hydrofacies Model

The maximum total thickness of the system is 425 m with maximum top and bottom elevations of 190 masl and 235 mbsl, respectively. Based on new boreholes, Salvany et al. (Salvany et al. 2011) redefined the sedimentological model of the Almonte-Marismas aquifer with nine lithologies or formations. In the present study we took Salvany's work and the updated geological map (GEODE) to define six hydrofacies (Figure 2). The six hydrogeological behaviors correspond to sediments of ages ranging from Neogene (Upper Miocene) to Quaternary. Between the Upper Miocene and the end of the Pliocene, the sediments are of marine origin, and from then until the Quaternary, they are of continental origin (Salvany et al. 2011). The numbering of hydrofacies did not follow a lithostratigraphic criterion. Figure 2 shows how hydrofacies from 1 to 4 emerge on the surface. The other hydrofacies, 5 and 6, do not appear in the geological map and constitute the semiconfined/confined sector of the aquifer.

Hydrofacies 6 corresponds to the Almonte sand unit described by Salvany et al. (Salvany et al. 2011), the most transmissive and deepest lithology. These sands are below the marsh and have a thickness between 10 and 110 m, which increases in a southerly

direction. Hydrofacies 5 corresponds to the Lebrija unit (Lower unit). With a maximum thickness of 120 m, it appears discontinuously as clay, sand, and gravel lens. Sometimes Hydrofacies 5 appears over Hydrofacies 6, and at other times it makes contact with the aquifer bottom (Figure 2). The marshlands are represented by Hydrofacies 4, formed by fine and medium plastic clays with a lot of organic matter and characterized by very low permeability. It behaves as a very slow aquitard. They are between 20 and 80 m thick, which is still increasing due to sediment transport. Hydrofacies 3 is formed by sands and corresponds to the Lebrija unit (Upper unit; Salvany et al. 2011). Hydrofacies 2 corresponds to the outcrop of silt at the surface that can be found in the geological map (Figure 2) and has little presence in the 3D hydrogeological model. Lastly, Hydrofacies 1 corresponds to the sands of the Abalarío and the Aeolian units that emerge on the surface in the geological map (Figure 2). They constitute the higher-recharge-rate area in the free aquifer sector. They also have mobile dunes and can reach a thickness of up to 150 m (Salvany et al. 2011).

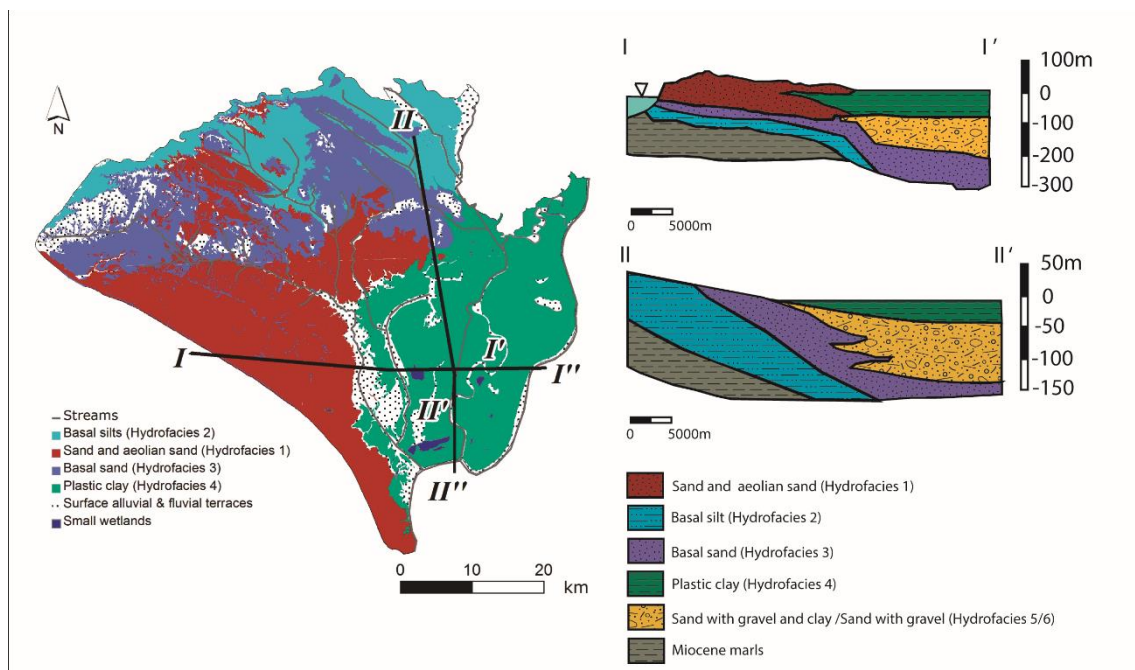


Figure 2. *Left:* Modified geological map (based on [35]). *Right:* Schematic of geological cross sections based on [6] and [32].

2.3. Geological and Geophysical Data

Geological and geophysical data were obtained from IGME (IGME 2007; IGME 2009). There are a considerable number of records with raw geological and geophysical

information (Figure 3): 460 well logs (more than 16,000 m) were performed between 1967 and 2009 and 895 vertical electrical soundings (VES) between 1967 and 1995, four seismic reflexion profiles of 7.8 km were made in 2007, and two different groups of reflexion seismic lines were performed in 1990 and 2002 (IGME 2007).

The biggest challenge is to extract facies information manually from such a huge number of well logs, with subjective and heterogeneous interpretations, and meter by meter. In addition, the arduous geophysical data recompilation carried out by the IGME in 2007 was carefully reviewed, providing a large quantity of data that had never been taken into account for the construction of the previous models. All raw data were analyzed and digitized in order to transfer the facies information to the corresponding hydrofacies. The resulting 3D hydrofacies values were formatted to be compatible with the geostatistics tools described in the following subsections.

2.4. Variogram Modelling

The experimental variogram, defined within the framework of a second-order stationary hypothesis, can provide an empirical description of the spatial continuity of one variable (Matheron 1965). The experimental variogram analysis gives information about the range, the sill, and the anisotropy coefficients that are necessary to determine the continuity, periodicity, and trends of the hydrofacies units (Trevisani and Fabbri 2010). Indicator (i.e., categorical) variables were used to represent the hydrofacies units for a given location within 2 m depth intervals and 500 X 500 m cells. The coding of hydrofacies units 1, 2, 3, 4, 5, and 6 to indicator variables (I) at any given point in space (u_α) was assigned by defining

$$I(u_\alpha, k) = \begin{cases} 1 & \text{if hydrofacie } h \text{ prevails a location } u_\alpha \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $k= 1, 2, 3, 4, 5, 6$.

Hydrofacies proportions, which are significant parameters when an indicator-based method is used (Seifert and Jensen 2000), were determined from hydrofacies digitalized values.

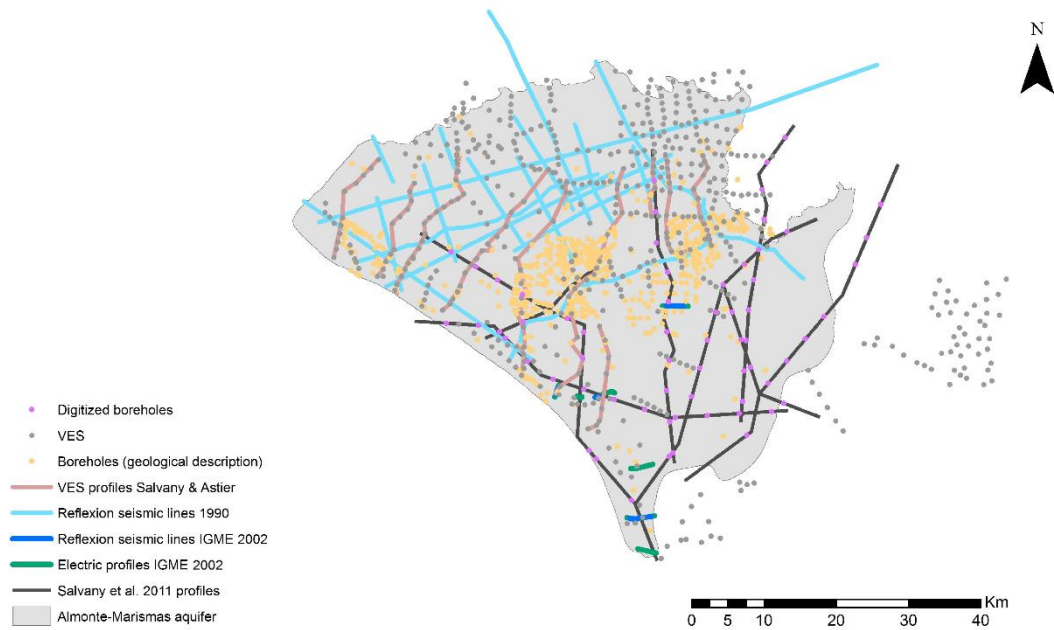


Figure 3. Spatial distribution of geological and geophysical information in Almonte-Marismas aquifer.

In the present study, 3D experimental variograms for each hydrofacies were computed and fitted by using Isatis software (Géovariances 2009). In comparison with other similar software as SGeMS (Remi et al. 2009), Isatis has an easier and faster interface to model and interpret the 3D variograms.

2.5. Sequential Indicator Simulation (SIS)

Geostatistical simulation methods generate nonunique results and overcome the smoothing effect problem (Caers 2000). Conditional geostatistical simulation application generates a set of possible equiprobable results that are conditioned to the available data, while having global statistical behaviors very close to those of reality, such as variograms, reproducing the fine variations. The justification for the use of SIS in this case is its easy implementation, the high level of control of the statistical variables that is exercised through the variograms, and the good geological resolution that is obtained for different regional scales (Deutsch 1998; Deutsch and Journel 1998; Goovaerts 1996; Soares 1998; Almeida 2010). SIS was applied to construct one simulation using a random seed value, 20 for the maximum number of conditioning values, and 30,000 m–30,000 m–500 m for the search ellipsoid value. The computation

time was 80 min using SGeMS software (Remi et al. 2009). A categorical transformation algorithm (TRANSCAT) was then run, again with SGeMS. This transformation is necessary to clean the small-scale variations and to match the target proportions while preserving the local statistics (Luckjan and Chalermyanont 2017). This last step is sometimes ignored by users, but it is necessary to obtain a smooth 3D model. For interactive 3D visualizations, Isatis software (Géovariances 2009) was used because it has a comfortable interface, it is easy to move the image, and it explores the 3D solids. However, Isatis does not allow SIS to be executed with several categorical variables at the same time.

2.6. Groundwater Modelling with MODFLOW and Model Muse GUI

The simulation of groundwater flow in the Almonte-Marismas aquifer was done with the widely known 3D finite difference model MODFLOW-2000 (Harbaugh 2005). The graphical user interface used to run MODFLOW was Model Muse software (Winston 1999), a free software package available for future applications, water management administrations, and other potential users. The model domain is oriented north–south and discretized into 500 X 500 m grid cells. These dimensions allowed the highest possible model resolution based on the computational limits, available data, and problem scale.

The basis of a good hydrogeological model begins with its robust results in steady-state. The Almonte-Marismas aquifer was modeled under steady-state conditions in order to compare (i) the new model, with heterogeneous hydrofacies read from the 3D geostatistical simulation and with appropriate permeability assigned to simulated lithofacies, and (ii) a model similar to the ones built before with homogeneous zones for both hydrogeological parameters. The new model has seven layers while the one similar to former models has two layers. Both models were calibrated independently to reproduce the piezometry observed in the 1970s, assuming that this period can correspond to a stationary state. The parameters calibrated were permeability values for each zone or hydrofacies unit (Table 1).

Table 1. Calibrated permeability values for each hydrofacies unit (m/ day).

Hydrofacies Code	Lithology	Kx = Ky	Kz
1	Dunes and aeolic sand	1.6	1.6
2	Silt basin	0.15	0.015
3	Sand basin	0.7	0.07
4	Plastic clay, wetlands	1.67×10^{-3}	1.67×10^{-5}
5	Clay, sand, and gravel	0.3	0.03
6	Sand and gravel	10	1

To characterize the impervious bottom of the aquifer using the available geological and geophysical data, kriging with an external drift trend of the aquifer thickness was computed (Wackernagel 1998; Chiles and Delfiner 1999). The total aquifer thickness was measured from the digital terrain model and the appearance of the low-permeability marls lithology at the bottom.

The boundary conditions were similar to those used in previous models (Guardiola-Albert and Jackson 2011; IGME 2011) and were selected to most closely match the physical system (Harbaugh 2005). North lateral flows were modeled with the General Head-Boundary package—water entering from the north border of the marshlands and draining in the rest of the northern limit. The River Guadalquivir contact is impervious, the Guadiamar and Tinto were modeled with the River Package, and the rest of the streams network was modeled with the Stream package. Atlantic Ocean contact was approached as cells of constant piezometric height without taking into account the effect of the tides, while also placing drain-type cells in the areas where there are springs in the coastal zone. The entire marshlands area was simulated with impermeable cells in Layer 1 for the zonification model. With the introduction of the hydrofacies model, the marshlands cells were activated. This decision could be taken as being due to the possibility that the 3D simulation can better define the interchanged gravel lens. Marshland cells became a zone of low hydraulic permeability (1.67×10^{-3} m/day). The phreatic evapotranspiration package simulates the areas of eucalyptus and pine trees. To implement the recharge, the aquifer was divided into 15 recharge subzones within which the infiltration process was assumed to be homogeneous (Custodio et al. 2009).

To evaluate both models' piezometry and water balance, differences were analyzed in the whole aquifer and in three ecosystems of special importance: La Rocina Stream, ecotones, and groundwater masses below the marshlands (Figure 1). Figure 4 shows the whole approach used in this study.

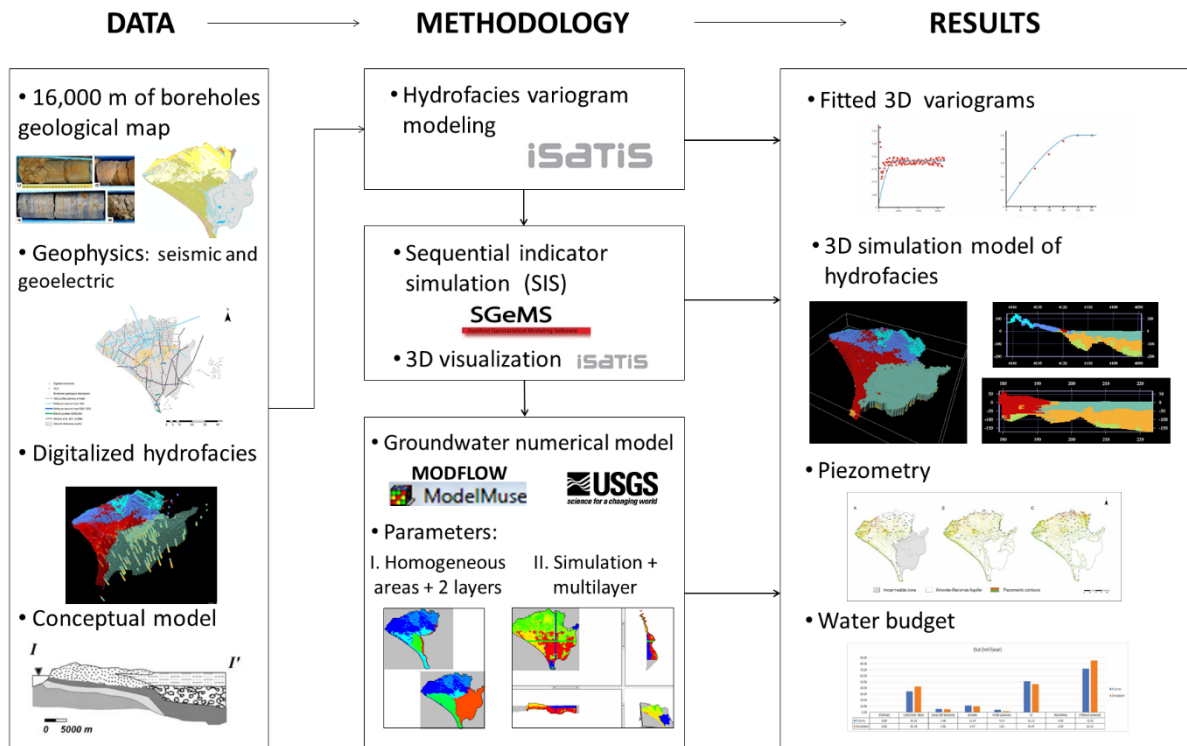


Figure 4. Flowchart of the approach used in the current study.

3. Results

3.1. Geostatistical Model

3.1.1. Variograms

Figure 5 and Table 2 present the numerical and graphical parameters of the variogram models. Except for the variogram models for marshland clays (Hydrofacies 4), which are Gaussian and power, the models are spherical or exponential. The vertical range is from almost one or two orders of magnitude lower than the horizontals, except in Hydrofacies 6 (sand and gravels). Eleven variogram models were obtained (Figure 4) with variations depending on the origin of the sediments.

Table 2. Experimental variogram model parameters. Hydrofacies codes: 1. Aeolian and dune sands; 2. Silt basin; 3. Sand basin; 4. Marshlands clay; 5. Sand and gravels with clay; 6. Sand and gravels.

Hydrofacies code	Direction	Model	Nugget effect	Sill	Range
1	Horizontal 0°	Spherical	0	0.09	150
1	Vertical	Spherical	0	0.01	47
2	Horizontal (omnidirectional)	Spherical	0	0.0045	15000
3	Horizontal 0°	Spherical	0	0.014	1500
3	Vertical	Exponential	0	0.05	900
4	Horizontal 45°	Exponential	0	0.06	7000
4	Horizontal 135°	Power	0	0.2	15000
4	Vertical	Gaussian	0	0.29	175
5	Horizontal 0°	Exponential	0	0.15	2400
5	Vertical	Exponential	0	0.13	800
6	Horizontal 0°	Exponential	0	0.1	500
6	Vertical	Spherical	0	0.4	300

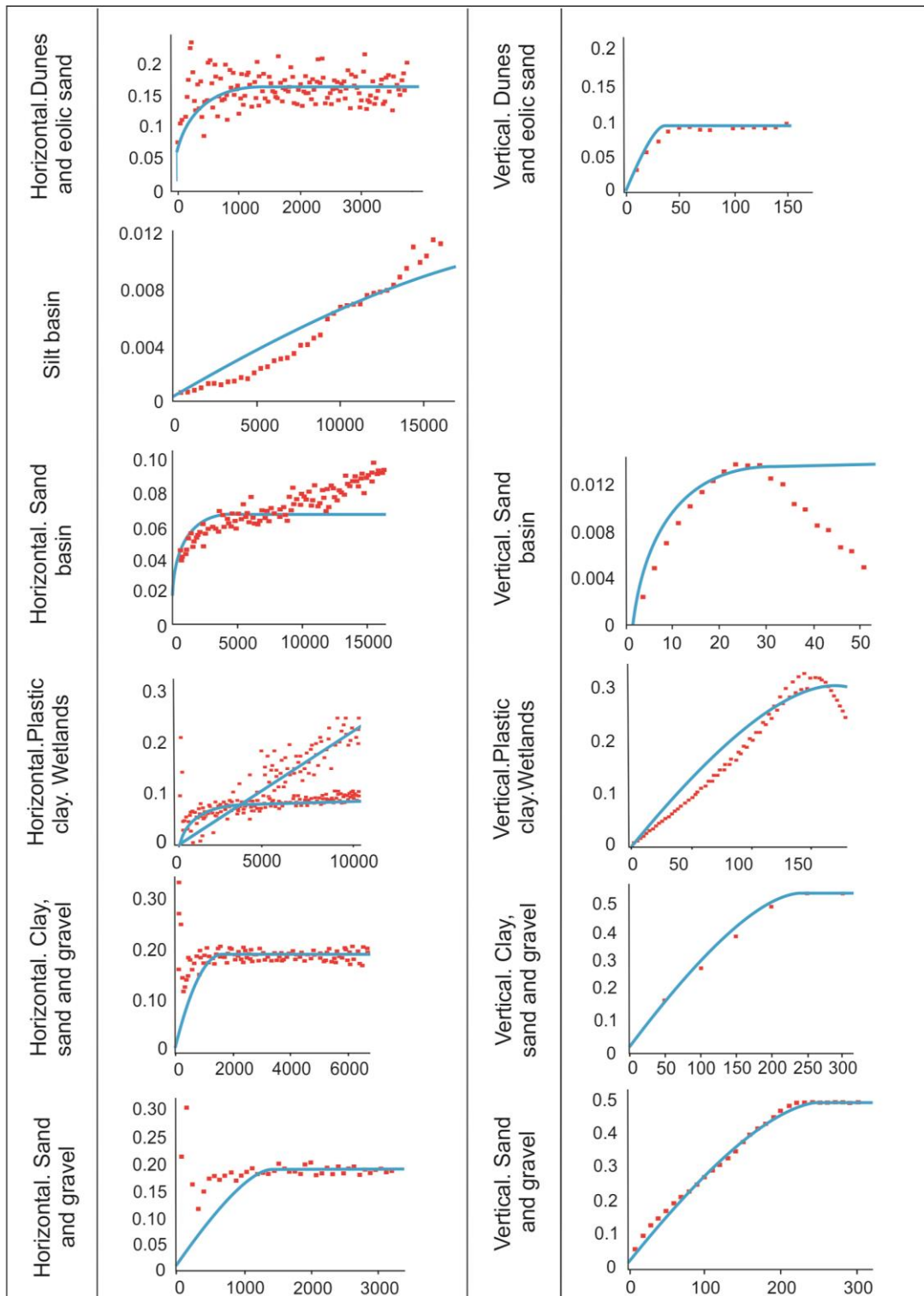


Figure 5. Modeled variograms (horizontal and vertical) of hydrofacies.

3.1.2. Sequential Indicator Simulation

Comparing the results with the interpretations by Salvany (Salvany et al. 2010; Salvany et al. 2011), the simulation was validated by both hydrofacies proportions and facies distribution (Figure 6). The proportions were 25% (Hydrofacies 1, Aeolian and dune

sands), 2% (Hydrofacies 2, silt basin), 12% (Hydrofacies 3, sand basin), 15% (Hydrofacies 4, marshlands clays), 24% (Hydrofacies 5, clay, sand and gravel), and 22% (Hydrofacies 6, sand and gravel).

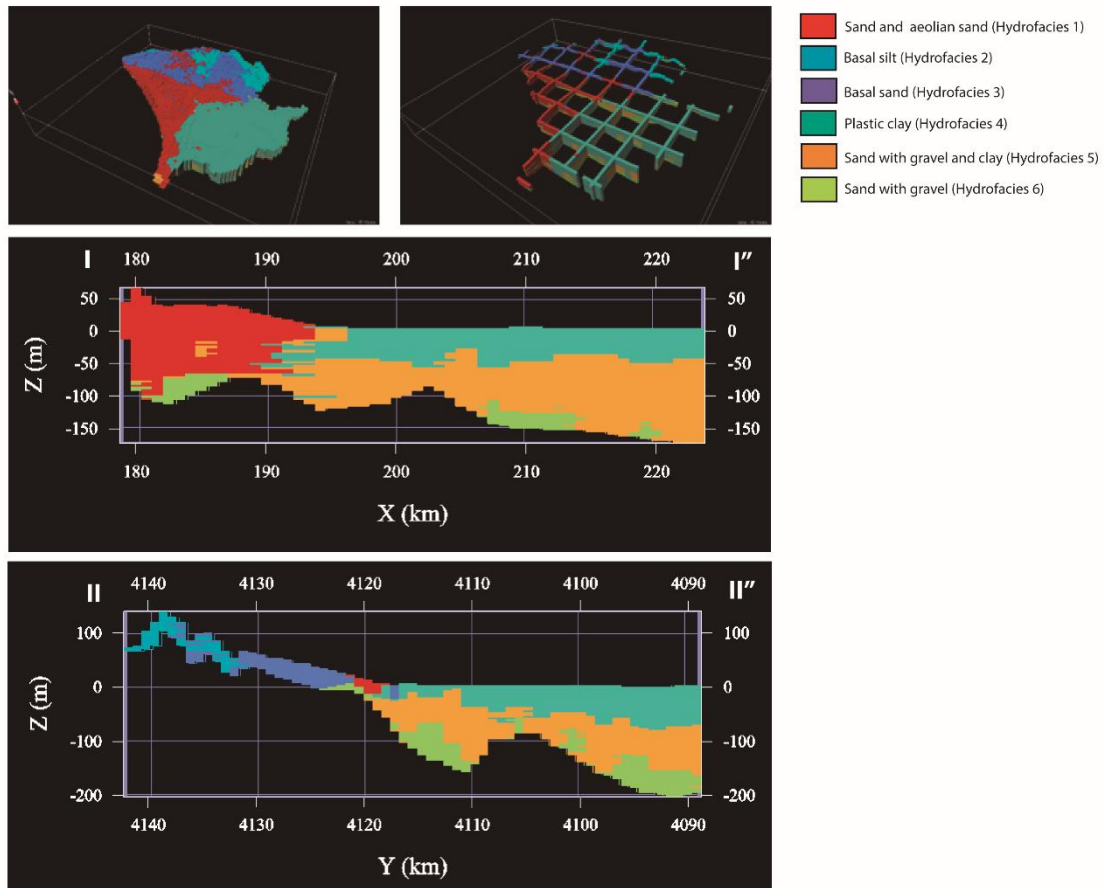


Figure 6. 3D views of SIS results with Isatis 3D Viewer and cross sections I-I'' and II-II''.

3.2. Groundwater Flow Model

3.2.1. Piezometry

Figure 7 shows piezometric contours for the groundwater model with the hydraulic permeability and storage coefficient approximated by zonification (Plot A for the upper layer and Plot B for the lower layer). Plot C in Figure 7 presents the piezometry for the groundwater model with the 3D geostatistical simulation, registered in Layer 6. Both models have similar flow patterns. The model with homogeneous zones for hydrogeologic parameters (Plots A and B in Figure 7) registers lower piezometric levels and a higher contour slope in the northeast zone (Plot C in Figure 7). We implemented the 3D simulation by changing the boundary conditions and activating the cells located

in the marshlands (gray area in Figure 7A), this area having the deepest isopiezometric values. Plot D in Figure 7 presents the differences between the piezometry represented in Plots B and C. In Plot D it can be seen that marshlands have similar piezometric values. Except for in the northern part of the aquifer and the marshlands, positive differences seem to be more abundant (Plot D in Figure 7), the piezometry in these areas being lower for the model with 3D simulation.

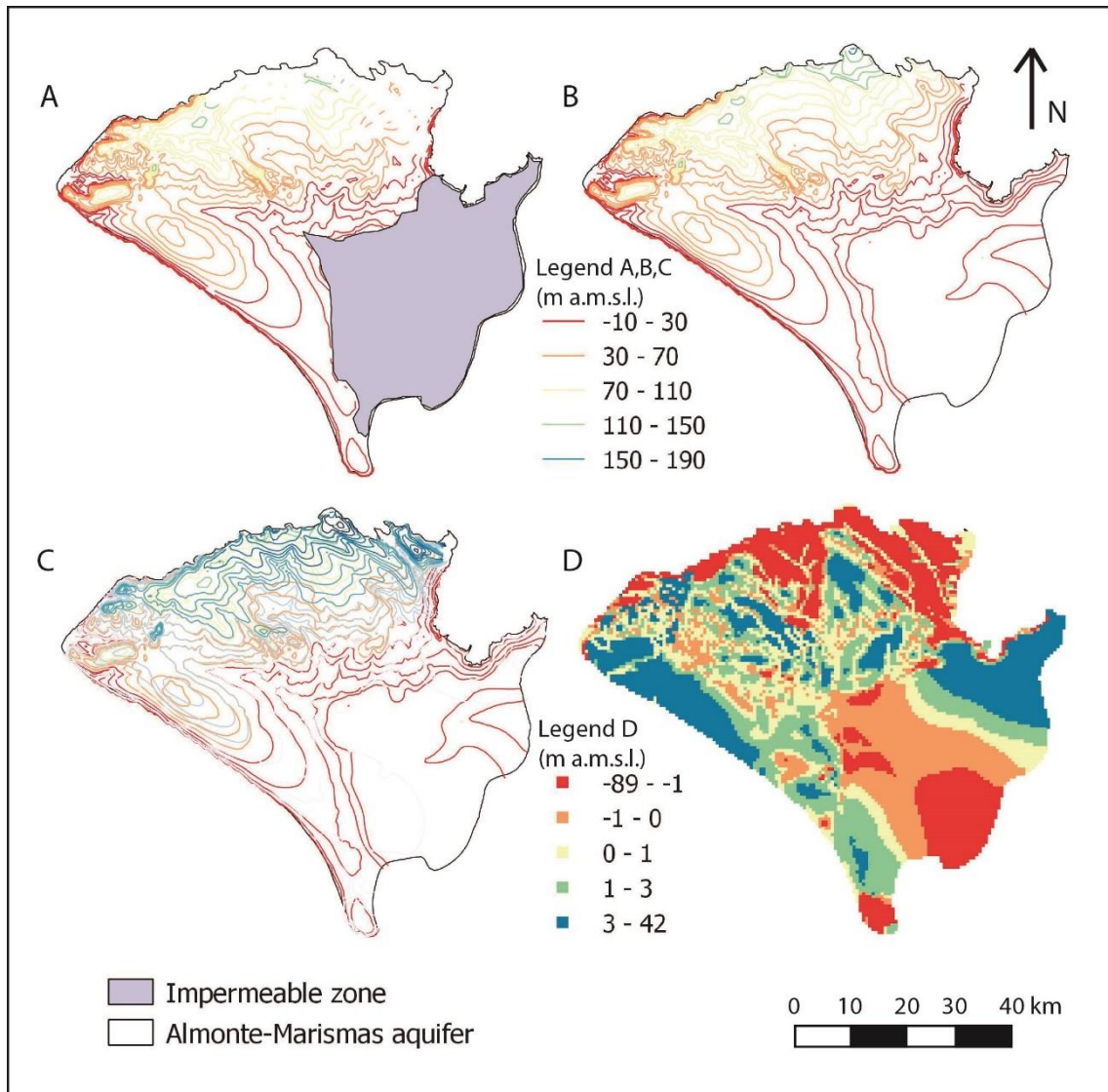


Figure 7. Piezometric contours of groundwater models: (A) first layer of the original model; (B) second layer of the original model; (C) model with hydrofacies simulation; (D) difference between (B) and (C).

3.2.2. Water Budget

Figure 8 shows the input and output budget values obtained by running both models (i.e., the permeability and storage coefficient entered with zonification or with the 3D simulation). This budget indicates that, when integrating the hydrofacies simulation, inputs from river and stream leakage increase by 3 and 10 hm³/year, respectively; while they descend for the general head boundaries by around 1 hm³/year. On the other hand, outputs through constant head and stream leakage boundaries rise with the geostatistical simulation by 8 and 13 hm³/year, respectively. Output terms decreasing in the model with higher heterogeneities are drains (1 hm³/year less), river leakage (reduction of 2 hm³/year), evapotranspiration (decrease of 5 hm³/year), and general head boundaries (decrease of 0.5 hm³/year). The highest differences are found in stream leakage for both input and output terms, while, as expected, the recharge does not change.

Figure 9 summarizes the interchange of fluxes between the Almonte-Marismas aquifer and some ecosystems of vital importance in Doñana (e.g., La Rocina Stream, the ecotones, and the groundwater masses under the marshlands). With respect to the former zoned model, introducing the geostatistical simulation to represent the geological heterogeneities considerably modifies the interchange of annual fluxes in these zones (Figure 9). These variations will be analyzed in Section 4.

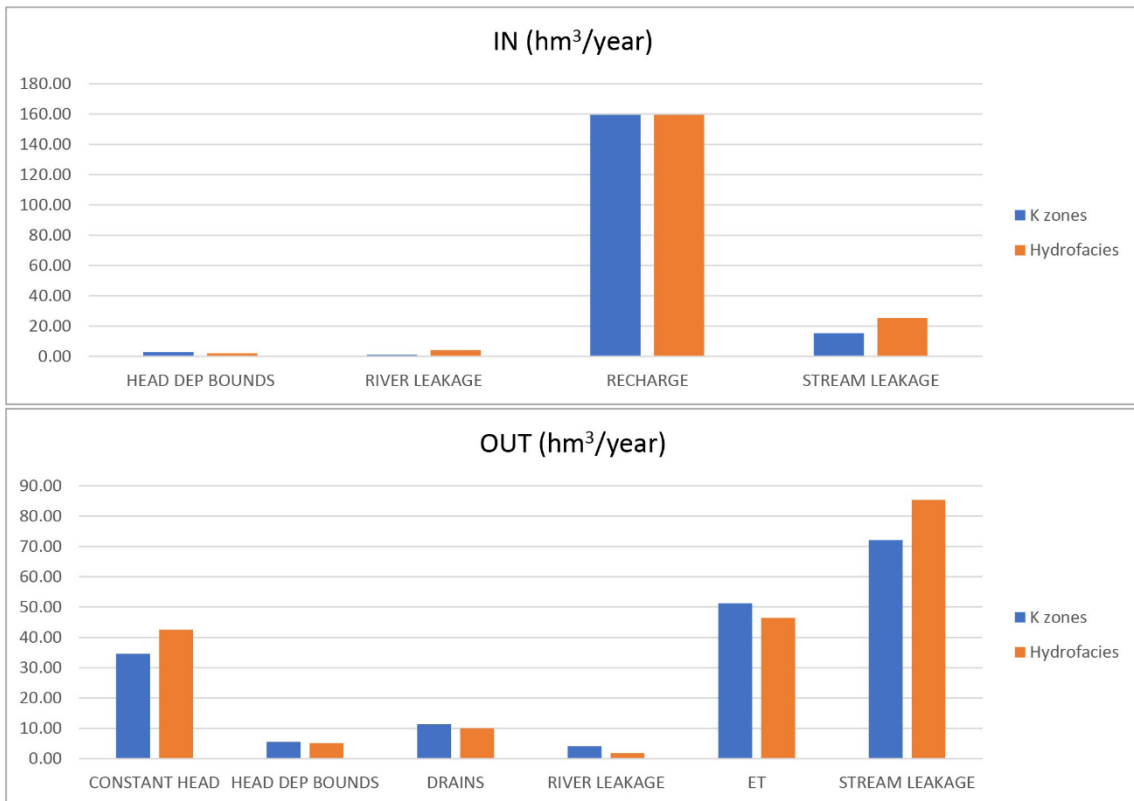


Figure 8. In and out water budget (hm³/year) comparison between the zoned model and the model with integrated 3D simulation.

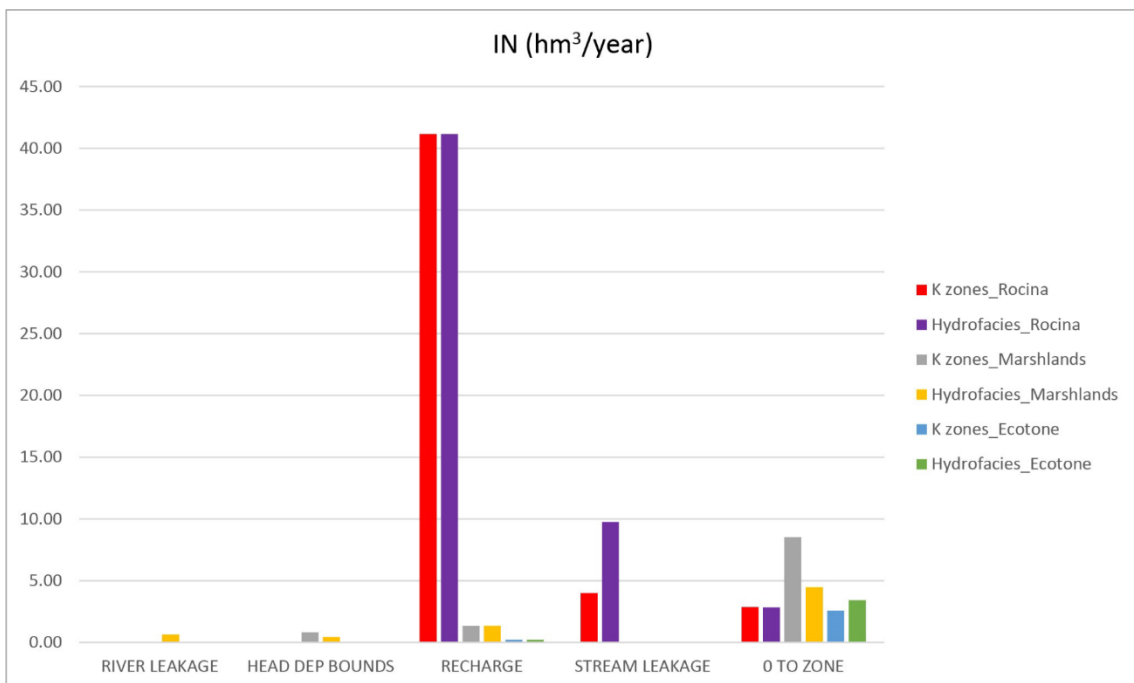


Figure 9. Cont.

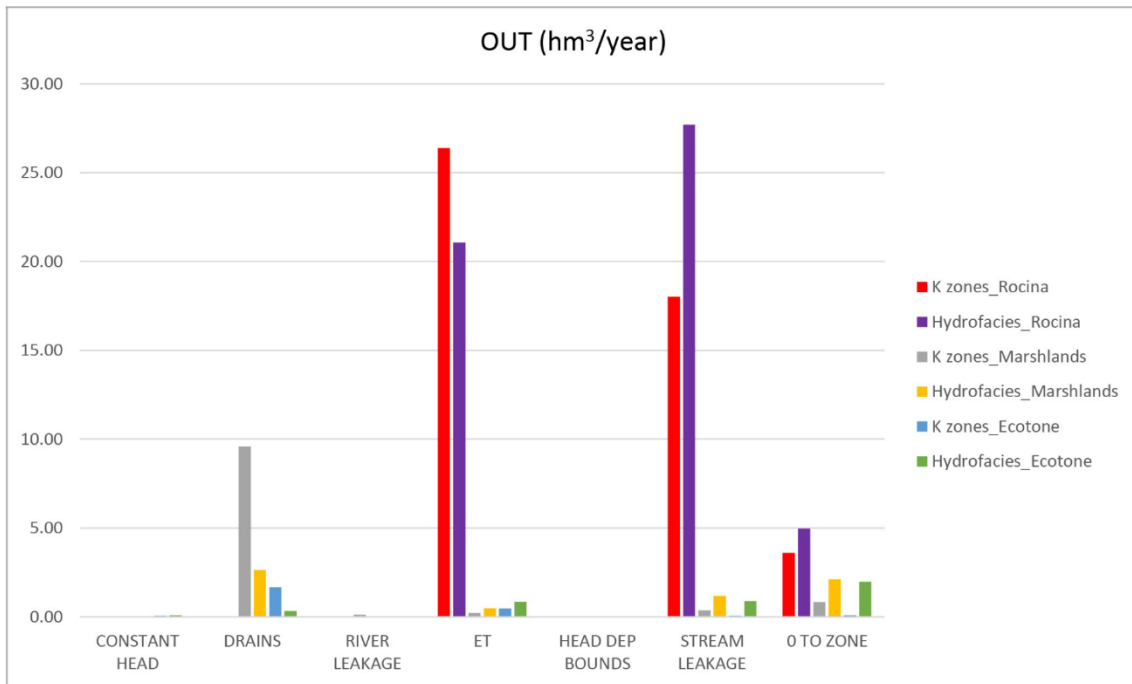


Figure 9. In and out water budget ($hm^3/year$) comparison between the original model (k_zones) and model with integrated simulation in different specific zones.

4. Discussion

4.1. Variogram Models

In five of the six variograms computed, the correlation range observed depends on the direction. This is explained by the lateral extent of the deposits, which is usually greater than their thickness; hence, the vertical range is much lower than the horizontal range (Gringarten and Deutsch 1999). The variogram of Hydrofacies 2, representing basal silt that emerges in a local zone at the northeast of the aquifer, is the only one that presents no anisotropy. The homogenous distribution of this kind of deposit explains this spatial behavior.

Hydrofacies 1 presents a highly continuous spatial correlation, indicated by the spherical model fitted to its experimental variogram. This fact is coherent with the soft Aeolian dune sands and ancient dunes of this hydrofacies. In the horizontal direction, the range is too small (150 m) in comparison to those of the other lithologies (greater than 500 m). The same occurs with the vertical directions, the range for Hydrofacies 1 being 47

m, while for the rest of the hydrofacies it is more than 175 m. These smaller range values can be explained by the low extent of Aeolian deposition processes.

Basal sands in the central and northern zone of the aquifer compose Hydrofacies 3. Its variogram ranges have medium values with the same order of magnitude for both the horizontal (1500 m) and vertical (900 m) directions. These values may correspond to the dynamics of these sand deposits, which are cemented and more ancient sands than Hydrofacies 1. In addition, Hydrofacies 3 has a slight horizontal trend.

This trend appears more clearly in Hydrofacies 2 (i.e., basal silt), with a sharper increase in the variogram values. The appearance of trends in sedimentary deposits indicates changes in some of the characteristics of the hydrofacies from the proximal to the distal part in depositional environments. These variations cause an increase beyond the threshold variance (Gringarten and Deutsch 1999). In these cases of nonstationary behavior, Gringarten and Deutsch (Gringarten and Deutsch 2001) recommend cleaning residual data in the original variogram to remove this kind of trend. However, in the present study this removal process was not performed in order to facilitate the simulation with SGeMS software. The elimination of the trend could be included in the future, constituting a potential improvement of this work. Equally for Hydrofacies 2 and 3, there is a distinctive smooth spatial correlation structure of earth deposits revealed by the spherical and exponential models used for the variogram fitting.

The plastic clays of Hydrofacies 4 arrive through the streams that flood the marshlands. They have different variogram ranges depending on the variogram orientation: 7000 and 15,000 m, respectively, for 45° and 135° , and 175 m in the z axis. The presence of a low range in vertical behavior, it being higher in the horizontal direction, is attributable to the considerable thickness of the hydrofacies. This 3D anisotropy may indicate a variation in the depositional direction associated with different depositional steps (Johnson and Dreiss 1989). The variety of sedimentological origins is also reflected by the diverse types of variogram model (i.e., exponential, power, and Gaussian models). The Gaussian model fit can be explained by the fact that it is a very continuous lithology on the surface. The vertical anisotropy and intermediate correlation lengths (i.e., 2400 m for horizontal and 800 m for vertical) of the variogram for Hydrofacies 5 are explained by the elongated and discontinuous lens interleaved with Hydrofacies 6. The

exponential models used to fit the corresponding experimental variograms show the continuity of these deposits.

The sixth and last hydrofacies presents quite small ranges (i.e., 500 m for horizontal directions and 300 m for vertical directions). These dimensions correspond very well to sands and gravels with a deltaic–alluvial origin. The exponential and spherical models agree on the considerable spatial continuity of the thickest hydrofacies in the area.

4.2. Geostatistical Simulation

Hydrofacies proportions quantify the hydrogeological behavior of the Almonte-Marismas aquifer. The conceptual model assumes that the largest amount of recharge occurs through sands and Aeolian sands. In total, 26% of the measured and simulated hydrofacies corresponds to these Aeolian sands, corroborating their major hydrogeological role in the aquifer. Hydrofacies 5 and 6 have a high groundwater storage power, supported by their 36% measured and simulated proportion. The prior geological model gives a conceptual model which is constructed from the surface geology (see profiles in Figure 2). The 3D simulations show improvements to this conceptual model. If we compare Profile II in Figures 2 and 6, the effect of the interpolated impervious aquifer bottom can be appreciated. The aquifer thickness from north to south is much greater in the geostatistical 3D model than in the conceptual profiles. Simulated sand and Aeolian sand hydrofacies (Figure 6) seem to be deeper than in the geological profiles (Figure 2). In Profile I, from west to east, 3D estimated facies (Figure 6) with a substantial gravel component (Hydrofacies 5 and 6) reveal a thickness much lower than that in the conceptual model (Figure 2), reducing these depths to almost half (i.e., from -300 m to -150 m).

Restructuring the sediment system that makes up the Almonte-Marismas aquifer into six hydrofacies gives more geological meaning to the layers used in the numerical model, naturalizing its behavior. The 3D simulation (Figure 6) makes it possible to incorporate two improvements. Firstly, one can include low-permeability cells representing plastic clays interbedded with the sand and gravel lenses of hydrofacies 5 and 6. Secondly, when the computing time allows it, a finer vertical discretization can be introduced,

improving the two-layered modelling. In this stationary case of the Almonte-Marismas model, it was possible to increase to seven layers.

When facies are discretized with indicator variables, it is not possible to use only indicator kriging results to estimate the 3D facies distribution because indicator kriging interpolation is not a direct representation of the hydrofacies, but a probability that these hydrofacies are present in a given pixel. However, raw indicator simulations have small-scale variations and do not match the target proportions. Thanks to TRANSCAT correction (Remy and Boucher 2009), the simulated realization is as smooth as an interpolated 3D field from kriging and follows the measured proportions. Still, this is a simulation result representing an equiprobable 3D image. Other estimates, different from this one, can be obtained through new simulations. Ideally, a high number of hydrofacies simulations could be obtained and input into the groundwater modelling to stochastically analyze the set of numerical results. However, the computing time required for the Almonte-Marismas transient groundwater model (153 min on a system with processor an Intel(R) Core(TM) i7-4750 CPU @ 3.60 GHz, x64-based processor, and 7.98 GB usable RAM), and the required ensemble size to perform an uncertainty analysis, makes this stochastic analysis impractical. In any case, this paper aims to highlight the possibilities of developing real heterogeneous groundwater numerical models (i.e., noncontinuous lithofacies without smooth limits) from geostatistical simulations for water resource management purposes. Hence, we tested the results with just one simulated realization. This is one of the drawbacks of the results presented here, but further research could be done from this starting point: this is the first time that 3D heterogeneous values of permeability have been considered in the highly heterogeneous Almonte-Marismas groundwater numerical model. Future research could also consider the use of MPS (Hu and Chuginova 2008), as recent studies have probed its efficiency in characterizing fluvial deposits (Li et al. 2015).

4.3. Groundwater Flow Model

As the hydrogeologic parameters have different spatial distributions in the two models studied and the number of layers is different, they were calibrated independently from each other. Hence, the aim of the comparison discussed here is not to provide an exact

analysis of the differences between the two models, but to give an idea of the implications for groundwater management decisions made with the results of these numerical models.

Upscaling flow is required to take into account the discrepancy between the scale at which we can characterize the medium with geostatistical simulations and the scale at which we can run the numerical model (Li et al. 2011). The methodology to couple calibration and hydraulic parameter upscaling has been already developed (Li et al. 2015). However, upscaling is not routinely used in real water management models because there is not an implementation in commonly used MODFLOW graphical user interfaces. Upscaling hydraulic permeability in the horizontal directions was not necessary in the present work; the discretization in the SIS is the same as that in the MODFLOW grid. However, for the vertical discretization, MODFLOW is not capable of dealing with the 2 m layers obtained with geostatistics, as dry cells cause frequent numerical instabilities. We have chosen the most abundant hydrofacies in the SIS to assign the hydrofacies to each of the seven layers that form the MODFLOW model. The impact of hydraulic permeability upscaling is greater in transport predictions than in flow and head results (Li et al. 2011). Consequently, 3D hydraulic permeability upscaling (Zhou et al. 2010; Li et al. 2011b) should be computed if future transport modeling is performed with the model presented here.

When comparing the model with homogeneous zones and the updated model with the heterogeneities, the greatest differences in the distribution of isopiezometric contours (Figure 7) are observed in the northern and western areas. In the northeastern zone, the variation is due to the presence of basal silts that modify the hydrogeological properties of the area with respect to the original. In this area the piezometry is higher in the model with 3D simulation, since the properties of the basic sands are introduced in the modeling. This is also reflected in the higher outputs through this general head boundary (Figure 8).

Plot C in Figure 7 shows how marshlands are not impermeable in the new modelling approximation. This allows for slow flow exchange in the marshlands, shown in the extent of the isopiezometric lines, which have less gradient and elevation. When compared with the version of the model with zones, the heterogeneous model drains 7

hm³/year less from the aquifer to the marshlands. This outcome seems to be more in accordance with reality, as the Doñana marshlands are basically fed by direct rainfall on the floodplain and by several watersheds (Serrano et al. 2006). The same happens in the ecotone water budget, with the groundwater contribution decreasing from 1.66 to 0.33 hm³/year. These amounts are more difficult to corroborate as there are no studies estimating outputs from the aquifer to the ecotones.

The significant dissimilarities observed in the budget outputs from the aquifer to the streams and sea (Figure 8) imply that groundwater contributions to these water bodies could be underestimated using homogeneous zones by values of about 18% and 23%, respectively. Studies of potential future water intrusion, more likely to happen in the global climate change framework (Sherif and Singh 1999), can reach quite different conclusions depending on the approximation used to characterize the spatial variability of hydrogeological parameters. The water management consequences of those results are also very important. The ongoing Hydrological Plan (HP; CHG 2015) is the main water management instrument in Doñana. In the HP the actual water reserves are computed considering lateral transfers to the sea and assuming that an optimal ecological status in the present ecosystems can be maintained. Variations of about 20% in these values can lead to considerable changes in the withdrawals allowed from the aquifer. La Rocina Stream is the main permanent tributary to the marshlands and a critical area for water management, with wide areas of berry cultivation irrigated by many illegal extractions (WWF 2016). The net interchange flux between La Rocina Stream and the aquifer is 30% higher in the model, with more spatial variability. Again, this corroborates the importance of using one model or another to evaluate groundwater reserves in Doñana. Moreover, all the deviations found in the present study highlight the need for effective numerical models and uncertainty quantifications in the design of strategies for groundwater resource management in response to future climate change (Guardiola-Albert & Jackson 2011).

5. Conclusions

The main goals of this study were to incorporate the spatial variability of subsurface geology into the numerical flow model of the Almonte-Marismas aquifer and to

determine whether this effort is relevant for management proposes. We were able to achieve these aims by performing a 3D geological characterization of the heterogeneities, analyzing 3D variogram models, and applying geostatistical SIS. We were also able to integrate this 3D hydrofacies structure into a steady-state MODFLOW model. Our findings confirm that water budget results can vary considerably with this new approximation. These changes are key when estimating the discharges from the aquifer to the streams, which are essential to maintaining the riparian ecosystems and the surface contribution to the marshlands. In addition, variations in the water budget estimation affect groundwater exploitation management decisions and studies of marine water intrusion in a global climate change framework. Further studies are needed to facilitate the computing operations needed to implement the 3D hydrofacies in the transient flow model (1975–2018) currently used in Almonte-Marismas aquifer management and to quantify the associated uncertainties in piezometric and water budget results.

Funding: This research was funded by the CLIGRO project (MINECO, CGL2016-77473-C3-1-R) of the Spanish National Plan for Scientific and Technical Research and Innovation and by the Ministry of Education, Youth and Sport of the Community of Madrid with ref. PEJ15/AMB/AI-0218, co-financed under the Youth Employment Operational Program, with financial resources from the Youth Employment Initiative (YEI) and the European Social Fund (ESF).

Acknowledgments: The three anonymous reviewers are thanked for their helpful comments. Spanish Geological Survey is acknowledged for providing the datasets used in this study.

Artículo II

Con el permiso de los coautores, esta sección reproduce el contenido del artículo:

Naranjo-Fernández, N., Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Rodríguez-Rodríguez, M., Fernández-Ayuso, A., Ruíz-Bermudo, F., Montero-González, E. (2020) Relevance of spatio-temporal rainfall variability regarding groundwater management challenges under global change. Case study in Doñana (SW Spain). *Stochastic Environmental Research and Risk Assessment*, DOI: 10.1007/s00477-020-01771-7.

Relevance of spatio-temporal rainfall variability regarding groundwater management challenges under global change.

Case study in Doñana (SW Spain)

Relevancia de la variabilidad espacio-temporal de las precipitaciones con respecto a los retos de gestión de los recursos hídricos bajo el marco del cambio global. Caso de estudio en Doñana (SO España)

Nuria Naranjo Fernández¹, Carolina Guardiola Albert², Héctor Aguilera², Carmen Serrano Hidalgo², Miguel Rodríguez Rodríguez³, Ana Fernández Ayuso³, Fernando Ruíz Bermudo² y Esperanza Montero González¹

1. Facultad de Ciencias Geológicas. Universidad Complutense de Madrid. Calle José Antonio Novais 12, 28040, Madrid; nuria.naranjo.phd@gmail.com; emontero@ucm.es;
2. Instituto Geológico y Minero de España. Calle Ríos Rosas - 23 - 28003 Madrid – España. c.guardiola@igme.es
3. Departamento de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, Carretera de Utrera km 1,41013 Seville, Spain

* Corresponding author: c.guardiola@igme.es

Abstract

Rainfall is the major contribution for groundwater recharge in arid and semiarid climates, therefore a key factor in water resources estimation. This work presents the results of an in-depth study in Doñana National Park concerning groundwater recharge behavior over a long period (1975–2016). The spatio-temporal kriging algorithm was used as a supportive tool to improve the reconstruction of the spatio-temporal rainfall variability. One of the main findings was that monthly recharge estimations range between 21 and 91% of the maximum rainfall, being overestimated in areas that also demonstrate spatial heterogeneity in rainfall distribution. In the light of these results, for water management purposes in the Mediterranean area, rainfall spatio-temporal scale is a critical aspect and it must be taken into account in groundwater reservoir allocation. Moreover, it is highlighted that local studies of rainfall and recharge, in an area of high ecological fragility, are essential to developing management strategies that

prevent climate change effects and guarantee optimal conditions for groundwater resources in the future.

Keywords: Global change impacts; Rainfall variability; Spatio-temporal kriging; Water resource management.

Resumen

La principal contribución de la recarga en climas áridos y semiáridos es la precipitación, además de un factor clave en la estimación de recursos hídricos. Este trabajo presenta los resultados de un estudio profundo en el Parque Nacional de Doñana sobre el comportamiento de la recarga de aguas subterráneas en un largo periodo (1975-2016). El kriging espacio-temporal ha sido utilizado como herramienta para mejorar la reconstrucción de la variabilidad espacio-temporal de la precipitación. Uno de los descubrimientos principales ha sido que la estimación mensual de la recarga supone entre un 21 y un 91% total de la lluvia máxima siendo sobreestimada en las áreas en las que se demuestra una heterogeneidad espacio temporal de la distribución de la precipitación. Tomando en cuenta los resultados obtenidos, en términos de gestión hídrica en el área Mediterránea, la escala espacio-temporal de la precipitación es un aspecto crítico que debe tenerse en cuenta en la localización de las reservas de agua subterránea. Además, se destaca que los estudios locales de precipitación y recarga, en áreas con elevada vulnerabilidad ecológica son esenciales para el Desarrollo de estrategias de gestión que previenen los efectos del cambio climático y garantizan unas óptimas condiciones para los recursos hídricos en el futuro.

Palabras clave: Impactos del cambio global; Variabilidad de la precipitación; Kriging espacio-temporal; Gestión de recursos hídricos.

1. Introduction

The natural environment of Doñana boasts great ecological wealth associated with the wetlands, shallow ponds and lagoons that form around its extensive sand dunes and marshlands. It was declared a Ramsar wetland in 1982, forcing the Spanish Government to adopt the necessary actions for the conservation of the high ecological value of this area. The high biodiversity wealth of these wetlands is threatened by a reduction of the water inputs due to the demands of water for agriculture, tourism and urban demand. Likewise, qualitative characteristics are diminished by industrial pollution and acid mine drainage. In addition, available water resources could decrease due to climate change variations in the future (Guardiola-Albert and Jackson 2011).

Much of the difficulty to establish sound water policies lies in the complexity of surface–groundwater interconnections, requiring site-specific data of many hydrological variables to produce reliable estimations (Dimitriou et al. 2017). The sedimentary nature of the underlying aquifer system gives high spatial heterogeneity to the recharge conditions and the groundwater fluxes that maintain large and complex ecosystems (Manzano et al. 2009). In addition, the spatio-temporal variability of recharge is important in response to meteorological forcing and, also, contributes to spatial distribution of groundwater resources. Moreover, in this Mediterranean area with Atlantic influence, where the margin between precipitation input and evapotranspiration output is small or even negative, the amount of groundwater recharge has a major influence on the water balance.

Water resources management in sensitive areas can be considered as one of the most difficult challenges nowadays. Recent studies have highlighted the importance of controlling local stressors, such as water extraction, to prevent undesired effects of climate change and ensure the maintenance of these places beyond thresholds of irreversible collapse of the entire ecosystem (Green et al. 2017). Therefore, the resource management plans must take into account global change variations in order to be cautious with the available water resource estimation during heavy drought periods. The hydrological management plan (HP, CHG 2015) estimates that in the Doñana area recharge, on average, is 46% of the total run-off, rising up to 82% for some periods and

groundwater bodies. In the Doñana area, water authorities use monthly or annual precipitation data from 80 rain gauges to estimate groundwater recharge with SIMPA (Integrated System for Modeling the Precipitation Contribution process, Álvarez-Rodríguez et al. 2005) and to make management decisions (CHG 2015, 2017). On the other hand, among the huge amount of scientific research published on hydrology in Doñana, some studies are based on just one to two rain gauges (Sacks et al. 1992; Dimitriou et al. 2017; Rodríguez-Rodríguez et al. 2017). In other words, the aforementioned studies do not take into account the influence of the spatio-temporal variability of rainfall on their conclusions.

There are plenty of studies about the influence of the scale of study on rainfall, regarding both the temporal and spatial scales (Buytaert et al. 2006; Nandargi and Mulye 2012; Moschou et al. 2013; Shi et al. 2014; Zurbieta et al. 2017; Deng et al. 2018). However, just a few of these studies analyze the influence of this variability on the evaluation of groundwater resources or hydrological processes (Narjary et al. 2014; Mumtaz et al. 2018). When groundwater resources mainly come from rainfall recharge, spatio-temporal variability of precipitation directly influences their quantification. This impact can be important when numerical flow models using MODFLOW 2005 (Harbaugh 2005), or previous versions of this code, are used to evaluate groundwater resources, as in the case of our study site (Guardiola-Albert et al. 2016). Recharge in MODFLOW is treated as an areal feature; being commonly computed in a number of previously established polygons on a monthly scale.

Geoestatistical interpolation of hydrometeorological data at daily time step is not new and research can be found in the literature (Carrera-Hernández and Gaskin 2007; Ly et al. 2011; Chappell et al. 2013; Yang et al. 2015). Moreover, there are also works in the Spanish territory estimating daily rainfall values with kriging (e.g. Herrera et al. 2012). However, none of them use the information of the temporal correlation structure. Militino et al. (2015) assume that when data are available in space and time, spatio-temporal models usually result in better estimates. Hence, they introduce time in the kriging estimation by adding it as a new covariate the average precipitation of each 5-day period within each month of the studied period in northern Spain. In 2012, Pebesma

(2012) published a tool to analyze spatio-temporal variability with geostatistics using the spatio-temporal variogram and kriging. It is a relatively new tool that has scarcely been used in rainfall or hydrogeological studies (De Iaco and Posa 2012; De Iaco et al. 2012; Mazione et al. 2019; Ruybal et al. 2019; Varouchakis and Hristopulos 2019). Although there are been many studies that deal with the evaluation of the spatio-temporal dependence of precipitation and discharge, spatio-temporal kriging has never been used in the Doñana area to evaluate the importance of the spatio-temporal variability of rainfall on hydrological and hydrogeological studies.

Underlying global climatic changes, an increase in precipitation extremes, as well as more prolonged and intensive droughts, are very likely in the study area (IPCC 2007; Guardiola-Albert and Jackson 2011; Green et al. 2017). One parameter which can help provide a better picture of drought conditions than monthly rainfall, is the number of rain days (Burgueño et al. 2005; Nandargi and Mulye 2012; Shi et al. 2014; Monjo and Martin-Vide 2016; Zurbietta et al. 2017), or alternatively the number of days with recharge. The term “rain day” indicates a day with rain higher than the detection limit (≥ 0.1 mm) has been recorded at any station. This parameter serves as a proxy for frequency of occurrence. For instance, the drought season may be marked by both fewer rain days and less rain per day as compared to periods of normal and above normal rainfall. Spatial distribution of number of rain days and its time trend has been depicted in wider areas than Doñana in the north of Spain (Burgueño et al. 2005). These kinds of studies are relevant in this sensible wetland, which is the most important wintering site for waterfowl in Europe (Scheffer et al. 2015), and where drought periods determine how the Water Management Administration allocates groundwater use rights.

There is a need to bind the gap between spatio-temporal variability of rainfall and suitable groundwater resources allocation. What is the adequate spatio-temporal scale of rainfall to study and manage the water resources in the Almonte-Marismas aquifer? Is it enough to work at annual or monthly scale? When should daily data be used? What is the range of spatio-temporal rainfall correlation within the aquifer? Are drought periods equally distributed around the different areas of the aquifer? The aim of this study is to answer these questions analyzing the influence of the spatio-temporal

variability of rainfall on a key aspect for groundwater resources in Doñana: recharge estimation controlling water resources availability under the effect of climate variations.

2. Study area and data

2.1. Almonte-Marismas Aquifer and its dependent wetlands

Almonte-Marismas Aquifer, in the Andalusia region, in southwest Spain, is located in a sedimentary basin that represents the last stage of the Guadalquivir River basin filling: the Guadalquivir Marshlands, with heights between 0 and 15 m.a.s.l. (Fig. 1). Administratively it belongs to the Hydrographical Confederation of the Guadalquivir (CHG is the Spanish acronym) and it is divided into 5 groundwater bodies (CHG 2015). Guardiola-Albert et al. (2016) highlight that this subdivision does not correspond to hydrogeological property areas. Its hydrogeological boundaries are the Guadalquivir River in the east, Atlantic Ocean in the south, Tinto River in the west and impermeable geological materials in the north. Basal sands, silts and eolian sands are the principal geological materials that form the Almonte-Marismas Aquifer, which is unconfined to the west and southwest, but from the east to the southeast it is confined by Doñana marshlands (Salvany and Custodio 1995).

Regarding the Doñana wetlands, which are located on the Almonte-Marismas aquifer, there is great variability in the geomorphological origin and hydrological behavior, but many of them depend on groundwater discharge (Custodio 2001; Manzano 2001). Most of the ponds, streams and springs are seasonal, although some are permanent, and their water comes mainly from the groundwater when it intercepts the surface of the land. This mechanism is becoming less frequent in ecologically significant areas as a consequence of increasing groundwater exploitations (Serrano and Serrano 1996). The natural flow is currently modified by pumping to meet agriculture and tourism demands.

2.2. Rainfall data

Rainfall data have been obtained from four Spanish public institutions: AEMET (Spanish State Meteorological Agency), Junta de Andalucía (Regional Government of Andalusia,

RGA), CGH and CSIC (Higher Council for Scientific Research). 113 weather stations (Fig. 1) have been studied: 77 from AEMET, 26 from RGA, 9 from CHG and 1 from CSIC.

There are records of rainfall data between 1903 and 2017, with some data gaps related to political or economic changes. Rainfall was analyzed in the 112 rain gauges measuring between 1975 and 2016 (Fig. 2; Appendix 1), but all recorded data were preprocessed in the same way, independently of the source. Missing values were left unfilled and measures under the detection limit (< 0.1 mm) were assumed to correspond to zero rainfall. Before 1975, there were many large time lags missing information, therefore they were discarded for this analysis and one rain gauge was not employed. The maximum daily precipitation intensity data was 168.2 mm/day (Appendix 1). In some cases, the daily record was complemented with temperature data (average, maximum and minimum) whenever the weather stations registered daily multi-parametric data.

As an example of spatial rainfall variability in the study area, Fig. 3 shows the differences between daily precipitation raw data of two rain gauges (a and b in legend of Fig. 1) located 347 m away (Palacio de Doñana (CSIC/ AEMET) and Almonte-Doñana (AEMET)). Figure 3 shows that there is a wide variation range, differing up to 90 mm for a given day.

3. Methodology

3.1. Spatio-temporal analysis and interpolation with geostatistics and R

The spatio-temporal experimental variogram provides an empirical description of the spatial and temporal continuity (i.e. spatio-temporal correlation). Once the experimental variogram has been computed, a model is also fitted for prediction. The separable variogram model is represented as the product of a spatial and temporal term (Gräler et al. 2016):

$$Y_{sep}(h, u) = sill \cdot (\bar{Y}_s(h) + \bar{Y}_t(u) - \bar{Y}_s(h)\bar{Y}_t(u))$$

where \bar{Y}_s and \bar{Y}_t are standardized spatial and temporal variograms with separate nugget effects. The overall sill parameter is denoted by “sill”. This type of model has also worked

well with other field data such as groundwater levels (Manziona et al. 2019; Ruybal et al. 2019; Varouchakis and Hristopulos 2019).

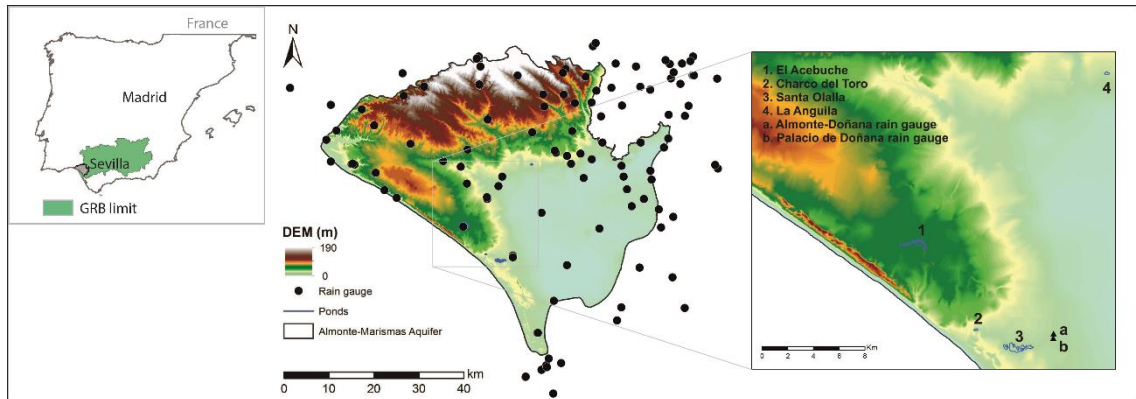


Figure 1. Rain gauge locations in the Almonte-Marismas aquifer.

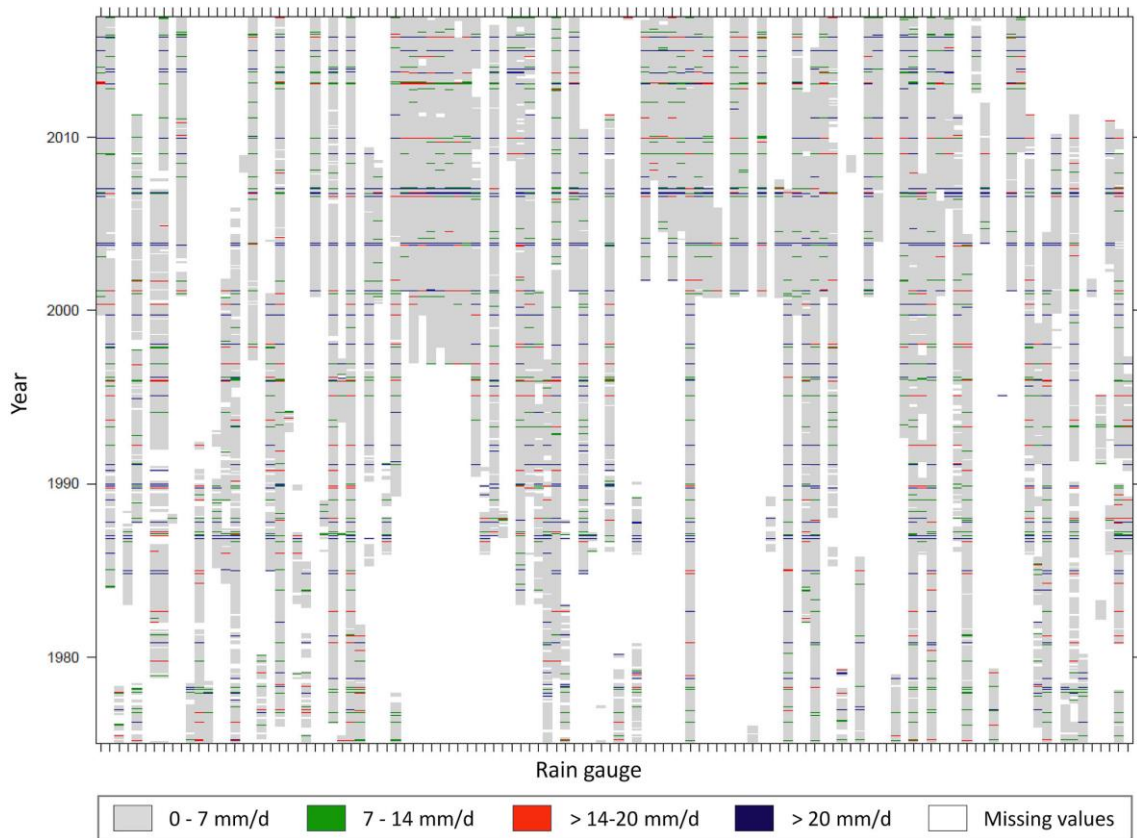


Figure 2. Temporal distribution of available precipitation measurements. The Y-axis represents the date when the rainfall was registered and the X-axis are the rain gauges.

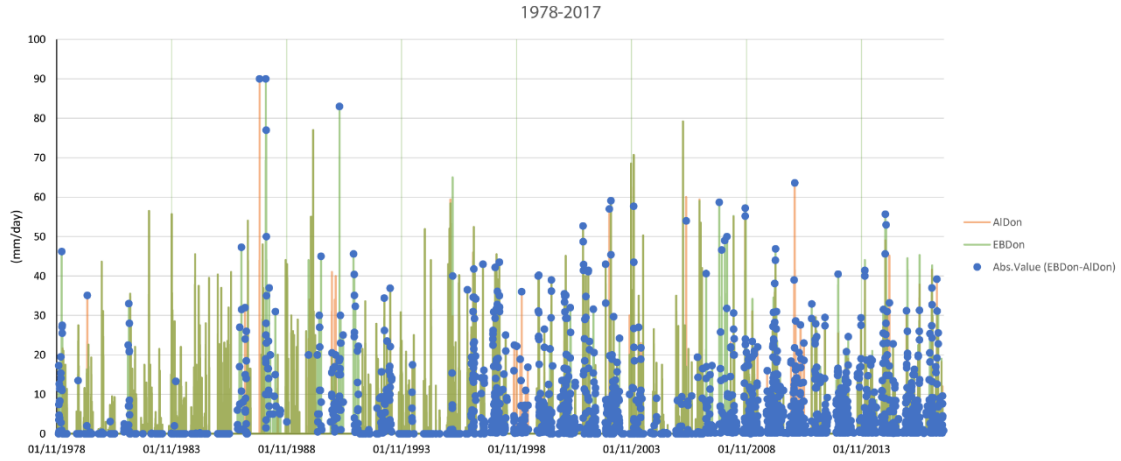


Figure 3. Blue dots show the differences in mm (absolute values) between raw daily precipitation data from Palacio de Doñana (green line) and Almonte-Doñana (orange line) rain gauges.

Note that the product covariance model fits data better than other spatio-temporal covariance models, such as the product-sum, metric and sum-metric models (Gräler et al. 2016), which we tried during model selection. *gstat* package allows to assess the goodness of fit of model and sample variogram giving the unweighted mean squared error between the modeled and experimental variogram surfaces (Pebesma 2012).

The spatio-temporal variogram model was then inputted into the spatio-temporal kriging algorithm to obtain rainfall estimations at unknown times and locations. Spatio-temporal estimation based on kriging methods is very similar to the spatial standard kriging. The approach of the associated predictor equations and the variogram model are the main differences. The spatio-temporal estimation expression is (Pebesma 2012):

$$\hat{Z}(s_0, t_0) = \sum_{i=1}^n \lambda_i Z(s_i, t_i)$$

where $\hat{Z}(s_0, t_0)$ is the estimated value of the variable in the unknown locations (s_0, t_0) and λ_i are the weights of the corresponding data sample $Z(s_i, t_i)$.

Experimental variograms, fitting variogram models and estimation of rainfall were generated by the *gstat* package R software (Pebesma 2012). Interpolated maps have a 500 x 500 cell size resolution.

Spatio-temporal kriging may not be appropriate in the context of non-stationarity. Christakos (2000) proposes a generalized spatiotemporal variogram model that can be associated with non-stationary random fields. Others authors remove the spatio-temporal trend, and then apply spatio-temporal kriging to the residues (Varouchakis and Hristopulos 2019). In the present study case, the stabilization of the spatio-temporal variogram sill, which can be seen in the Results section, and the use of a subset of points to perform kriging (i.e. neighborhood), allows relaxing the stationary assumption to local intrinsic spatio-temporal stationarity (Wikle et al. 2019).

The performance of the interpolation was checked in 10 rain gauges spatially distributed (Table 1); five with a high number of missing values (90%) and the other five with low degrees of missing information (up to 25%). Each time series data was split in two, choosing first 50% of the series to be available and the last 50% of the series for testing. The performance of the estimation was assessed using the coefficient of determination (R^2), mean absolute error (MAE) and normalized root mean squared error (NRMSE).

The number of rain days provides information on how precipitation accumulates over a few or several days. From the interpolated values, rain days at an annual rate were computed in all the rain gauge locations when the estimated value was higher than 0.1 mm.

Table 1. Summary statistics in daily precipitation for the 10 series selected for interpolation validation (SD: standard deviation; MV: missing values; n: number of available records).

Rain gauge	Mean (mm)	Median (mm)	SD (mm)	Range (mm)	Rainy days (%)	MV (%)	n
AIRAi	1.03	0.00	5.42	70.20	34.36	97.86	326
MoRAi	1.34	0.00	6.05	70.50	32.00	97.70	350
AzLBi	2.45	0.00	7.55	65.00	16.67	96.61	516
LPVil	1.82	0.00	6.74	65.00	10.78	95.43	696
ViAPu	2.37	0.00	7.46	77.00	19.20	90.52	1,443
ALAce	1.65	0.00	6.12	101.50	15.40	25.32	11,365
Chipi	1.49	0.00	5.91	128.50	19.31	24.21	11,535
EBDon	1.52	0.00	5.86	90.00	15.35	13.34	13,189
AzLJu	1.63	0.00	6.25	112.50	17.06	6.03	14,302
MOINM	1.53	0.00	6.15	106.10	13.18	5.66	14,357

To show how the interpolation preserves marginal properties and spatio-temporal intermittency, Bennet et al. (2018) suggested a wide range of statistics which are important from a hydrological perspective. Among other statistics, they analyzed rainfall

occurrence/amounts, wet/dry distributions, seasonality, spatial gradients or temporal correlations across range of time scales. In the present work, the following statistics were computed before and after the interpolation in all the rain gauges: (i) median and variance of monthly rainfall, when 20 or more days are measured, in each month of the year, (ii) mean and variance of daily values for every year and (iii) the median percentage rate of number of rain days in a month and the number of days with rain data in that month. In order to have visible and interpretable plots, and to look into spatial gradients, all these statistics were computed over the whole studied period and 15 recharge zones (see next Sect. 3.2.).

For uncertainty analysis, a high number (e.g. 100) of spatio-temporal simulations need to be performed. However, at the spatio-temporal scale of the problem presented here and with the existing spatio-temporal tools, obtaining a considerable number of simulations is too computationally expensive. Another way to evaluate precipitation interpolations is to use an objective verification based on a subsequent application such as stream flow simulations with a hydrological model (Ly et al. 2013; Chen et al. 2019).

Another application of the spatio-temporal variogram are predictive precipitation maps, which can be generated for any time point based on the optimal temporal space variogram model through geostatistical simulations algorithms (Ruíz-Villanueva et al. 2013).

3.2. Groundwater recharge

The groundwater numerical model that supports the Extraction Plan in the Almonte-Marismas Aquifer for the Guadalquivir River Basin Authority (Guardiola-Albert et al. 2016) is developed with MODFLOW 2005 (Harbaugh et al. 2005). The MODFLOW Recharge package was originally developed to treat recharge as an areal feature. Hence, the study area was divided into recharge subzones within which the infiltration was assumed to be homogeneous (Figure 4). This division in 15 zones was based on the study carried out by Custodio et al. (2009), following geological and hydrogeological criteria (Figure 4). Groundwater recharge was then computed using Trasero software for hydrological time series analysis (Padilla and Delgado 2013).

Then, in the same way as for the aforementioned supporting groundwater numerical model, one rain gauge per zone was selected. These rainfall data series will be used to calculate the recharge. Prior to entering the daily data in Trasero, gaps in rainfall and temperature series were infilled. The spatio-temporal distribution of temperature variation is much smoother (i.e. higher correlation) than that of precipitation and temperature does not show the zero value problem. Therefore, a regression method is suitable for temperature imputation. We used the missForest method (Stekhoven and Bühlmann 2012), which is a non-parametric method for missing data imputation based on the random forest algorithm. This method has shown high performance in datasets including mixed-data types and complex interactions and non-linear relations, particularly with increasing correlation (Stekhoven and Bühlmann 2012; Tang and Ishwaran, 2017). It was performed with the missForest package from R (Stekhoven 2013) that can be run in parallel and provides much faster calculations than spatio-temporal kriging.

The potential evapotranspiration was calculated using the Hargreaves method (Hargreaves 1989). The Thornthwaite empirical method (Thornthwaite 1948) was used to estimate the effective rainfall value (ER), and the curve number method (CN) from the Soil Conservation Service (Mockus 1964; SCS 1975) was employed to determine the direct runoff fraction of ER. The CN is a simple, widely used and efficient method for determining the approximate amount of runoff from a rainfall even in a particular area. This method only requires rainfall amount and curve number. The curve number is based on soils, plant cover, impermeable areas, interception, and surface storage. Hargreaves, Thornthwaite and CN are included in the Trasero software.

In the CN method, a runoff threshold was estimated based on tabulated values that depend on soil type, vegetation cover, hydrological characteristics and land uses, the runoff threshold for each sub-area was calculated and then the direct runoff values were determined.

The natural recharge of each subzone was obtained by subtracting the direct runoff value from the ER. Daily and monthly recharge values were assessed by using daily and monthly temperature and precipitation data. Monthly aggregated values were computed from daily rainfall and temperature data. Daily estimated recharge values

were aggregated to give monthly recharge values for comparison with monthly recharge values computed from monthly rainfall data.

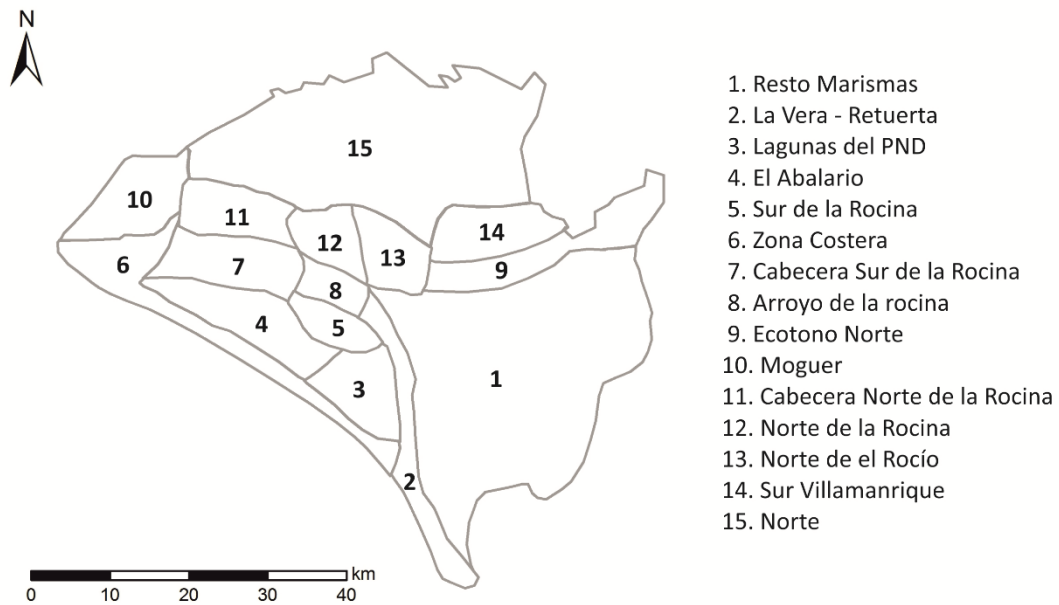


Figure 4. Recharge subzones distribution.

When studying a climate series, it is important to look at the inflection points to segment the climate series into sub periods and to carry out an in-depth study of the series (e.g. Yue and Wang 2002; Samdi and Zghoul 2006; Arun et al. 2012). The inter-annual variability of recharge and days with recharge was analyzed computing paired t-test (Zimmerman 1997) for two time periods (1975-1998 and 1999-2016). For a visual study of trend, loess regression in R has been used, which is a non-parametric approach that fits multiple regression lines in a local neighborhood. Figure 5 summarizes the methodological approach used in the study.

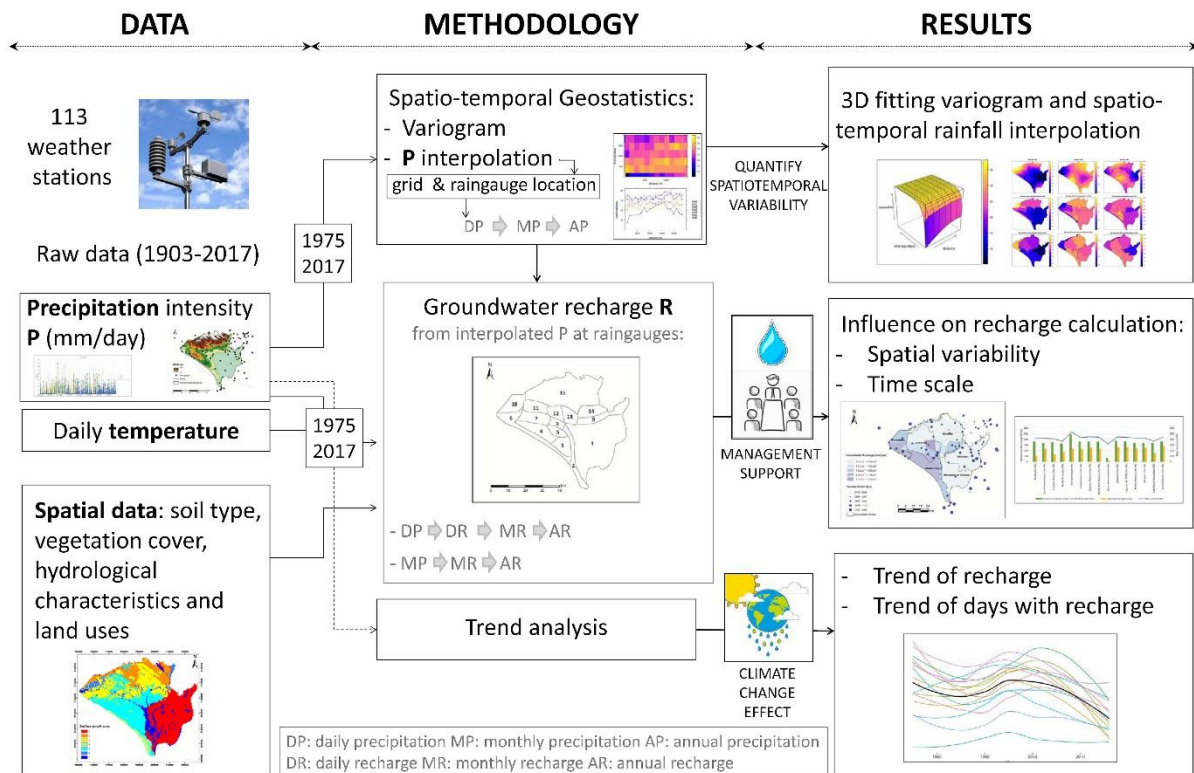


Figure 5. Flowchart of the methodological approach used in the current study.

4. Results and Discussion

4.1. Spatio-temporal rainfall interpolation

The top graph in Figure 6 shows the spatio-temporal model variogram in 3D. This separable variogram model (Gräler 2015) was chosen because it gave the best fit among all the models tested. The lower graphs in Figure 6 show the experimental (on the left) and modeled (on the right) spatial variograms for different time lags. The good fit obtained with the spatio-temporal variogram model can be visualized. The mean absolute error given by the fitting gstat tool is 0.89 mm. The total fitted spatio-temporal sill was of 36 mm²/day. Among all variogram parameters, the nugget is the most unpredictable one because of the lack of proximity between samples (Armstrong, 1998). Measurement errors were taken into account in the standardized spatial and temporal models by means of the partial nuggets (i.e., 0.2 mm²/day and 0.5 mm²/day, respectively). The other contribution to the nugget, the microscale variation of rainfall, was neglected because of the following constrains: (i) when large lag tolerance is used in the experimental variogram calculation (here 4,000m), there is an overestimation of

the experimental nugget (Negreiros et al. 2010), (ii) there is bibliography pointing out that nugget is negligibly small in the case of rainfall intensity data at daily averaging interval (Muthusamy et al. 2016), and (iii) in this smooth topographical area, there is no data available on microscale rainfall variations that would allow to give a certain value for this component. The spatial component was modeled using a spherical variogram with a 45 km range. The temporal component was fitted with one exponential model with a range of 6 days. These range values indicate that rainfall presents spatial correlation up to 45 km and temporal correlation to 6 days. Spatially, this has an important interpretation, as this is roughly half of the size of the Almonte-Marismas aquifer. Spherical model represents continuous variables, which is in accordance with the lack of spatial complexity of daily rainfall regime in a smooth topography area, such as Doñana. The explanation lies in the atmospheric circulation patterns and their effects on rain episodes (Burgueño et al. 2005).

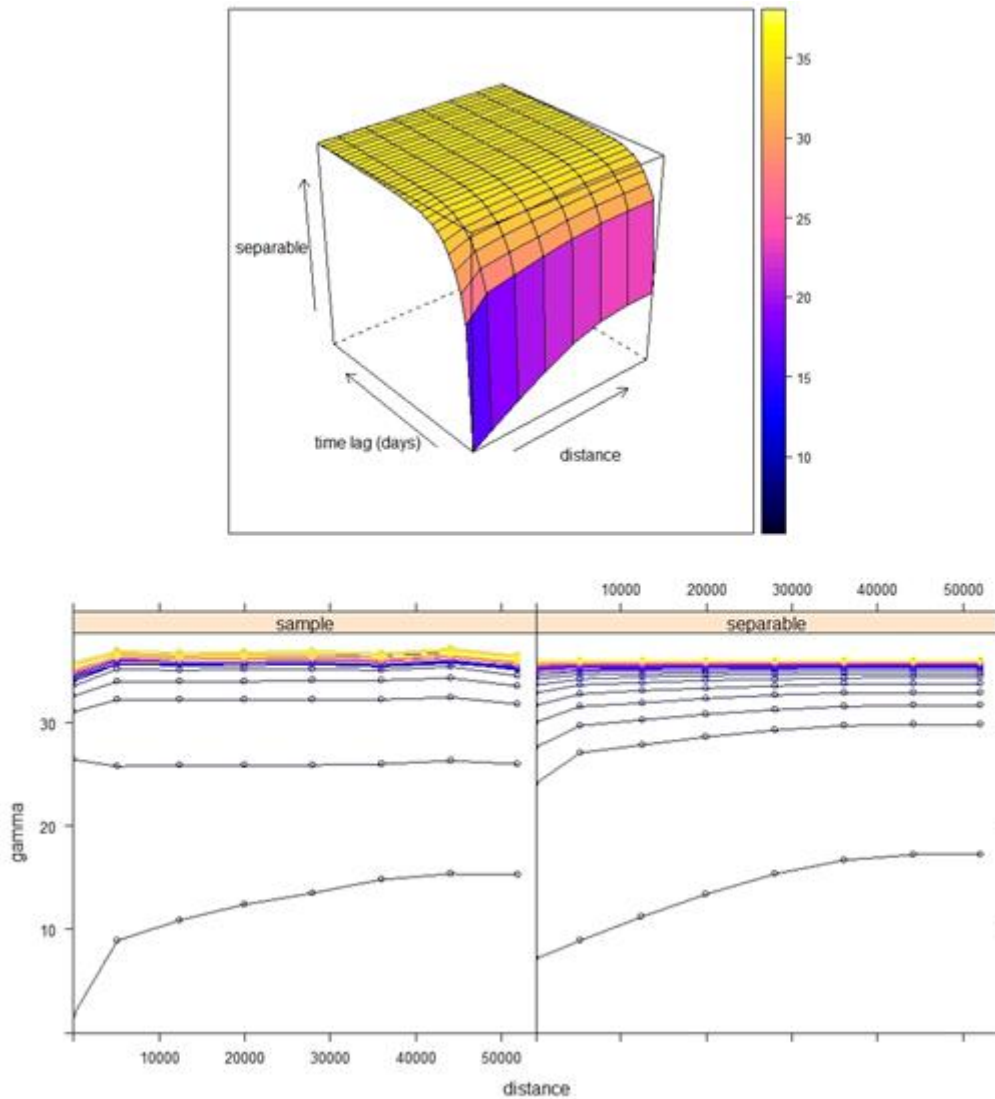


Figure 6. Spatio-temporal variogram of rainfall. Upper graph is the spatio-temporal variogram model, lower left graph represents the spatial experimental variogram for different temporal lags, and the lower right graph shows the corresponding fitted model. Colors given in the legend are valid for upper and lower graphs.

The variogram model was then used to interpolate maps or complete series through spatio-temporal kriging. Taken into account the rainfall range and standard deviation of the 10 test data series (Table 1), error metrics obtained indicate a good performance of the spatio-temporal kriging interpolation with median values of $R^2 = 0.69$, $MAE = 1.16$ mm and $NRMSE = 0.05$.

Appendix 2 and 3 provide information about the rainfall interpolation accuracy. Appendix 2 shows observed vs predicted and bias plots for the 10 rain series chosen to test the performance of the interpolation. It can be seen that when there is a high percentage of missing values, interpolations are less accurate (Figure A2.1 a to e) than

predictions of rain gauges with low percentage of missing values (Figure A2.1 f to j). However, in general, interpolated values are a good approximation of observed values (Figure A2). Only in two rain gauges (Figures A2.1 d and h) R2 values are low. Bias plots reflect a general overestimation of low rainfall and underestimation of higher precipitation intensities, which is a consequence of the kriging smoothing effect (Figure A2.2). All these results confirm the inherent difficulties when daily rainfall is interpolated (e.g. heavily skewed distribution and large spatio-temporal variability). Appendix 3 includes three figures plotting the statistics calculated in each recharge zone to check the interpolation performance. Figure A3.1 shows that kriging smooths monthly median and variance rainfall, rainfall occurrence and rainfall amounts. Mean and variance of interpolated daily values within the year are consistent with measured values in all the recharge zones, with some smoothing and variance decrease associated to the kriging interpolation method (Figure A3.2). In the majority of recharge zones the number of rain days after interpolation increases (Figure A3.3). This is the main drawback of the kriging method. Due to the spatio-temporal correlation given by the variogram (e.g. 6 days range), in the imputation process there are a high quantity of days with very low precipitation (around 1 mm or less). This causes that, even they do not trigger recharge, they are counted as rain days. The other important disadvantage is the high computing times required for each interpolation for a long period at a daily time step. The main advantage of the spatio-temporal kriging is the capability to estimate precipitation consistently, with encouraging results under extreme missingness (up to 98% of missing values).

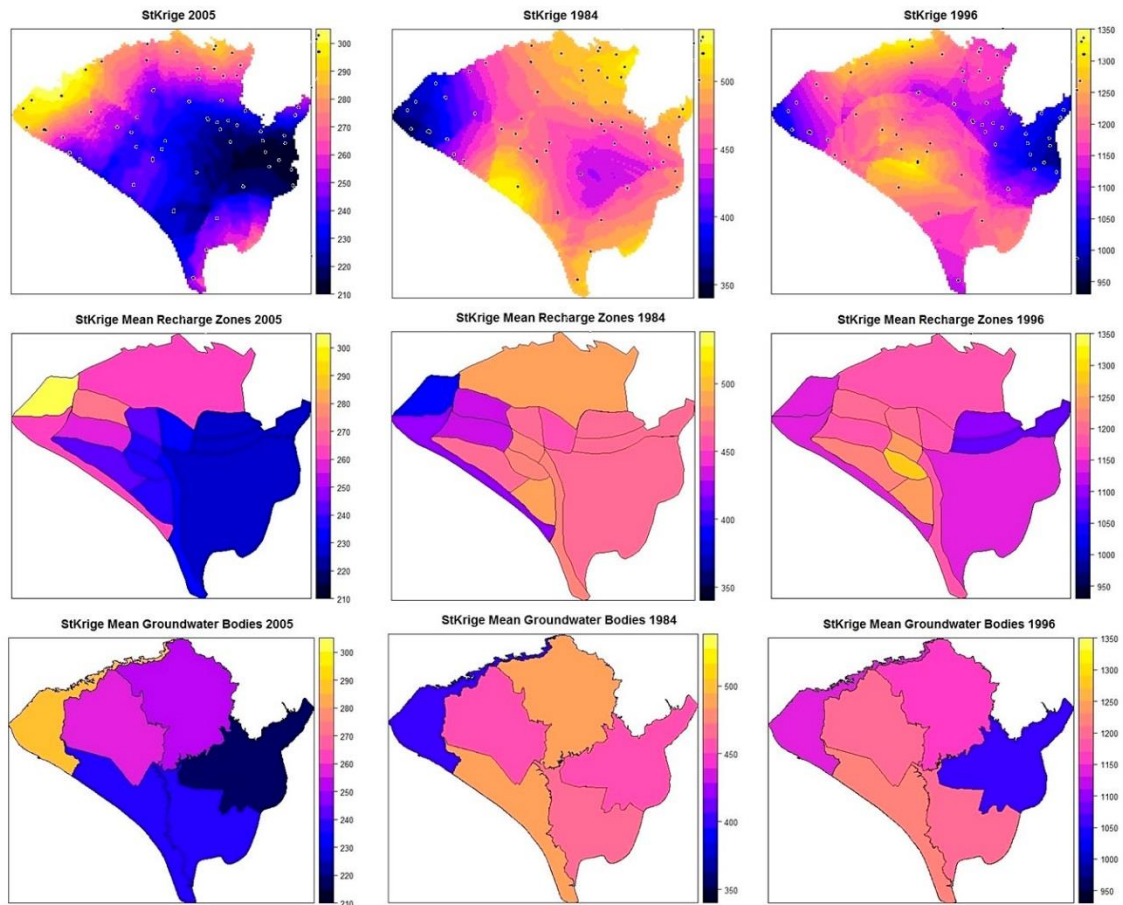


Figure 7. Rainfall (mm) interpolations with a resolution of 500x500 m pixels; dry year (2005), average year (1984) and wet year (1996). 500x500 m pixels interpolated values are represented in the first row, recharge subzones in the second row and groundwater bodies in the third row. Dots in the first row show the rain gauges located in the Almonte-Marismas aquifer.

The first row in Figure 7 represents three examples of the interpolated annual rainfall maps in dry (2005), average (1984) and wet (1996) years, respectively. Mean values within recharge subzones and groundwater bodies correspond, respectively, to second and third rows. The La Rocina groundwater body (Figure 8) is one of the most critical areas for water management in Doñana, with wide areas of berries irrigated by many illegal extractions (WWF 2016). As expected, mean values of rainfall for La Rocina groundwater body (third row in Figure 7) differ from spatially distributed rain values (first row in Figure 7) or from recharge zone rain values (second row in Figure 7). This implies that in dry periods, as 2005, the mean rainfall value in La Rocina can be overestimated or underestimated by 30 mm/year. However, in a wet year, like 1996, the rain calculation bias can reach up to 100 mm/year. These differences show that, in this study area, with high spatio-temporal rainfall variability, the use of an average or

single value of rainfall could considerably increase or decrease the estimation of natural resources, especially in wet periods.

The total number of rain days during the study period (1975-2016, i.e. 15,341 days) is shown for the interpolated rainfall series at rain gauge locations (Figure 8). There is a wide range of rain days among all the stations without a clear spatial pattern, between 2,110 and 5,309 days. Values shown in Figure 8 reflect the high spatial variability of rain days within some groundwater bodies (e.g. Almonte) or in general at the northern recharge subzones. Again, this finding contrasts with the flatness of Doñana's topography.

4.2. Spatial variability influence in recharge calculation

Understanding groundwater recharge is essential for successful management of water resources (Healy 2010). This work proposes using 15 recharge subzones (Figure 4) (Custodio et al. 2009). The mean annual recharge estimations obtained in this study are shown in Figure 8. Today CHG manages water in Doñana area, computing water reserves (CHG 2015) based on monthly groundwater recharge estimations obtained from SIMPA (closed source soil water balance method similar to the open source method used in this work, Álvarez-Rodríguez et al. 2005). The on-going HP (CHG 2015) is the main water management instrument in Doñana. In the HP, monthly recharge values are aggregated to annual values within each groundwater body (Figure 8) to compute actual water reserves. These limits of groundwater bodies (CHG 2015) within the Almonte-Marismas aquifer are defined to help water administration, but do not have clear hydrogeological meaning (Guardiola-Albert et al. 2016). For example, inside the aforementioned La Rocina groundwater body, there are four recharge subzones. Among them, mean recharge values range from less than 40 mm/year to more than 252 mm/year. The recharge value used for management decisions will be an average of them. The HP cannot handle detailed values in each pixel to manage the water resources, but it is highlighted here that errors associated with aggregated values should be considered when making relevant decisions such as how much water can be extracted from this part of the aquifer each year. Moreover, HP computes groundwater reserves, roughly approximating the fluxes between the aforementioned groundwater bodies, as their

contacts are difficult to interpret. Uncertainties in groundwater recharge related to the mean spatial values will be an additional inaccuracy factor for these misinterpreted interchange fluxes.

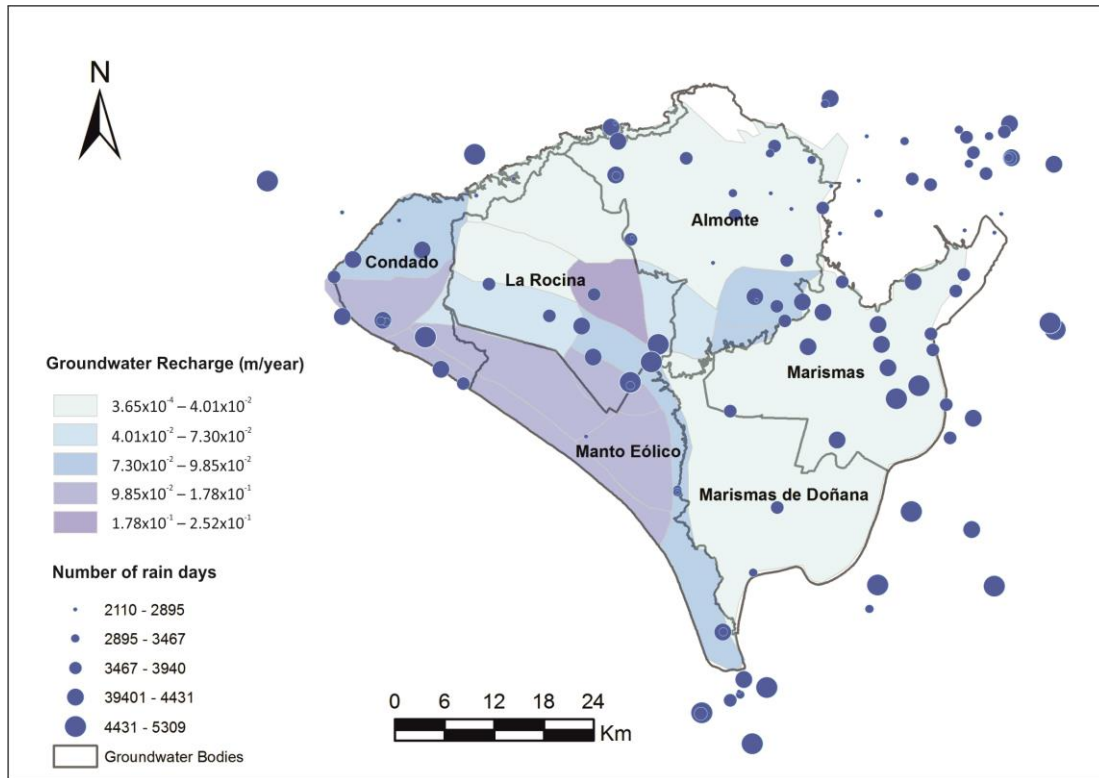


Figure 8. Mean annual groundwater recharge values per zone in Almonte-Marismas aquifer and total number of rain days from the interpolated series in each station between 1975 and 2016. Groundwater bodies division is also plotted.

For each recharge subzone, Figure 9 depicts, maximum (upper graph) and median (lower graph) monthly aggregated recharge values from daily estimations in green bars and their corresponding monthly recharge calculation (in yellow) and rain (with a blue line) over the same month. It reveals high spatial variability in maximum and median values comparing the 15 recharge subzones; ranging from 67 to 488 mm in the maximum values and from 23 to 78 mm in median values. Each recharge subzone has a dominant soil texture (i.e. sandy, silty, clayey) that conditions the water-holding capacity and runoff threshold parameters that control the amount of water than can be stored in the soil and infiltrated from the surface, respectively. Finer textured zones such as Resto Marismas with the clay present in the marshland, have greater water-holding capacity and lower runoff threshold, resulting in small recharge values. Sandy soils such as those in Sur Rocina show the opposite behavior leading to higher recharge rates. These results

for the monthly scale also highlight the importance of dividing the aquifer into recharge subzones with hydrogeological criteria. Continuing with the example of La Rocina groundwater body, it can be seen that recharge estimations within its limits vary from 332 to 448 mm for the maximum recharge month and from 33 to 62 mm for the median recharge month.

Scanlon et al. (2002) stated that spatial variability in recharge on a local scale may not be critical for water-resource assessment. However, groundwater recharge rates differences of more than 200 mm/year are found between the different recharge zones in the Almonte-Marismas Aquifer. In the context of the existing climate change (Guardiola-Albert and Jackson 2011) and the critical situation of the Almonte-Marismas Aquifer regarding water scarcity, overexploitation and droughts events reported by scientists (Green et al. 2016) and ecologists (WWF 2016), these differences in recharge estimations are significant for aquifer sustainability.

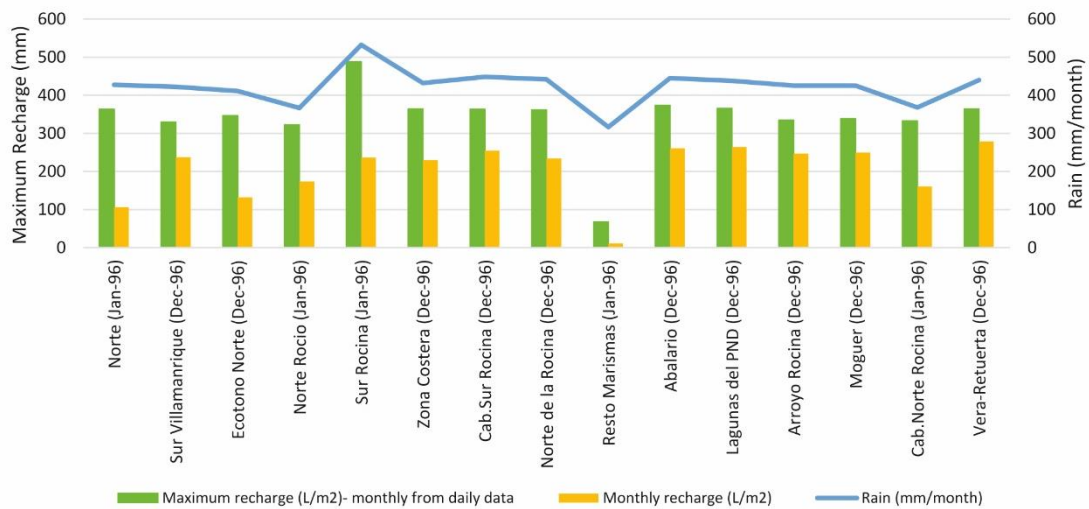


Figure 9. Cont.

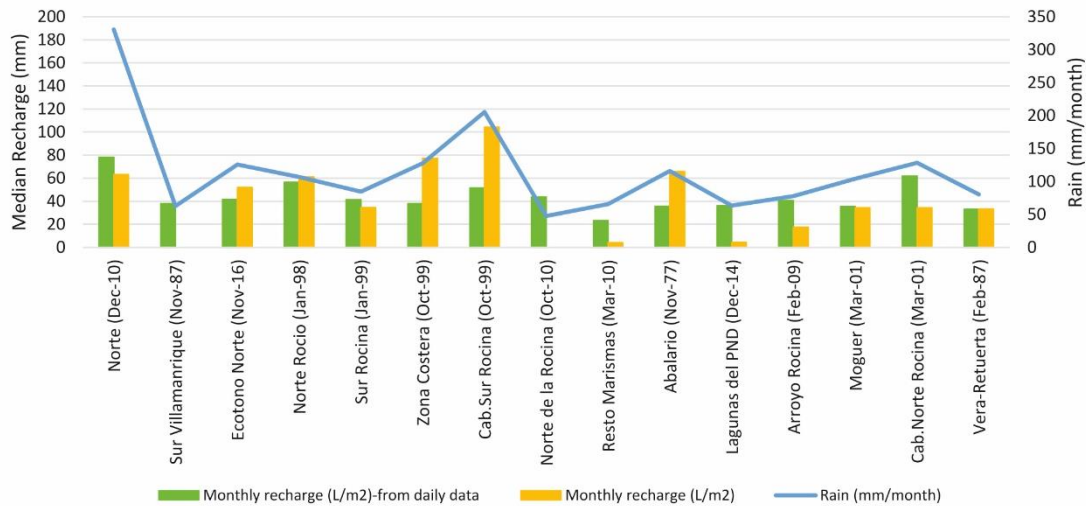


Figure 9. Maximum (upper graph) and median (lower graph) monthly aggregated recharge from daily data, including monthly recharge and rain corresponding to the same month. The corresponding month appears in brackets next to the subzone name (X-axis).

4.3. Time scale influence in recharge calculation

Recharge values at decadal time scales is generally required for water-resource planning (Scanlon et al. 2002), but this paper intends to analyze in depth the effect of time scale on water resources estimations. As Scanlon et al. (2002) support, it is assumed here that daily recharge computations are more accurate than monthly recharge calculations. The temporal variogram range indicates that there is a temporal correlation of six days that was taken into account for the daily interpolations.

Maximum and median monthly aggregated recharge values from daily estimations (green bars in Figure 9) are compared with its corresponding monthly recharge calculation (in yellow) and rain (blue line) over the same month. January and December 1996 are the months with maximum recharge in all recharge subzones. It is observed that, generally, there is a high discrepancy between the maximum aggregated daily recharge (green bar) and monthly recharge (yellow bar), the former exceeding the later (upper graph in Figure 9). The same is found by other authors (De Silva 2010; USGS 2008) and is explained by Scanlon et al. (2002); if water balance is calculated on a daily time step, precipitation sometimes greatly exceeds evapotranspiration on a single day and extreme precipitations events responsible for recharge are dampened out when averaging over longer time periods.

Analyzing median monthly recharge (lower graph in Figure 9), the same result as for maximum monthly recharge is found for some of the subzones: there is an underestimation in recharge results when large time steps are used (e.g. Cabecera Norte Rocina). However, for other subzones, median monthly recharge is higher than median aggregated daily recharge (e.g. Cabecera Sur Rocina subzone). This is contrary to what it can be seen in a wet period or in other aforementioned studies (De Silva 2010; USGS 2008). The reason for these differences when comparing the time scale effect on recharge estimation can be attributed to the rainfall intensity in the studied period and the water balance method used. This method uses a “residual” approach whereby all the variables in the water budget equation are measured or estimated, except recharge (Scanlon et al. 2002). The recharge estimate is set equal to the “residual” and hence it depends on the accuracy with which the other components in the water budget equation are interpolated. This limitation is critical when the magnitude of the recharge rate is small compared to that of the other variables as evapotranspiration (Scanlon et al. 2002). In each time step, the method computes the amount of water exceeding evapotranspiration and field capacity storage as ER. The soil water storage status depends on the situation after the previous step. This is ideal when the balance is performed on a daily basis, but misleading results may be obtained when using the method on aggregated monthly data. The effect of using either monthly rainfall data or daily to monthly aggregation on monthly recharge calculations is appreciated when there are a number of days in the month with low precipitation, lower than evapotranspiration, which do not produce any recharge. However, if this rain is aggregated to a monthly scale before computing the recharge, this value will be higher than evapotranspiration and non zero recharge might appear when it did not exist on the daily scale. To show the influence of this time scale effect in dry periods, non recharge months have been calculated for both recharge estimated series: monthly and aggregated daily to month. Differences in the number of months of zero recharge values were plotted in Figure 10. Daily recharge results (added to month scale) register between 2 and 38 more months with zero recharge than monthly recharge results (Figure 10). This indicates that monthly recharge estimations could be greatly overestimated in dry periods.

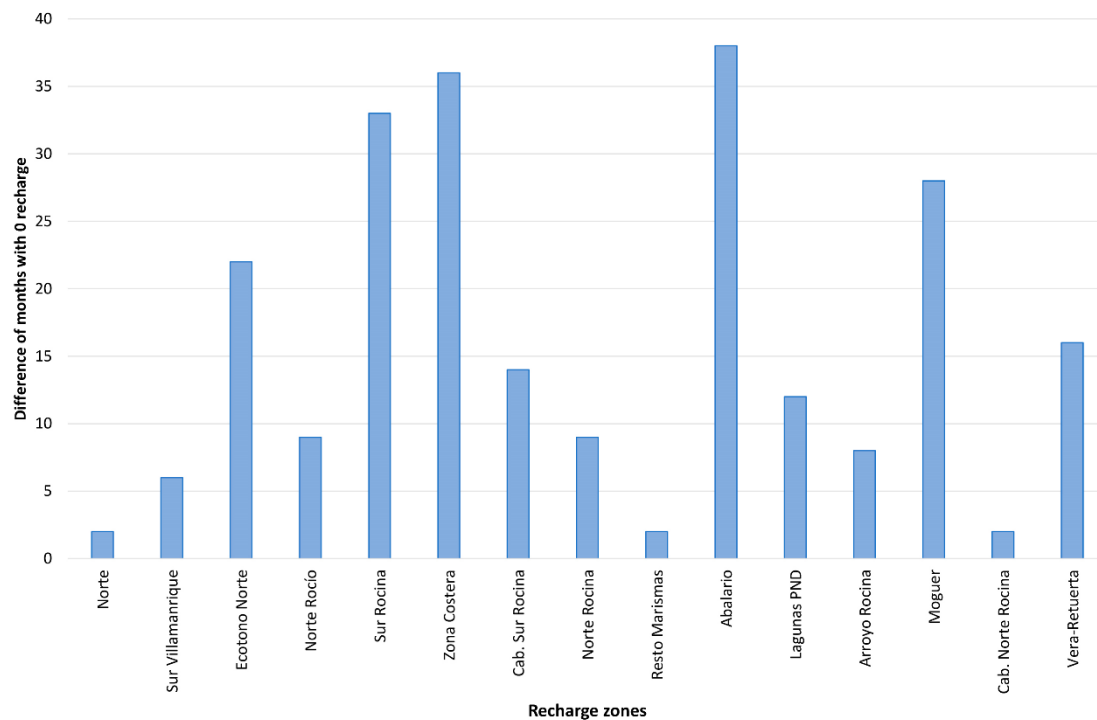


Figure 10. Difference of number of months with no recharge between recharge from daily data and recharge from monthly data.

This variability of relationships between daily and monthly results demonstrates that it is important to choose the appropriate time scale when the available water resources can be underestimated or overestimated depending on the time scale and the rainfall intensity in the studied period. An incorrect choice of the time scale in the calculation of water balances could lead to inappropriate water management decisions and could increase aquifer overexploitation (De Silva and Rushton 1996). This discrepancy can be even higher in the case of Doñana as HP uses historical mean values to compute groundwater reserves, instead of using the previous year's water balance. To face this issue, over the last three years the CHG have supported their decisions in La Rocina groundwater body with updated monthly recharge time series. These groundwater recharge estimations could be improved by using updated daily recharge values. Besides, the same should be done in the rest of the groundwater bodies existing in the Almonte-Marismas Aquifer which maintain all the Doñana ecosystems.

In addition, recharge values show a significant average difference in both the total amount of groundwater recharge and the number of days with recharge higher than 0.01 mm/day in two periods (Figure 11). Between 1975-1998 and 1999-2016 this difference is confirmed by paired samples t-tests ($t_{\text{amount}} = 11.3$, $p < 0.001$; $t_{\text{days}} =$

8.6, $p < 0.001$). Varouchakis et al. (2018) and Narjary et al. (2014) found similar trends in their studies on the rain days in Greece and India, respectively. In the Shaying River (China), Shin et al. (2013) found positive trends of the rainfall concentration index, indicating an increase in the uneven distribution of rainfall and drought occurrences. Zubieta et al. (2016) found that, in the Mantaro River basin (Peru), rainfall concentration index series do not show a significant trend. The decline in the number of rain days but no reduction of total annual rainfall implies an increase in rainfall intensities, leading to a reduction in soil moisture content, higher surface runoff and sediment losses (Narjary et al. 2014). This denotes a similar recharge decrease than it is seen here. All these changes could generate longer drought periods and strongly reductions of groundwater recharge, so a reduction of water resources as well.

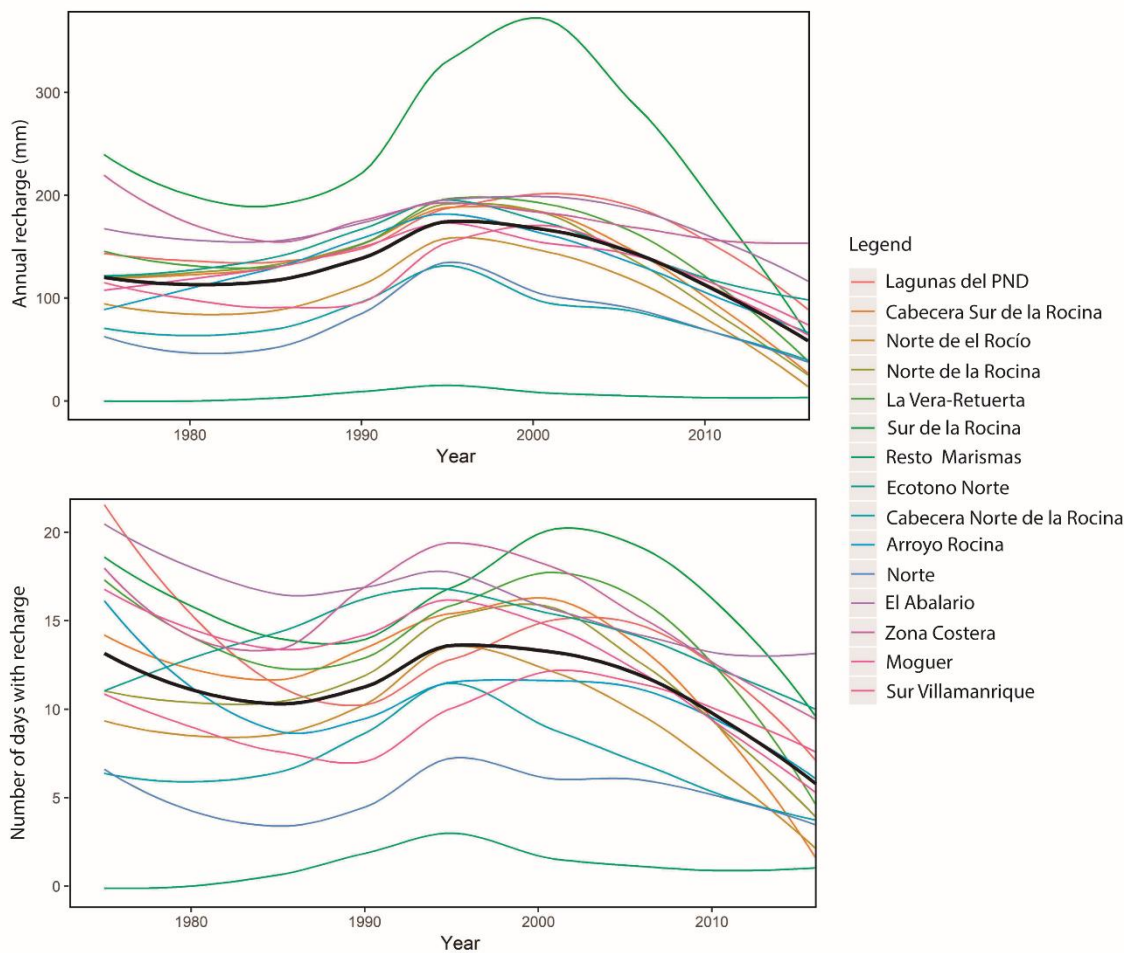


Figure 11. Loess adjustment curves of total amount of groundwater recharge calculated and the number of days with recharge per natural year in period 1975-2016. Grey values correspond to the 15 recharge zones values series. The black line is the average curve taken from the 15 loess curves.

5. Conclusions

In this study case, monthly recharge values range between 21% and 91% of the maximum rainfall. This confirms that, even in this small, flat area, recharge estimations are spatially variable due to different soil properties, land uses and rainfall spatial heterogeneities. Besides, the temporal scale is regarded as crucial to recharge assessments, especially considering the foreseeable effects of climate change on arid regions, namely an increase of daily temperature and precipitation extremes. When mean values of 1975-1995 and 1996-2016 are compared, the 40-year period covered shows a significant decrease of the annual recharge and the number of days with effective recharge. The distribution of the same amount of annual rain in less rain days means that the difference in the calculation between monthly and daily recharge can be higher. One of the adaptations to climate change that water management agencies should implement is to calculate the available water resources with a daily time step. In the present work it has been demonstrated that a small amount of daily rain does not produce recharge when computed in monthly time steps. Consequently, there can be an overestimation of the available water resources. There are also underestimations in recharge calculations when using monthly time steps, that can be attributed to extreme precipitations events, responsible for recharge, that are dampened out when averaging over longer time periods. In the light of these conclusions, new water management plans in Doñana need to take into account finer spatio-temporal scales for efficient water allocation.

Acknowledgements: We thank the reviewers and editors for their remarks that helped to improve the manuscript. This research has been funded by CLIGRO Project (MICINN, CGL2016-77473-C3-1-R) of the Spanish National Plan for Scientific and Technical Research and Innovation and it is also part of the activities subsidized within the National System of Youth Guarantee (MINECO activity with reference PEJ-2014-85121 and Ministry of Education, Youth and Sport of the Community of Madrid with ref. PEJ15/AMB/AI-0218)) co-financed under the Youth Employment Operational Program, with financial resources from the Youth Employment Initiative (YEI) and the European Social Fund (ESF).

Appendix 1

Table 2. Summary statistics of 112 rain gauges used for the spatio-temporal interpolation.

RainGauge	σ	Med.	Max.	Min.	1st Q.	3rd Q.	Var.	SD	%MV
5787E	1.47	0.00	28.00	0.00	0.00	0.00	29.36	5.52	99.82
A22SA	1.47	0.00	99.00	0.00	0.00	0.10	32.63	5.71	57.68
ALAce	1.64	0.00	101.50	0.00	0.00	0.00	37.46	6.12	25.32
ALBeL	2.54	0.00	85.00	0.00	0.00	0.00	64.43	8.03	96.41
ALCab	1.44	0.00	65.00	0.00	0.00	0.00	33.38	5.78	89.03
ALCdA	1.68	0.00	129.00	0.00	0.00	0.00	46.50	6.82	47.05
ALDo2	1.64	0.00	90.00	0.00	0.00	0.00	38.16	6.18	46.67
ALDon	1.51	0.00	90.00	0.00	0.00	0.00	33.17	5.76	23.69
ALERo	3.35	0.00	72.20	0.00	0.00	0.00	95.68	9.81	98.82
ALJdA	1.66	0.00	93.00	0.00	0.00	0.20	37.38	6.11	62.45
ALjSV	2.32	0.00	70.00	0.00	0.00	0.00	49.03	7.01	94.84
ALLCan	1.56	0.00	102.00	0.00	0.00	0.00	34.01	5.83	65.81
ALLMe	2.24	0.00	72.00	0.00	0.00	0.00	52.14	7.22	93.71
ALLRo	1.82	0.00	96.00	0.00	0.00	0.00	44.17	6.65	87.59
ALMar	1.52	0.00	110.00	0.00	0.00	0.00	36.63	6.05	67.69
ALMim	2.10	0.00	125.00	0.00	0.00	0.00	59.23	7.70	42.78
ALRAi	1.02	0.00	70.20	0.00	0.00	0.10	29.26	5.42	97.86
Areno	1.16	0.00	93.60	0.00	0.00	0.00	21.18	4.60	55.52
AzCan	2.09	0.00	60.00	0.00	0.00	0.00	39.41	6.28	91.81
AzCsB	1.68	0.00	123.80	0.00	0.00	0.00	42.87	6.55	65.23
AzDeN	1.61	0.00	87.00	0.00	0.00	0.00	38.46	6.20	5.78
AzDhP	1.49	0.00	32.00	0.00	0.00	0.00	17.51	4.19	97.90
AzFAO	1.87	0.00	102.50	0.00	0.00	0.00	49.73	7.05	90.47
AzHRa	1.80	0.00	80.00	0.00	0.00	0.00	35.88	5.99	86.44
AzJdA	1.47	0.00	100.00	0.00	0.00	0.20	31.76	5.64	60.90
AzLBi	2.45	0.00	65.00	0.00	0.00	0.00	56.88	7.55	96.61
AzLJu	1.63	0.00	112.50	0.00	0.00	0.00	39.12	6.25	6.03
AzLTi	1.63	0.00	54.00	0.00	0.00	0.00	36.49	6.04	92.02
BMAEM	1.60	0.00	140.00	0.00	0.00	0.00	38.15	6.18	2.09
BoEvi	1.49	0.00	65.00	0.00	0.00	0.00	28.59	5.35	85.68
BolCo	1.67	0.00	96.70	0.00	0.00	0.00	40.27	6.35	53.26
BolMC	1.71	0.00	120.00	0.00	0.00	0.00	51.33	7.17	82.64
BolMi	2.17	0.00	70.00	0.00	0.00	0.00	50.01	7.08	92.61
Bonar	1.58	0.00	114.00	0.00	0.00	0.00	40.49	6.36	30.69
BoRai	1.59	0.00	97.80	0.00	0.00	0.10	36.70	6.06	61.46
C06SA	1.09	0.00	147.50	0.00	0.00	0.00	23.45	4.84	51.09
C07SA	1.15	0.00	131.50	0.00	0.00	0.10	24.39	4.94	59.54
C08SA	1.09	0.00	76.50	0.00	0.00	0.00	20.33	4.51	52.49
C09SA	1.18	0.00	93.10	0.00	0.00	0.00	24.05	4.90	53.97
C11SA	1.20	0.00	81.50	0.00	0.00	0.00	23.28	4.83	51.09
C13SA	1.07	0.00	119.50	0.00	0.00	0.00	23.96	4.89	53.76
C14SA	1.25	0.00	95.60	0.00	0.00	0.00	25.12	5.01	50.85
CaCe2	1.60	0.00	98.50	0.00	0.00	0.00	39.76	6.31	43.82
CaCes	1.98	0.00	85.00	0.00	0.00	0.00	52.09	7.22	93.21
CamEC	1.76	0.00	109.10	0.00	0.00	0.00	41.95	6.48	33.15
CasCu	2.55	0.00	60.50	0.00	0.00	0.00	67.64	8.23	97.21

RainGauge	σ	Med.	Max.	Min.	1st Q.	3rd Q.	Var.	SD	%MV
ChECA	1.41	0.00	126.00	0.00	0.00	0.00	29.75	5.46	79.25
Chipi	1.49	0.00	108.00	0.00	0.00	0.00	33.52	5.79	24.21
CoEda	1.62	0.00	140.00	0.00	0.00	0.00	39.93	6.32	37.15
CoRSe	1.80	0.00	93.00	0.00	0.00	0.00	41.80	6.47	66.00
DHTor	1.68	0.00	110.00	0.00	0.00	0.00	40.44	6.36	57.84
EBDon	1.51	0.00	90.00	0.00	0.00	0.00	34.30	5.86	13.34
EspLo	2.13	0.00	60.20	0.00	0.00	0.00	44.78	6.69	86.78
EsRai	1.40	0.00	101.20	0.00	0.00	0.20	28.31	5.32	61.18
GinCo	1.94	0.00	148.00	0.00	0.00	0.00	49.54	7.04	43.06
GTAEM	2.64	0.00	45.50	0.00	0.00	0.00	54.53	7.40	98.62
HinCA	4.34	0.00	40.50	0.00	0.00	5.38	82.25	9.24	99.82
Hinoj	1.79	0.00	87.00	0.00	0.00	0.00	42.97	6.56	45.44
HiToC	2.60	0.00	90.00	0.00	0.00	0.00	69.30	8.33	92.82
Hu11	3.23	0.00	94.00	0.00	0.00	0.20	120.48	11.00	98.63
Hueva	1.88	0.00	70.10	0.00	0.00	0.00	45.33	6.73	84.00
IFAEC	1.58	0.00	93.60	0.00	0.00	0.20	37.11	6.09	64.00
IFALP	1.51	0.00	95.40	0.00	0.00	0.20	29.69	5.45	76.83
IFChi	1.40	0.00	114.00	0.00	0.00	0.20	30.04	5.48	70.12
IFLTC	1.63	0.00	93.60	0.00	0.00	0.20	39.09	6.25	62.40
IMJdA	1.40	0.00	85.80	0.00	0.00	0.20	30.15	5.49	69.58
LCINM	1.77	0.00	110.20	0.00	0.00	0.00	43.96	6.63	1.75
LCJdA	1.69	0.00	82.40	0.00	0.00	0.00	39.48	6.28	60.62
Leb_I	1.50	0.00	104.40	0.00	0.00	0.20	30.92	5.56	60.37
LebII	1.47	0.00	66.80	0.00	0.00	0.20	29.24	5.41	87.57
LPR_I	1.49	0.00	93.40	0.00	0.00	0.20	31.98	5.66	60.40
LPRAi	1.34	0.00	80.20	0.00	0.00	0.20	27.46	5.24	61.35
LPRiB	0.87	0.00	22.30	0.00	0.00	0.00	10.17	3.21	99.42
LPRII	1.46	0.00	144.60	0.00	0.00	0.20	31.41	5.60	60.77
LPVil	1.82	0.00	65.00	0.00	0.00	0.00	45.40	6.74	95.43
LpyVF	1.49	0.00	103.40	0.00	0.00	0.20	30.97	5.57	83.85
Lucen	1.58	0.00	153.00	0.00	0.00	0.00	41.18	6.42	25.40
M09SA	1.57	0.00	111.70	0.00	0.00	0.00	36.77	6.06	58.10
MaAst	1.72	0.00	112.00	0.00	0.00	0.00	44.95	6.70	32.79
MOINM	1.53	0.00	106.10	0.00	0.00	0.00	37.79	6.15	5.66
MoJdA	1.55	0.00	92.80	0.00	0.00	0.20	33.79	5.81	60.23
MoMad	1.33	0.00	101.00	0.00	0.00	0.00	32.40	5.69	38.29
MoMaz	2.68	0.00	63.00	0.00	0.00	0.00	51.73	7.20	94.22
MoRAi	1.32	0.00	70.50	0.00	0.00	0.10	36.54	6.05	97.70
MoSho	1.23	0.00	72.00	0.00	0.00	0.00	27.56	5.25	76.33
Niebl	1.62	0.00	168.20	0.00	0.00	0.20	38.78	6.23	60.61
PadLF	1.37	0.00	115.10	0.00	0.00	0.00	35.32	5.94	69.24
PFAEM	1.34	0.00	45.00	0.00	0.00	0.00	18.26	4.28	92.81
PFPit	1.43	0.00	132.20	0.00	0.00	0.00	34.18	5.85	42.03
PiMed	1.73	0.00	97.00	0.00	0.00	0.00	41.73	6.46	7.08
PiSan	1.83	0.00	103.20	0.00	0.00	0.00	54.74	7.40	56.44
RonEs	1.43	0.00	160.00	0.00	0.00	0.00	31.70	5.63	2.35
SaLMa	1.51	0.00	84.00	0.00	0.00	0.20	31.56	5.62	62.05
SBBGu	1.74	0.00	135.50	0.00	0.00	0.00	51.83	7.20	74.00
SBCdA	1.64	0.00	109.70	0.00	0.00	0.00	43.73	6.61	51.88
SBINM	1.50	0.00	145.40	0.00	0.00	0.00	34.48	5.87	33.43
SBJAB	1.65	0.00	121.00	0.00	0.00	0.20	40.66	6.38	89.96

RainGauge	σ	Med.	Max.	Min.	1st Q.	3rd Q.	Var.	SD	%MV
SBJAI	1.41	0.00	104.00	0.00	0.00	0.20	34.60	5.88	80.95
SBPla	1.82	0.00	78.80	0.00	0.00	0.00	33.32	5.77	91.41
SeRAi	1.50	0.00	119.80	0.00	0.00	0.10	35.87	5.99	62.01
SeTab	1.50	0.00	56.80	0.00	0.00	0.00	27.61	5.26	80.50
SLMay	1.81	0.00	117.00	0.00	0.00	0.00	49.01	7.00	44.08
SMCQu	1.56	0.00	91.00	0.00	0.00	0.00	33.79	5.81	63.05
STab1	1.66	0.00	80.30	0.00	0.00	0.00	37.88	6.15	33.15
STab2	1.50	0.00	102.40	0.00	0.00	0.00	36.33	6.03	76.61
Trebu	1.65	0.00	95.00	0.00	0.00	0.00	36.34	6.03	32.31
ViAPu	2.37	0.00	77.00	0.00	0.00	0.00	55.64	7.46	90.52
ViGua	1.93	0.00	50.40	0.00	0.00	0.20	36.52	6.05	97.51
ViIRY	1.02	0.00	83.00	0.00	0.00	0.00	20.53	4.53	89.23
ViICA	1.60	0.00	73.00	0.00	0.00	0.00	36.38	6.03	75.19
ViJu	1.49	0.00	104.20	0.00	0.00	0.00	35.17	5.93	31.28
ViILM	1.90	0.00	87.00	0.00	0.00	0.00	50.44	7.10	79.64

These statistics were computed from the measured period in each of the rain gauges: mean (m), median (MED.), maximum (MAX.), minimum (MIN.), first quartile (1st Q.), third quartile (3rd Q.) and standard deviation (SD).

Appendix 2

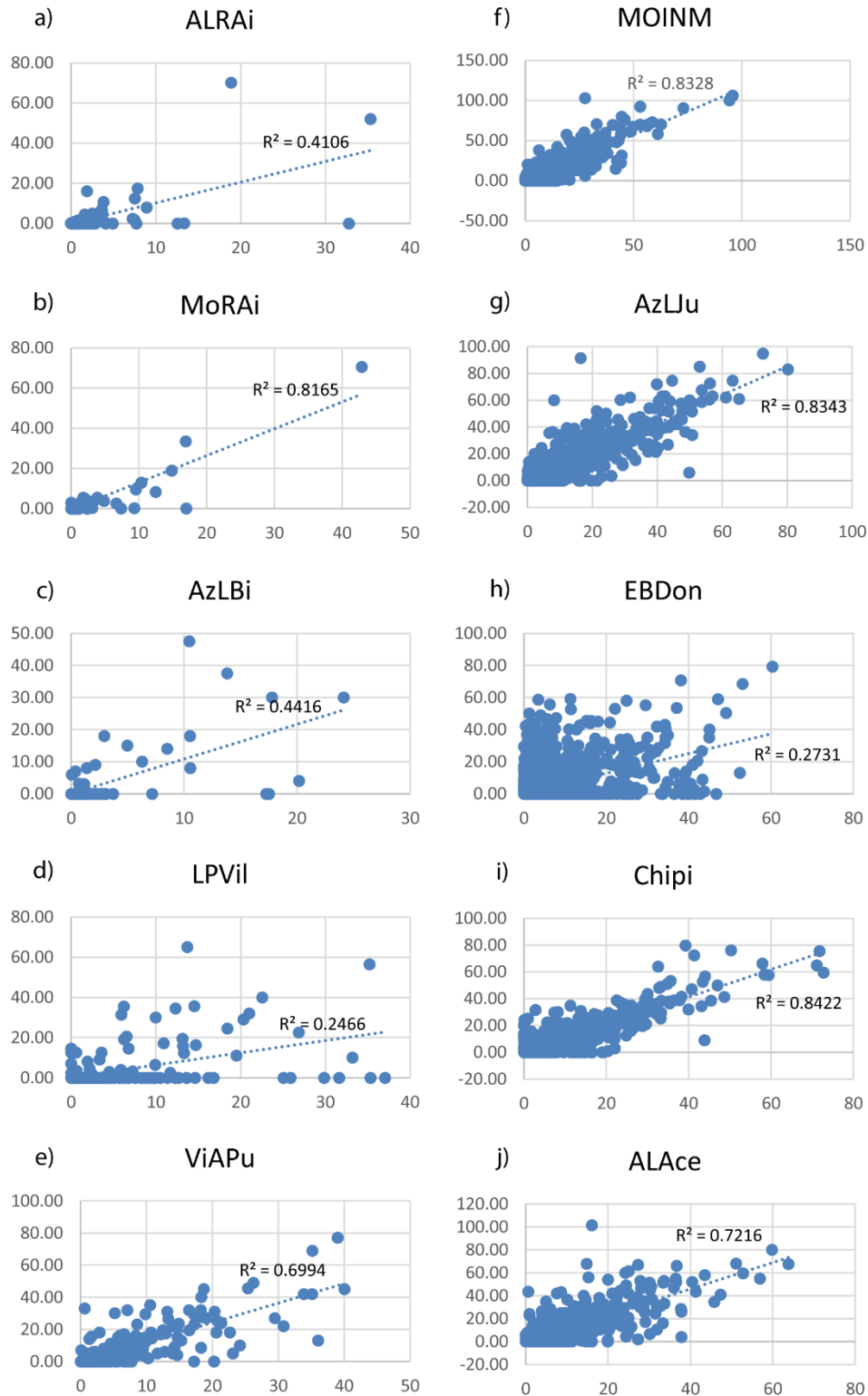


Figure 12. Observed versus predicted daily rainfall for the 10 rain gauge stations chosen to test the goodness of the interpolation. Figures a to e correspond to the five rain gauges with a high number of missing values (90%). Figures f to j represent five rain gauges with low degrees of missing information (up to 25%).

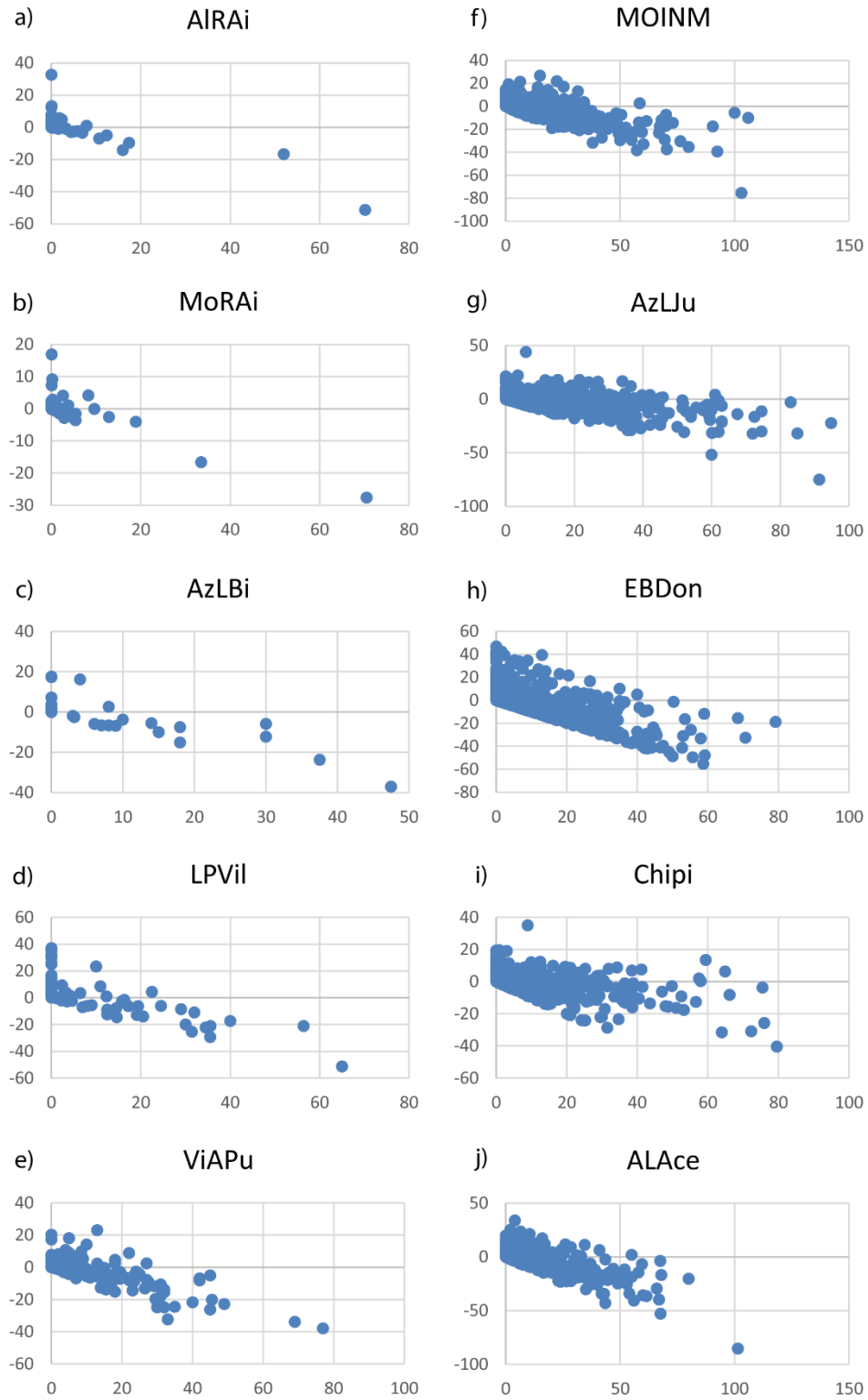


Figure 13. Daily rainfall bias plots for the 10 rain gauge stations chosen to test the goodness of the interpolation. Figures a to e correspond to the five rain gauges with a high number of missing values (90%). Figures f to j represent five rain gauges with low degrees of missing information (up to 25%).

Appendix 3

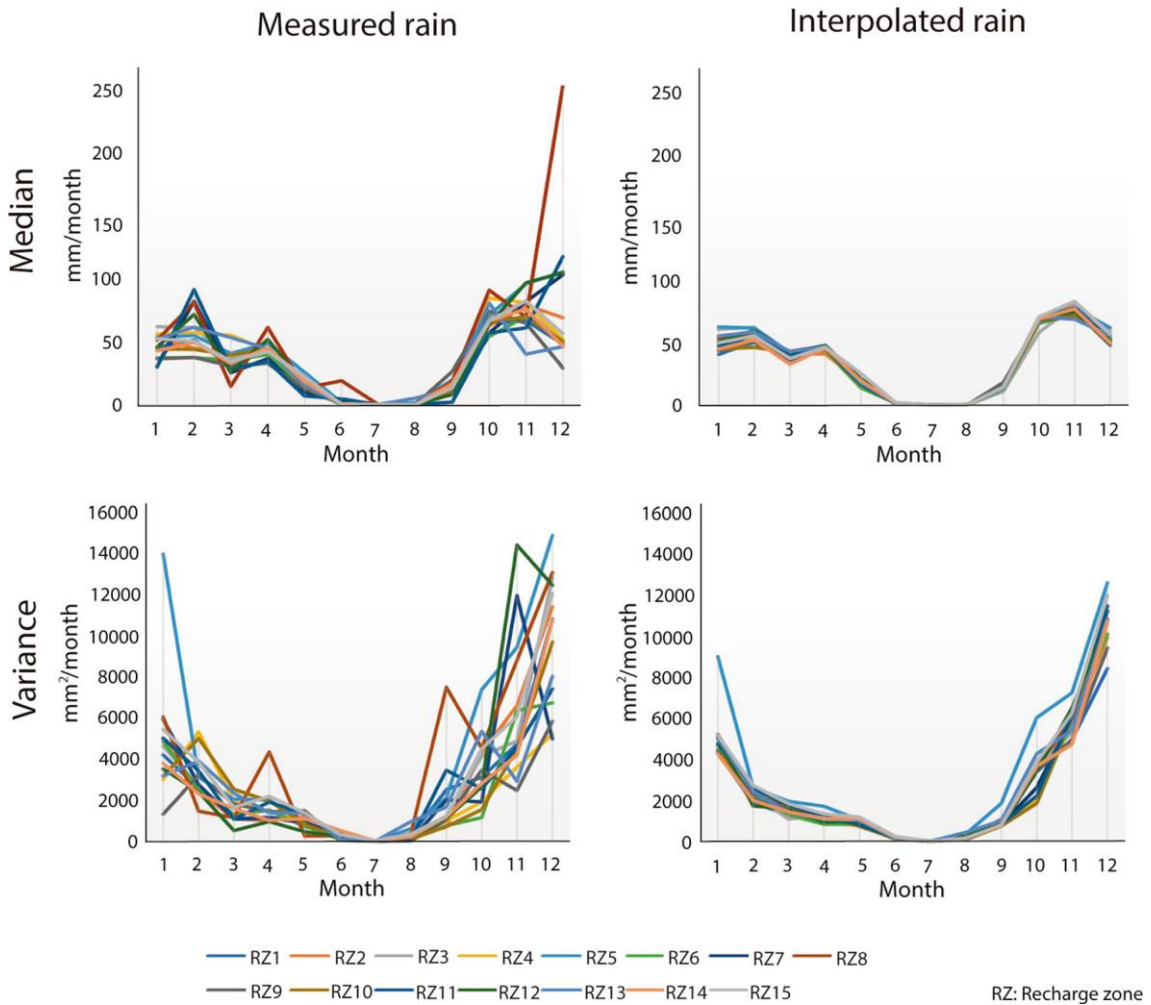


Figure 14. Median and variance of monthly rainfall, when 20 or more days are measured in each month of the year. Values are computed over the studied period for every rain gauge laying inside each recharge zone defined in Figure 4.

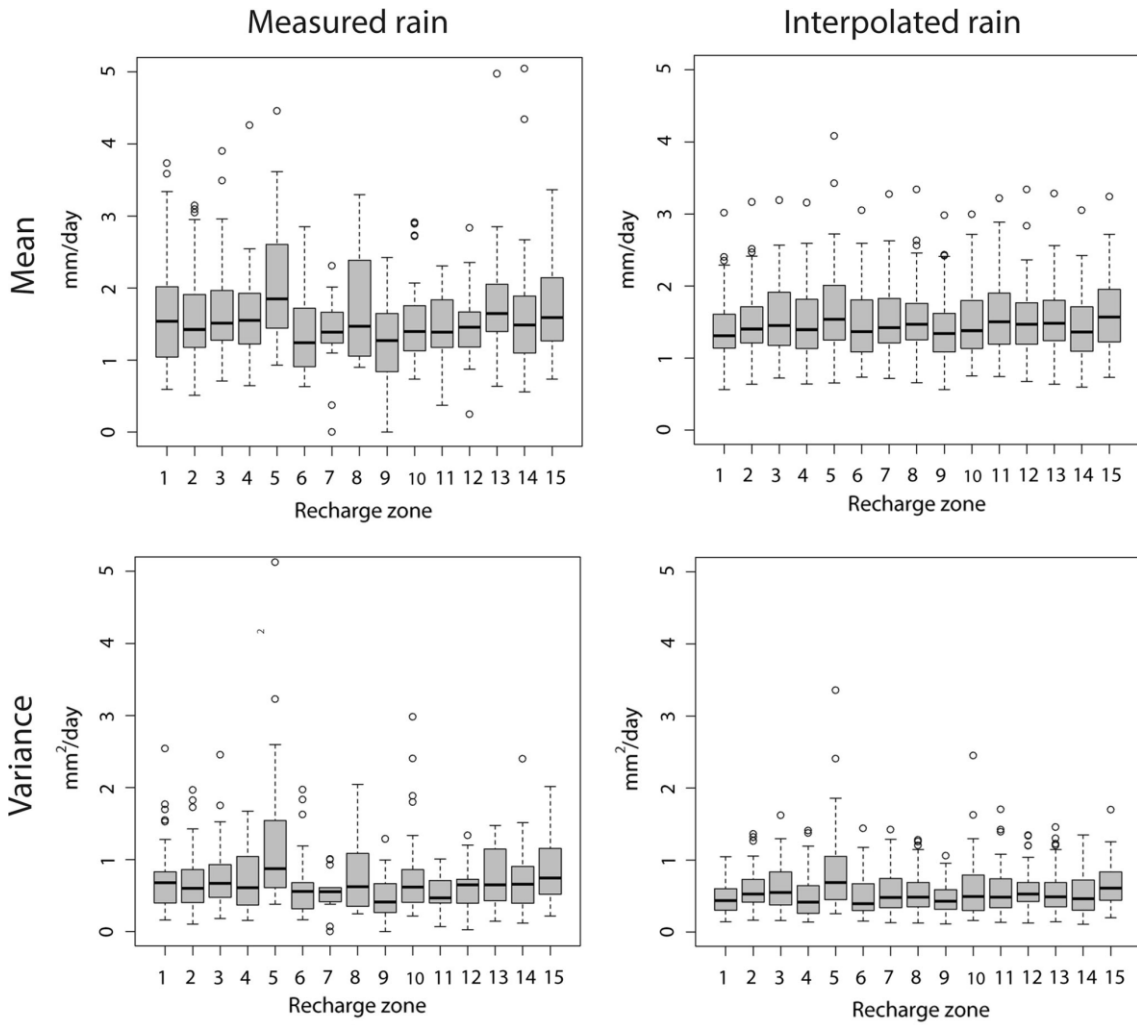


Figure 15. Mean and variance of daily values for every year of the studied period averaging the values of rain gauges lying inside each recharge zone defined in Figure 4.

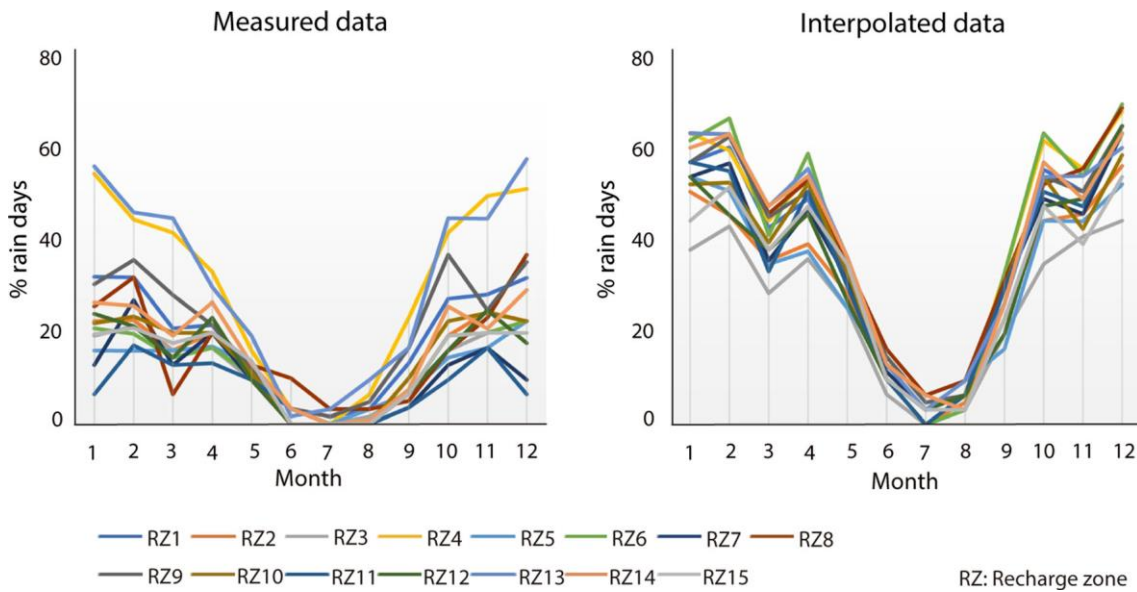


Figure 16. Median percentage rate of rain days in a month and the number of days with data in that month. This ratio was computed for all rain gauges lying inside each recharge zone defined in Figure 4.

Artículo III

Con el permiso de los coautores, esta sección reproduce el contenido del artículo:

Naranjo-Fernández, N., Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Montero-González, E. (2020) Clustering groundwater level time series of the exploited Almonte-Marismas aquifer in Southwest Spain. *Water*, 12, 1063, doi: 10.3390/w12041063.

Clustering groundwater level time series of the exploited Almonte-Marismas aquifer in Southwest Spain

Agrupación de series temporales de niveles de agua subterránea en el explotado acuífero de Almonte-Marismas en el Sureste de España

Nuria Naranjo Fernández¹, Carolina Guardiola Albert², Héctor Aguilera², Carmen Serrano Hidalgo² y Esperanza Montero González¹

1. Facultad de Ciencias Geológicas. Universidad Complutense de Madrid. Calle José Antonio Novais 12, 28040, Madrid; nuria.naranjo.phd@gmail.com; emontero@ucm.es;

2. Instituto Geológico y Minero de España. Calle Ríos Rosas - 23 - 28003 Madrid – España. c.guardiola@igme.es

* Corresponding author: c.guardiola@igme.es

Abstract

Groundwater resources are regularly the principal water supply in semiarid and arid climate areas. However, groundwater levels (GWL) in semiarid aquifers are suffering a general decrease because of anthropic exploitation of aquifers and the repercussions of climate change. Effective groundwater management strategies require a deep characterisation of GWL fluctuations, in order to identify individual behaviours and triggering factors. In September 2019, the Guadalquivir River Basin Authority (CHG) declared that there was over-exploitation in three of the five groundwater bodies of the Almonte-Marismas aquifer, Southwest Spain. For that reason, it is critical to understand GWL dynamics in this aquifer before the new Spanish Water Resources Management Plans (2021-2027) are developed. The application of GWL series clustering in hydrogeology has grown over the past few years, as it is an extraordinary tool that promptly provides a GWL classification; each group can be related to different responses of a complex aquifer under any external change. In this work, GWL time series from 160 piezometers were analysed for the period 1975-2016 and, after data pre-processing, 24 piezometers were selected for clustering with k-means (static) and time series (dynamic) clustering techniques. Six and seven groups (k) were chosen to apply k-means. Six characterized types of hydrodynamic behaviours were obtained with time series clustering (TSC). Number of clusters were related to diverse affections of water

exploitation depending on soil uses and hydrogeological spatial distribution parameters. TSC enabled us to distinguish local areas with high hydrodynamic disturbance and to highlight a quantitative drop of GWL during the studied period.

Keywords: groundwater level hydrographs; k-means clustering; time series clustering; water resource management

Resumen

Las aguas subterráneas son normalmente la principal fuente de suministro en zonas de climas áridos y semiáridos. Sin embargo, los niveles piezométricos en acuíferos semiáridos están sufriendo un descenso general debido a la explotación antrópica y a las repercusiones del cambio climático. Las estrategias de la gestión de las aguas subterráneas requieren una caracterización de la fluctuación de los niveles piezométricos, con la finalidad de identificar comportamientos individuales y los factores desencadenantes. En septiembre de 2019, la Confederación Hidrológica del Guadaquivir (CHG) declaró sobreexplotadas tres de las cinco masas de agua subterránea que componen el acuífero Almonte-Marismas, suroeste de España. Por este motivo, es crítico entender la dinámica de los niveles piezométricos en este acuífero antes de desarrollar los nuevos planes de gestión (2021-2027). La aplicación de las técnicas de agrupamiento en series piezométricas ha crecido en los últimos años ya que es una herramienta extraordinaria para la clasificación de niveles piezométricos; cada grupo pertenece a un grupo de respuestas obtenidas por cambios externos en el acuífero. En este estudio, las series temporales de 160 piezómetros han sido analizadas para el periodo entre 1975-2016, después de ser pre-procesadas, 24 piezómetros han sido seleccionados para realizar un cluster de k-medias (estático) y un cluster de series temporales (dinámico). Seis y siete grupos (k) han sido elegidos para aplicar k.medias. Seis comportamientos hidrogeológicos se han obtenido con el cluster de series temporales (TSC). El número de grupos está relacionado con las diversas afecciones producidas por la explotación de las aguas subterráneas dependiendo de los usos del suelo y de la distribución espacial de los parámetros hidrogeológicos. TSC nos permitió distinguir áreas con elevadas alteraciones en la hidrodinámica local y subrayar cuantitativamente los descensos de niveles piezométricos en el periodo estudiado.

Palabras clave: Hidrogramas del nivel piezométrico; cluster k-medias; cluster de series temporales; gestión de recursos hídricos.

1.Introduction

The increase of groundwater exploitation for intensive agricultural irrigation contributes to transformations in the natural hydrodynamics of a given area. Several studies have shown strong changes in natural groundwater flow in over-exploited areas (Álvarez-Cobelas et al. 2001; Castellazi et al. 2016; Lubis 2017). Climate change drives increasing modifications in groundwater uses. In many cases, the decrease of groundwater levels (GWL) involves ecological degradation, especially in wetland areas (Custodio 2002; Diaz-Paniagua et al. 2010). Effective groundwater management strategies require a deep characterisation of GWL fluctuations, in order to identify individual behaviours and triggering factors.

Doñana Natural Space, located in the southwest of Spain, is one of the largest protected wetlands in Europe (Green et al. 2017). The study of historical piezometric databases could be a particularly challenging work in the Doñana wetland area because of heterogeneity in the records and a high number of missing values. While data were formerly obtained manually, piezometers have recently evolved to be part of an automated system, continuously recording GWL fluctuations. This means that the size of the dataset has increased, making the management and interpretation of the GWL data an important issue.

Clustering has frequently been applied in hydrology, e.g. in hydrochemistry studies to distinguish hydrochemical groups of water (Irawan et al. 2009). Besides this, it has been used to classify GWL hydrographs (Bloomfield et al. 2015), to study the changes in GWL produced after an earthquake in Japan (Nakagawa et al. 2019) and to evaluate subsurface flow patterns (Rinderer et al. 2019). Recently clustering was also used to study hydrochemical impacts on groundwater by altering the interaction between groundwater and surface water (Yuan et al. 2020) and to evaluate the missing data of monitoring wells (Ashgarnina & Petroselli 2020).

There are many clustering techniques but k-means is one of the most popular, unsupervised algorithms applied in the geosciences and, also, in hydrogeology (Marin Celestino et al. 2018). The aim of k-means is to divide n observations into k groups. It is relatively simple to implement, scales to log database and guarantees convergence

(Bottou & Bengio 1995). Nevertheless, the number of clusters, k , has to be pre-defined and the method is very sensitive to outliers. In addition, k -means does not consider the temporal correlation structure of the time series, which is an important limitation to detect changes in temporal patterns in different groups. To solve this limitation and handle dynamic data, time series clustering (TSC) techniques have been developed, modifying the similarity definition of the data or the prototype extraction function to an appropriate one (Sardá-Espinosa 2019).

In September 2019, the Guadalquivir Basin Authority (CHG in Spanish acronym) activated the administrative process to declare 3 of the 5 groundwater bodies that comprise the Almonte-Marismas aquifer to be at risk of not reaching a good quantitative state. For that reason, it is critical to understand GWL dynamics in this aquifer before the new Spanish Water Resources Management Plans (2021-2027) can be developed. Qualitative studies of the historical GWL data in the Almonte-Marismas aquifer exist (UPC 1999; Olías Álvarez & Rodríguez Rodríguez 2013; CHG 2019) but they examined the data by visual inspection and none of them give a clear categorisation of GWL behaviour. In general, to analyse the characteristics and summarise the patterns of the complex fluctuation of GWL, it is important to rely on objective classifications (Nguyen et al. 2015).

The consequences of climate change, such as increasing evaporation or more extreme hydrological events (e.g. droughts), aggravate the anthropic pressure, posing a serious threat to the Almonte-Marismas aquifer's hydrogeological state (Guardiola-Albert & Jackson 2011). The over-exploitation of the area is attributed to uncontrolled, unsustainable and, in many cases, illegal growth of intensively irrigated agriculture over recent decades (WWF 2016). Unregulated and illegal pumping must be taken into account in groundwater management evaluations (Erostate et al. 2020). When long GWL time series are available, one way to evaluate the agricultural expansion effect on the aquifer dynamics is a deep study of the GWL by using clustering tools.

In the present work, the aim is to apply clustering methodologies to improve the understanding of the Almonte-Marismas aquifer dynamics, obtaining information that will be useful to develop sustainable groundwater resource management plans. For that purpose, monthly GWL from 1975 to 2016 were clustered using three different

methods: visual observation (as a reference for a qualitative analysis), k-means (MacQueen 1967) and TSC (Sardá-Espinosa 2019). To the best of the authors' knowledge, this is the first time that TSC has been applied to a medium size GWL database with a 41 year time series.

2. Materials and Methods

2.1. Study site

The Almonte-Marismas aquifer, in Southwest Spain, is located at the margins of the mouth of the Guadalquivir River (Figure 1a). This aquifer hosts the Doñana marshlands, which was flooded by groundwater and surface water and declared to be RAMSAR wetlands in 1982 (Custodio et al. 2009). At the same time, Doñana water supports the socio-economic activities developed on the area, such as agriculture, industry, mining and tourism.

geological terms, it is composed of a complex sedimentary system of sand and gravels with some clay lenses, basal sands, silts and aeolian sands (Salvany & Custodio, 1995; Salvany et al. 2011; Figures 1a and 1b). The average precipitation between 1975 and 2016 was approximately 550 mm/year and the average annual temperature was 25 °C. The maximum daily precipitation was registered at 168.2 mm in November 2001. From June to September the maximum temperature was 45 °C and the minimum temperature in winter months was 6 °C. The sub-humid Mediterranean climate with an Atlantic influence results in high seasonality (Custodio et al. 2009). The main recharge sources are rainfall, irrigation return flow and lateral inflow from the northeast. The average recharge is 200 hm³/year, primarily during spring and autumn (Custodio et al. 2009). The main groundwater flow runs from northwest to southeast, recharging mainly from the northern and southwestern areas and discharging into the Atlantic Ocean onto the marshland, rivers and streams that flow from the higher regions in the north (Figure 1c). There is no control on how much water is withdrawn for agriculture but there are estimated increases in groundwater abstraction for irrigation, from 11 hm³/year at the beginning of the 1970s to 80 hm³/year in the mid-1990s, continuing to increase up to more than 130 hm³/year, currently (CHG 2019).

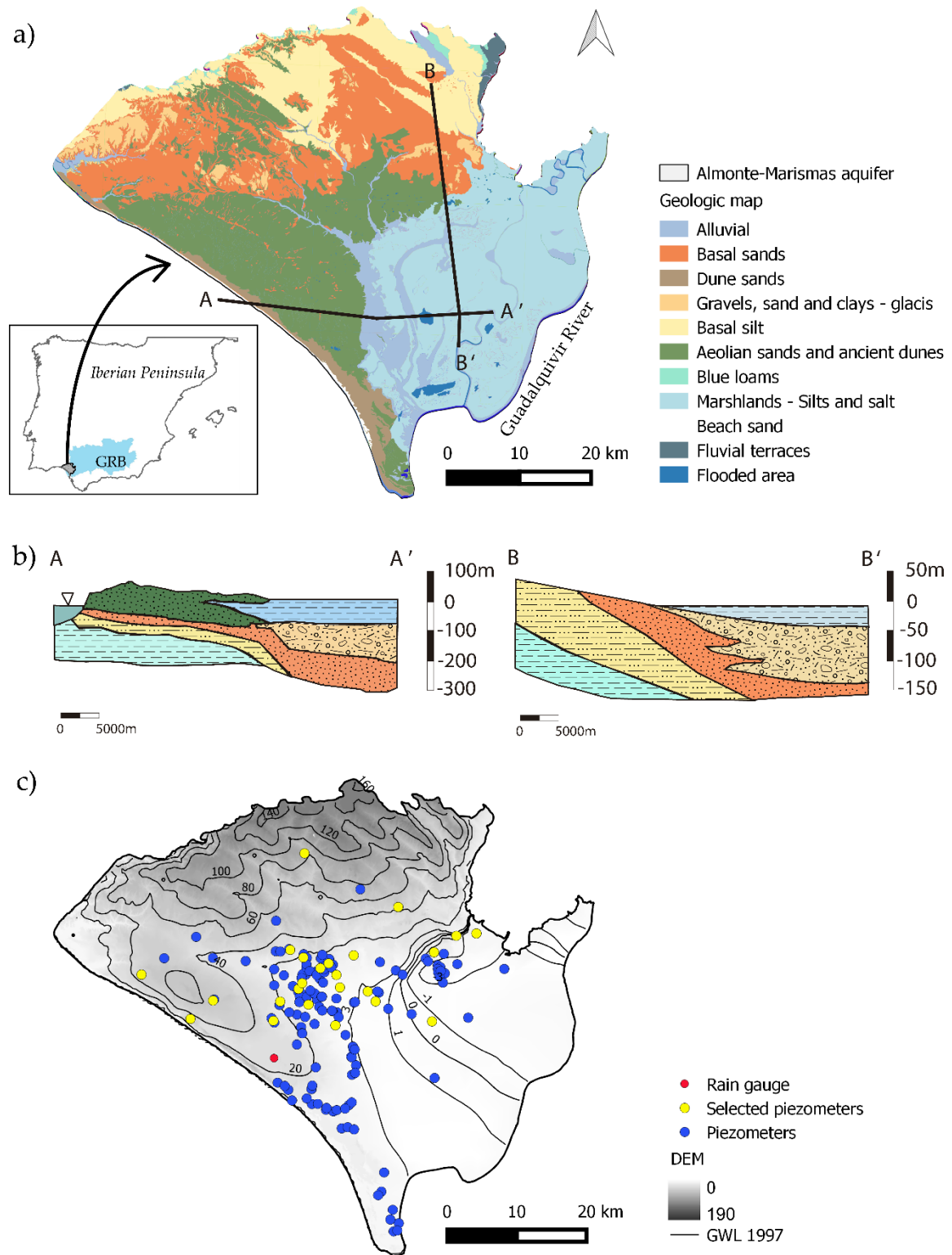


Figure 1. a) Location and geological map. b) Cross sections of Almonte-Marismas aquifer (Guadalquivir River Basin, GRB). c) Piezometers and Almonte-Acebucho rain gauge locations and piezometric surface at the end of 1997, taken from Guardiola-Albert et al. 2009. DEM: Digital Elevation Model.

The topography of the region falls from approximately 150 m above sea level (masl) in the north, to less than 1 masl in the marshland area on the coast to the south. In

The Almonte-Marismas aquifer's complex sedimentary origins, ecological richness and groundwater over-exploitation are three challenging factors for groundwater resource management decisions (Salvany et al. 2011; Carnicer i Cols 2007). The combined effect of agricultural activities, industry and tourism have added to climate change and led to a general decrease of GWL, changing the hydrodynamics in the area (Custodio et al. 2009). As a result, several groundwater-dependent temporal ponds have disappeared (Díaz-Paniagua et al. 2008).

2.2. Data collection and pre-treatment

Groundwater monitoring has historically been carried out in the Almonte-Marismas aquifer since 1903. Unfortunately, this GWL series does not retain continuity. Piezometers were first installed in 1970 (Custodio et al. 2009), making it possible to have an entire database with 160 points (blue dots in Figure 1c) with different degrees of temporal continuity. Data from these piezometers (with less than 80% missing values (MV)), have been collected by the Spanish Geological Survey (IGME) and the CHG and have been used in the present work (yellow dots in Figure 1c). The period of time chosen, from October 1975 to September 2016, follows a hydrological year and registers the most data recorded. Agricultural activity grew considerably in the mid-1990s. Consequently, in order to analyse the evolution of the aquifer dynamics due to anthropic actions, the study period was subdivided into two sub-periods of twenty and twenty-one hydrological years: (1) October 1975 – September 1995 and (2) October 1995 – September 2016.

The measurements, as an absolute elevation above or under sea level, have been homogenised to a monthly time step, taking the average of all data recorded within each month. In addition, outliers and duplicates were discarded.

To study GWL responses to precipitation, rainfall data were taken from the Almonte-Acebucho rain gauge (Figure 1c), which is maintained by the AEMET (Spanish State Meteorological Agency). Cumulative deviation from the mean monthly rainfall (CDR) was calculated to represent precipitation variations in both sub-periods.

Figure 2 shows the whole approach that was followed in this study. To define a starting point (and before applying more advanced statistical techniques), visual classification of

the piezometric time series was performed. This exercise is essential to developing a global understanding of the time series.

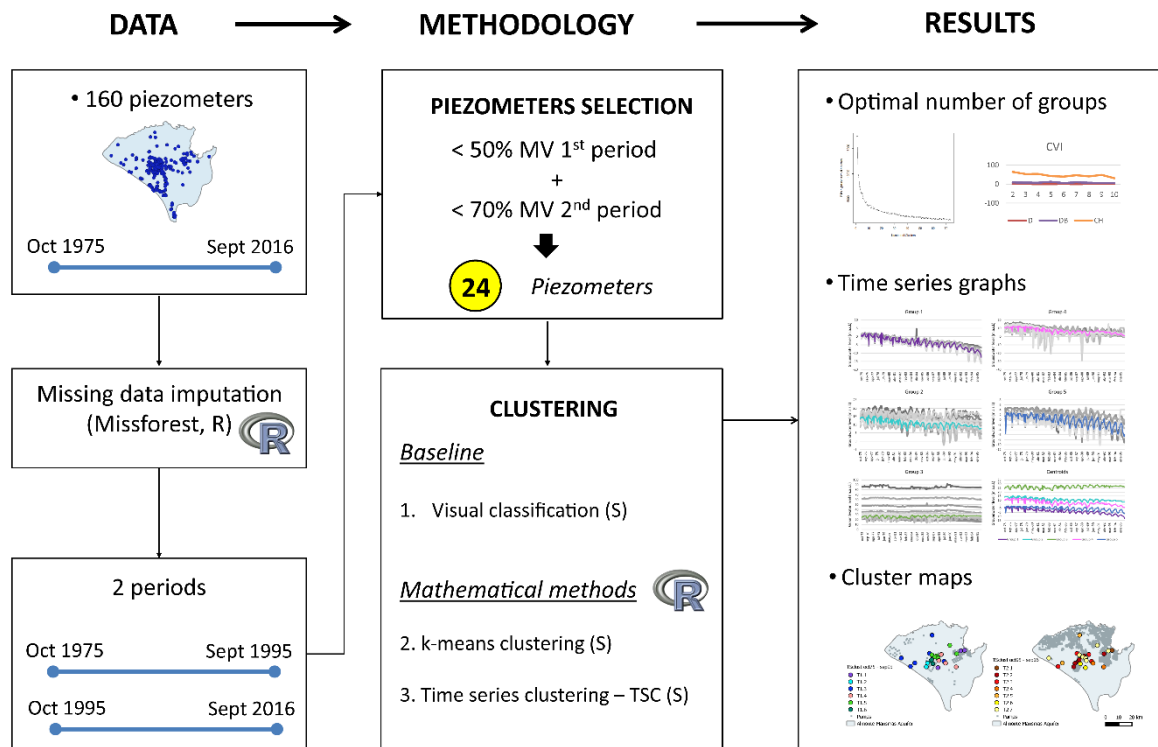


Figure 2. Flowchart of the methodological approach used in the current study. S: standardisation as pre-treatment.

MV in the GWL time series have been infilled using all the available information (160 piezometers) and the missForest algorithm (Stekhoven & Bühlmann 2012). MissForest is a non-parametric, iterative imputation method, based on the random forest algorithm (Breiman 2001). The advantage of missForest is that it does not make any assumptions about the distribution of data or imputation models (Aguilera et al. 2019). It uses an iterative imputation scheme by training a random forest on the observed values of each variable in a first step. Then, it predicts the missing values and proceeds iteratively until the stopping criterion is met or the user-specified maximum number of iterations is reached. The algorithm continuously updates the imputed matrix variables, assessing its performance between iterations. This assessment is done by considering the difference between the previous and the new imputation results. As soon as this difference increases for the first time, the algorithm stops and keeps the last imputation matrix. The Normalized Root Mean Square Error (NRMSE; equation 1) has been obtained to evaluate the performance of the method:

$$NRMSE = \sqrt{\frac{\text{mean}(X^{true} - X^{imp})^2}{\text{var}(X^{true})}} \quad (1)$$

where X^{true} is the complete data matrix, X^{imp} is the imputed data matrix, $mean$ and var are, respectively, the short notation for the empirical mean and variance computed over the continuous missing values only. Good performance leads to a value of NRMSE close to 0 and bad performance results in a value approaching 1.

To perform a consistent GWL clustering it was decided to only use piezometers with a low percentage of MV in both study sub-periods (i.e. October 1975 – September 1995 and October 1995 – September 2016). The average MV percentage for the entire study period was $T = 51.5\%$. For the first and second sub-periods the average MV percentages were $P_1 = 64.1\%$ and $P_2 = 39.5\%$, respectively. That is the reason why a more restrictive MV threshold was applied in order to select which piezometers were used in the cluster analysis in the first sub-period, compared to the second one. The following expressions (relating P_1 and P_2 to T (equations 2 and 3)) were used to decide which MV thresholds define what piezometers from the 160 available data points remained for the cluster analysis:

$$MV_1 = P_1(2 - r_1) \quad (2)$$

$$MV_2 = P_2(1 + r_2) \quad (3)$$

where $r_1 = P_1/T$ and $r_2 = P_2/T$.

Substituting P_1 , P_2 , r_1 and r_2 for their corresponding values in the equations yields an MV_1 threshold of 48.4% for the first sub-period and an MV_2 threshold of 69.8% for the second sub-period. Rounding up these figures, the piezometers that have less than 50% MV for the first sub-period and 70% MV for the second sub-period have been kept, resulting in 24 piezometers in total (yellow dots in Figure 1c). The general statistics for these 24 piezometers can be seen in Appendix 1-Table 2.

To avoid the scaling effects of GWL values while clustering, the imputed monthly piezometric series were standardised. This step was carried out before visual and k-means clustering and it is already implemented in the algorithm used to perform TSC.

2.3. *k-means clustering*

K-means clustering is one of most common methods for exploratory data analysis. This technique can identify groups in the dataset that show similar patterns or equal trends. Specifically, k-means divides up the data into pre-defined, non-overlapping groups, where every data point belongs to only one group or cluster. First of all, the random partition method randomly assigns a cluster to each observation and then the algorithm proceeds to the update step, computing the initial mean to be the centroid of the cluster's randomly assigned points.

The standard algorithm of k-means (equation (4); Hartigan & Wong 1979), minimises the sum of squares of the distance between data points within a cluster and the centroid of that cluster:

$$W(C_K) = \sum_{X_i \in C_k} (x_i - \mu_k)^2 \quad (4)$$

In the present case study, x_i is the data vector with monthly measurements corresponding to piezometer i , X_i is the set of piezometers that belongs to cluster C_K and μ_k is the centroid, computed as the mean value of the data vector belonging to cluster C_K . Each data vector (x_i) is assigned to a cluster when its distance to the centroid, $W(C_K)$, is at a minimum.

The appropriate number of clusters is given by the point where the sedimentation graph changes the trend, drawing an “elbow” (Ketchen & Shoock 1996). The sedimentation graph plots a clustering score as a function of the number of clusters and the elbow method is a heuristic method for interpreting and validating consistency. The score is generally a measure of the input data on the k-means objective function (e.g. sum of $W(C_K)$ across clusters). The k-means algorithm was run with the R package factextra (Kassambara & Mundt 2017).

2.4. Time series clustering (TSC)

The TSC method considers time series distribution. As with the k-means method, it has a stochastic nature due to its random start. TSC was run with the R package *dtwclust* (Sardá-Espinosa 2019).

In order to classify all piezometers, partitional clustering (Paparrizos & Gravano 2015) was selected to create partitions. The specification of the number of clusters (k) is tedious work that must be performed before starting the algorithm. The optimal k was identified by using internal clustering validity indices (CVIs). Cluster validity is a set of techniques that have been developed to find the best k without any previous class information. Based on similar works (Maulik 2002; Krszczuk & Hurley 2010) y, three internal CVI indexes were used to choose k : Dunn (D), Davies-Bouldin (DB) and Calinski-Harabasz (CH). D is a ratio-type index and its cohesion is calculated by the distance to its nearest neighbour and the separation by the maximum cluster diameter (Dunn 1973). DB is one of the most popular indices and it was based on the distance between centroids (Davies & Bouldin 1979). CH cohesion is estimated based on the distance from the points in a cluster to its centroid (Calinsky & Harabasz 1974).

Depending on the distance of the disturbance and the aquifer characteristics, GWL responses can be registered with a lag of days or months from one location to another. Considering this singularity of GWL behaviour, the most appropriate distance to measure GWL time series is Shape-Based Distance (SBD) (Sardá-Espinosa 2019). SBD warps the time series and it is based on the cross-correlation with coefficient normalisation ($NCCC$) and shape extraction as a prototype function (5) [35]:

$$SBD(x, y) = 1 - \frac{\max(NCCC(x, y))}{\|x\|_2 \|y\|_2} \quad (5)$$

where $\| \cdot \|_2$ denotes the standard L2 norm of the series and $SBD(x, y)$ values range between 0 (perfect similarity) and 2 (unequal time series).

3. Results and discussion

3.1. Imputation piezometry and visual classification

Monthly imputation of the 41 year time series for the 24 piezometers with the missForest algorithm resulted in an NRMSE of 0.05, indicating a good performance of the imputation scheme. This small error can mainly be explained by the lack of uneven piezometer variations at a monthly scale, which is present in other kinds of time series (e.g. temperature), and by the high number of piezometers available for imputation.

Table 1 shows how many piezometers integrate every cluster for each sub-period. Figure 3 represents the spatial distribution of the different clusters obtained. Visual classification results for both study sub-periods are shown in the first row of Figures 3a and 3b and in Figures 4 and 5. A layer corresponding to a rough spatial distribution of the groundwater extractions in the two different sub-periods (Guardiola-Albert et al. 2009; Manzano-Arellano et al. 2009) is also included in Figure 3 (in grey). To ease the interpretation, the raw piezometric time evolution for each visual cluster is shown in Figure 4 (clusters named V1.1, V1.2, V1.3 and V1.4 for the first sub-period) and Figure 5 (clusters named V2.1, V2.2, V2.3 and V2.4 for the second sub-period). Piezometric time series are plotted in the primary Y axis with grey lines; the mean of each GWL cluster is shown in the secondary Y axis, with coloured lines.

It is difficult to compare our GWL classes (Figures 3 and 4) with the ones obtained from the 15 year series through visual inspection by UPC (UPC 1999) and, in addition, those published in Manzano et al. 2002. These authors only showed GWL evolution at 7 points within the same study area but without trying to find analogous piezometric evolution at different points.

Visual analysis found four different groups for both sub-periods. All visual cluster groups have a different spatial distribution when comparing the first and second sub-periods studied (Figures 3a and 3b). The main difference between the two sub-periods is that the clustering spatial distribution is more irregular in the second one. This could be attributed to an increase in the number of pumping wells (grey zones in Figure 3), in

addition to climate change effects that could reduce the recharge rate (Naranjo-Fernández et al. 2020).

At the end of the first sub-period there is a generally descending trend of GWL and a maximum average local oscillation of 5 m (Figure 4). The series that comprise groups V1.2 and V1.3 have piezometers with GWL between 0 and 80 masl, while V1.1 and V1.4 have a lower range, between -10 and 30 masl. This reflects the fact that series with visually similar GWL oscillations are located in different areas of the aquifer (Figure 3a). All of the groups exhibit time variations that may be due to extractions, as well as the recharge-discharge pattern characteristics of the hydrological cycle (e.g. compared with the CDR line). Visually, it looks like V1.1 and V1.3 are highly affected by pumping due to their cyclical oscillation over a year, ranging from 2 to 6 m. Group V1.4 seems to be located in an area exploited in the first half of the sub-period, but not in the second period. Until 1990, cluster V1.2 fluctuations followed the same temporal evolution that Custodio (Custodio 2002) assigned to undisturbed GWL. Furthermore, the reduced influence of anthropic actions is corroborated by the fact that piezometry has the same trend as the CDR curve (Figure 4). Among the four cluster groups, V1.4 presents the highest average level decrease of around 7 m (from 1975 to 1995). The main feature revealed by these results is that, between 1975 and 1995, anthropic impacts were already affecting the aquifer as GWL was not recuperating after the pumping cycle.

For the second sub-period, Figure 5 shows that the order of the groups from highest to lowest inter-annual oscillation is V2.2, V2.1, V2.3 and V2.4. All groups have a descending trend but at varying degrees. The V2.4 graph (Figure 5) shows the same trend as the CDR and the smallest decrease (of up to 2 m) reflects low anthropic intervention. V2.1, V2.2 and V2.3 have GWL between -10 and 40 masl and GWL in V2.4 is between 20 to 90 masl because of the altitude at the different locations. As in the first study sub-period, the wide GWL ranges are a consequence of the classification criteria (GWL behaviour) and not just because of location or average values. For example, V2.4 presents a greater GWL range because two piezometers (yellow dots in Figure 3b) are located at the highest topographical area and groundwater table. V2.3 is the group with the highest drawdown in this second sub-period, up to 6 m on average. Observed drawdowns are in accordance with the piezometric descent that was estimated for the whole of the

Almonte-Marismas aquifer between 1994 and 2012 (Olías-Álvarez & Rodríguez-Rodríguez 2013); this ranged from 2 to 6 m. Cluster V.2.1 shows high fluctuations of up to 4 m, attributed to pumping effects. However, time variations are similar to the ones seen in CDR, which also shows the climate response. This feature could be explained by merging climate oscillations and pumping effects seen at the locations in this cluster V2.1 (Figure 3).

Visual cluster results allow us to have an overview of the piezometer responses along each sub-period. Also, this approach is a qualitative baseline for evaluating and comparing the results of the mathematical methods.

Table 1. Number of piezometers in each obtained cluster. Sub-period 1: from October 1975 to September 1995. Sub-period 2: from October 1995 to September 2016.

Method	Sub-period	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
<i>Visual</i>	1	10	4	7	3	-	-	-
	2	10	5	6	3	-	-	-
<i>k-means</i>	1	1	2	5	1	3	12	-
	2	1	8	4	1	5	1	4
<i>TSC</i>	1	3	2	2	3	4	6	4
	2	1	6	4	2	2	7	2

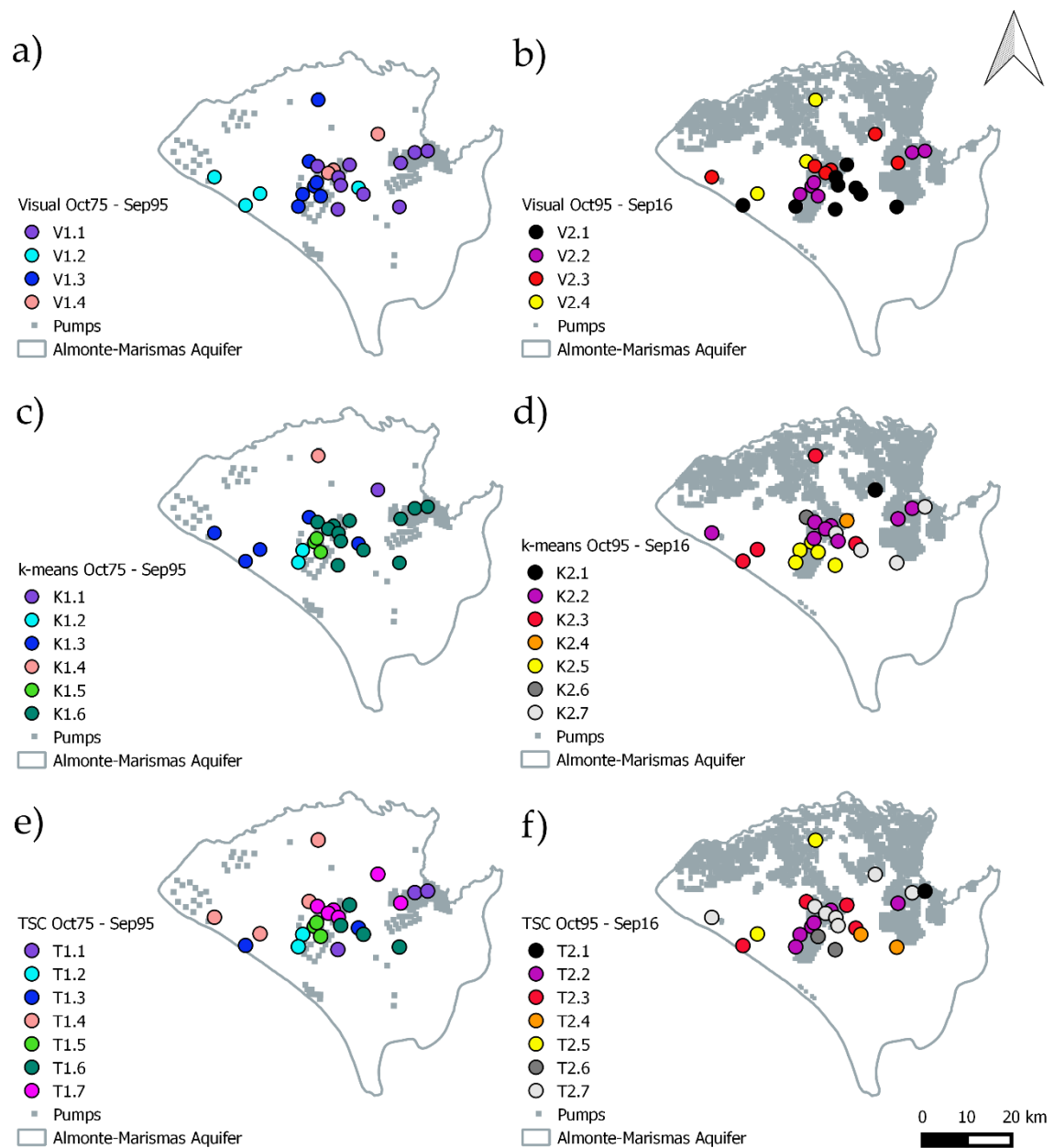


Figure 3. Spatial distribution of piezometer groups for the three clustering methods applied. An approximate distribution of the pumping wells [Guardiola-Albert et al. 2009, Manzano-Arellano et al. 2009] is shown in grey. **a:** Visual clustering for first sub-period (October 1975 to September 1995). **b:** Visual clustering for second sub-period (October 1995 to September 2016). **c:** k-means clustering for first sub-period. **d:** k-means clustering for second sub-period. **e:** TSC for first sub-period. **f:** TSC for second sub-period.

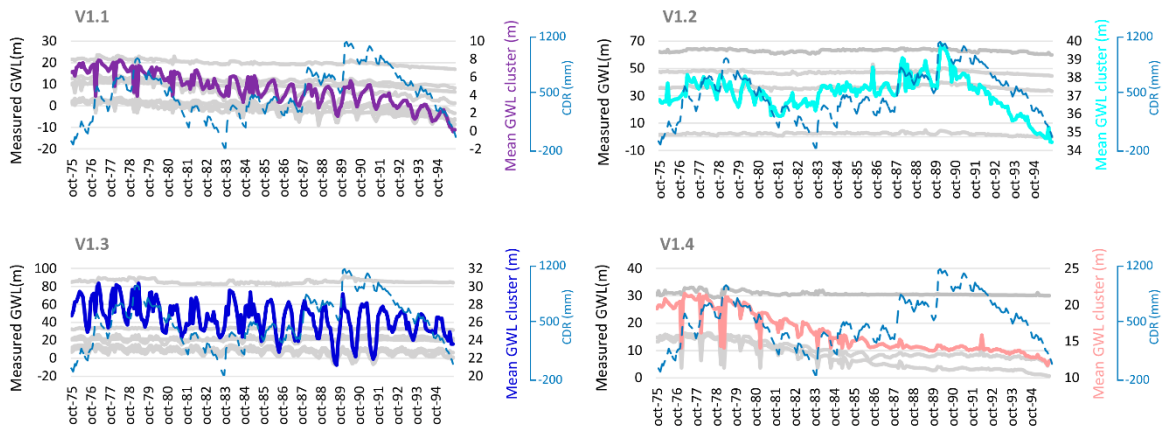


Figure 4. GWL series (grey lines), mean of the GWL cluster series (coloured lines) from October 1975 to September 1995 by visual clustering and cumulative deviation from the mean monthly rainfall (CDR).

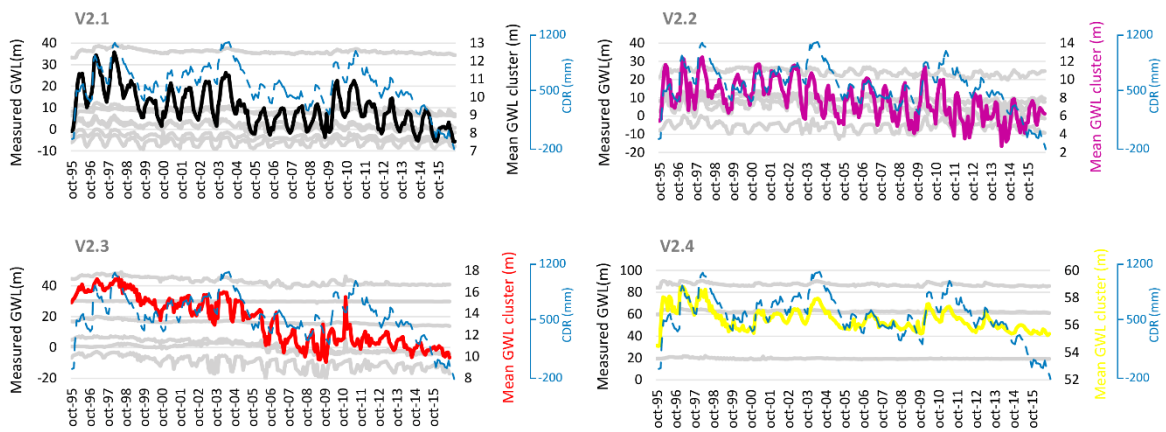


Figure 5. GWL series (grey lines), mean of the GWL cluster series (coloured line) from October 1995 to September 2016 by visual clustering and cumulative deviation from the mean monthly rainfall (CDR).

3.2. k-means clustering

As explained in the Section 2, when applying k-means clustering, the number of clusters k has been defined with the aid of the sedimentation graph (Figure 6). Both sedimentation curves for first (dark blue) and second (red) study sub-periods change their trend at the point which indicates the optimal number of clusters (i.e. the elbow). Both curves show smoothly decreasing trends and, given that the maximum number of clusters is 24, points around the second inflexion point in each period were tested and a final value of k chosen as a trade-off between complexity and representativeness. Consequently, $k=6$ and $k=7$ were taken to be the optimal number of clusters to run the k-means algorithm for the first and the second study sub-periods, respectively (Figure

6). Initially, it can be interpreted that the k-means method distinguishes more GWL variation groups that were found by visual classification. However, two and three groups are formed by just one piezometer in each sub-period. These clusters with a single piezometer correspond to singular time series variations that k-means cannot group.

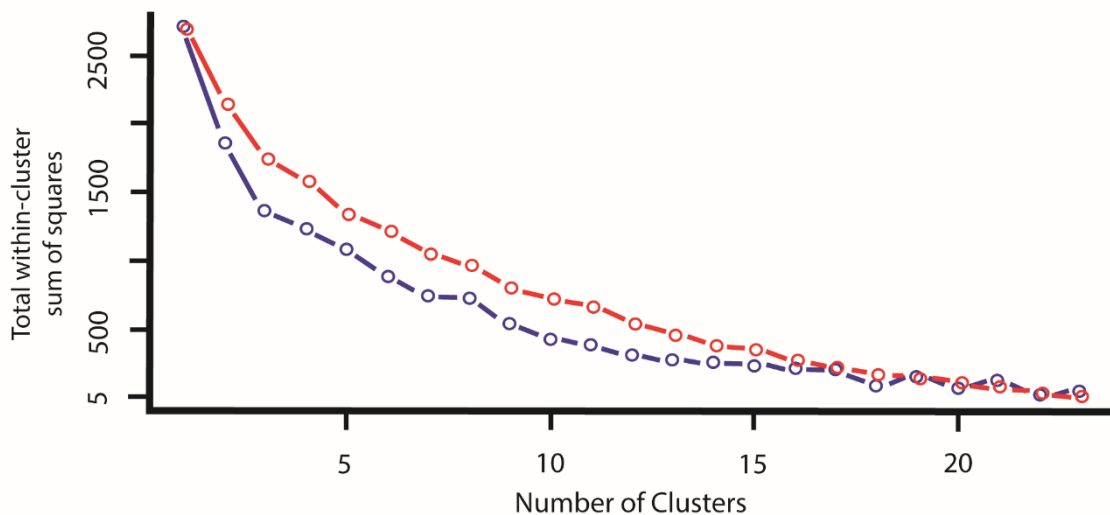


Figure 6. *k-means sedimentation graphs of cluster sub-periods. Blue dots represent the first sub-period (from October 1975 to September 1995) and red dots the second sub-period (from October 1995 to September 2016).*

As a visual classification, k-means has more spatial continuity in the cluster groups in the first sub-period than in the second period (Figures 3c and 3d). For both sub-periods and each k-means cluster, Figures 7 and 8 plot the raw piezometric series (grey), and the average raw series (coloured).

In the first sub-period, K1.3 is the group with the highest GWL range (Figure 7) due to the sparse spatial distribution of the piezometers. For example, there is one point located in the piezometric dome (Figure 1c), where important recharge rates take place. K1.3 is almost the same group as V1.2, with just one piezometer more and the same trend as the CDR. These groups of piezometers do not show fluctuations affected by direct pumping, but a decreasing trend can be seen since 1990. Groups K1.2, K1.5 and K1.6 have GWL between -10 and 30 masl and they correspond to piezometers located in the pumping areas with high anthropic changes in the natural hydrodynamics. K1.2 and K1.5 (Figure 7) have an average time series that can correspond to V1.3 (Figure 4) and K1.6 (Figure 7) corresponds to V1.1 and V1.4 (Figure 4). These groups are affected

by groundwater extractions and are differently classified when using visual or k-means clustering. K1.5 and K1.6 average series show clear cycles that are attributed to the effects of groundwater extraction, as well as an important, descending trend of up to 8 m, on average (Figure 7). Spatially, they are located in an area of high agricultural development (Figure 3c). Finally, both groups K1.1 and K1.4 are composed of a single piezometer (Table 1), located in regions not affected by irrigation pumping during this first sub-period (Figure 3c).

For the second sub-period, it was observed that the GWL variations are greater than the first sub-period (Figure 8). Groups K2.1, K2.4 and K2.6 are composed of a single piezometer (Figure 8, Table 1). K2.1 is the same piezometer as K1.1. It presents a different GWL pattern than the other piezometers (with small oscillations from October 1984) and an important drawdown at the beginning of 2009 (Figures 7 and 8). These abrupt variations suggest that there could be errors in the data acquisition from this piezometer. K2.4 temporal behaviour indicates that aquifer exploitation started at the beginning of 2007 near this point and this could be the reason why the GWL behaviour does not fit with other groups. The opposite behaviour was seen in K2.6, with smooth variations during the second half of the sub-period. K2.3 does not show a clearly descending trend, with piezometers in very different areas of the aquifer (Figure 3d) and appearing as a smoothed signal of the CDR. The average series of K2.2, K2.5 and K2.7 have a decreasing tendency (Figure 8) due to the growing groundwater extractions in the nearby areas (Figure 3d). In particular, the K2.2 group presents the most pronounced drawdowns of up to 8 m, and mainly comprises piezometers located in central areas of the aquifer. It appears to correspond to an area that began to be exploited later than zone K2.5 but, as piezometry suggests, with a greater intensity reflected by a faster descent since the year 2000. The average oscillation patterns are more regular for clusters K2.5 and K2.7; this can be explained by irrigation and geological heterogeneity.

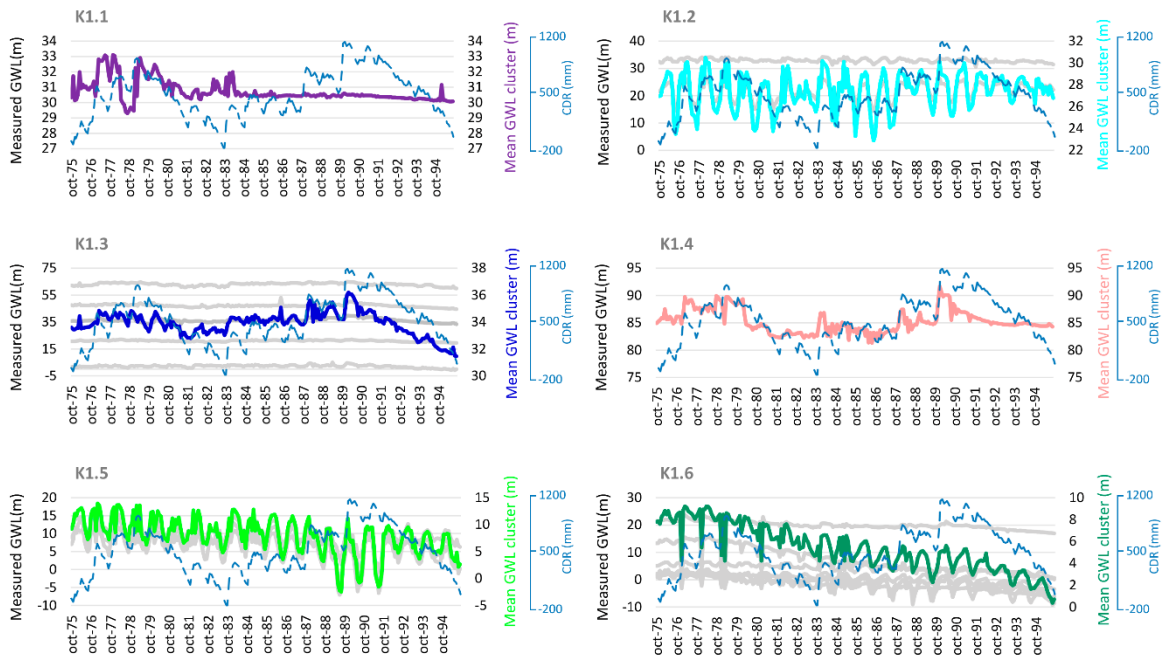


Figure 7. GWL series (grey lines), mean of the GWL cluster series (coloured line) from October 1975 to September 1995 (first sub-period), classified by k-means clustering and cumulative deviation from the mean monthly rainfall (CDR).

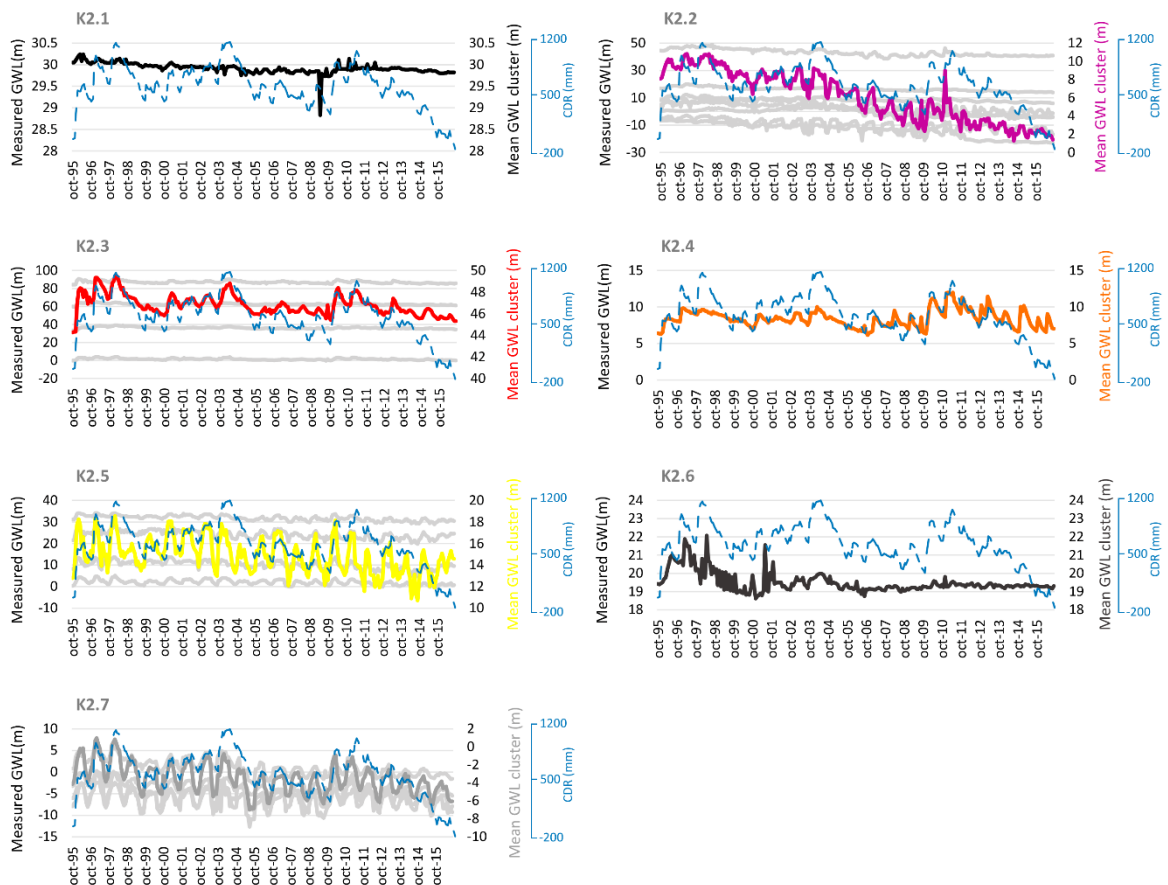


Figure 8. GWL series (grey lines), mean of the GWL cluster series (coloured line) from October 1995 to September 2016 (second sub-period), classified by k-means clustering and cumulative deviation from the mean monthly rainfall (CDR).

When we compare k-means results with visual classification however, some k-means groups represent a merge of hydrogeological behaviour due to exploitation for irrigation of different types of crops (e.g. berries, citrus) or other uses (e.g. drinking water supply), it could be seen in the different patterns of K2.2, K2.5 y K2.7 (Figure 8). On the other hand, groups K1.3 and K2.3 have strong similitudes with the CDR graph. Similar k-means results were found by Boomfield et al. 2015 with a 30-year monthly GWL series. These were in a limestone and chalk aquifer located in the UK. However, they highlight the relationship between GWL and geological characteristics for 74 points, spatially distributed across 2000 km². In the present study (and by comparing cluster spatial locations (Figure 3) with the geological distribution (Figure 1a)), it was observed that GWL are not always dependent on geology. However, cluster groups are in agreement with the findings in Boomfield et al. 2015, regarding the identification of

anthropogenically impacted groundwater hydrographs showing declining trends (e.g. clusters K1.6 or K2.2).

3.3. Time series clustering

Figure 9 shows the CVI values for a different number of clusters in both sub-periods. Following the criteria explained in Section 2, k was set to 7 in both sub-periods; D, DB and CH present an increase (marked with a red line in Figure 9). Only one of the clusters in the second sub-period had just one piezometer.

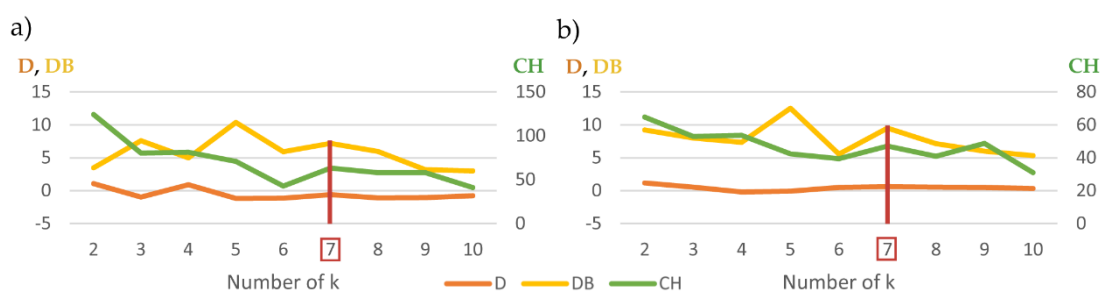


Figure 9. Cluster Validity Indices (CVI): Dunn (D), Davies-Bouldin (DB) and Calinski-Harabasz (CH). X axes represent the number of clusters (k) and Y axes are the sum-of-squares. D and DB are represented in the primary Y axis and CH in the secondary Y axis. **a:** CVI for first sub-period (October 1975 to September 1995). **b:** CVI for second sub-period (October 1995 to September 2016).

For the first sub-period, the TSC map (Figure 3e) reveals a spatial homogeneity for clusters T1.2, T1.4 and T1.5, which are located in the north and centre of the aquifer. The average raw series of groups T1.3 and T1.4 (Figure 10) do not seem to have been directly influenced by agricultural pumping, due to the shape and size of their oscillations and they follow the CDR. They correspond to clusters V1.2 and K1.3, with the same decreasing trend from 1990 (Figures 4, 7 and 10). The T1.1 average time series is very similar to V1.1 and T1.7 to V1.4, both are affected by pumping cycles (Figures 4 and 10). The rest of the groups (T1.2, T1.5 and T1.6) discriminate between different patterns that were not detected by k-means or visual classification. All of them merge annual hydrological cycles and irrigation extractions.

In the second sub-period, the spatial distribution of TSC groups is not homogeneous (Figure 3f). As in the former cluster methods, and regarding cluster series graphs (Figure 11), the general trend in the second sub-period is a GWL decrease. More specifically, T2.2, T2.4, and T2.7 have an outstanding average decrease between 2 and 10 m. Only

T2.1 is composed of one piezometer (Figure 11, Table 1). T2.3 and T2.5 have similar average series and are not spatially adjacent to the irrigation areas (Figure 3f) but are closely related to the CDR. T2.2, T2.4, T2.6 and T2.7 reveal the annual oscillations associated with extractions, with average amplitudes of approximately 4 m (Figure 11). It is difficult to find correlation between these four TSC groups in the second sub-period and the visual clusters. This can be explained by the TSC's ability to integrate different temporal hydrodynamic variations such as increasing pumping, evolution in the types of crops (Fundación Doñana 21 2003) and seasonality of GWL oscillations (Fernández-Ayuso et al. 2019).

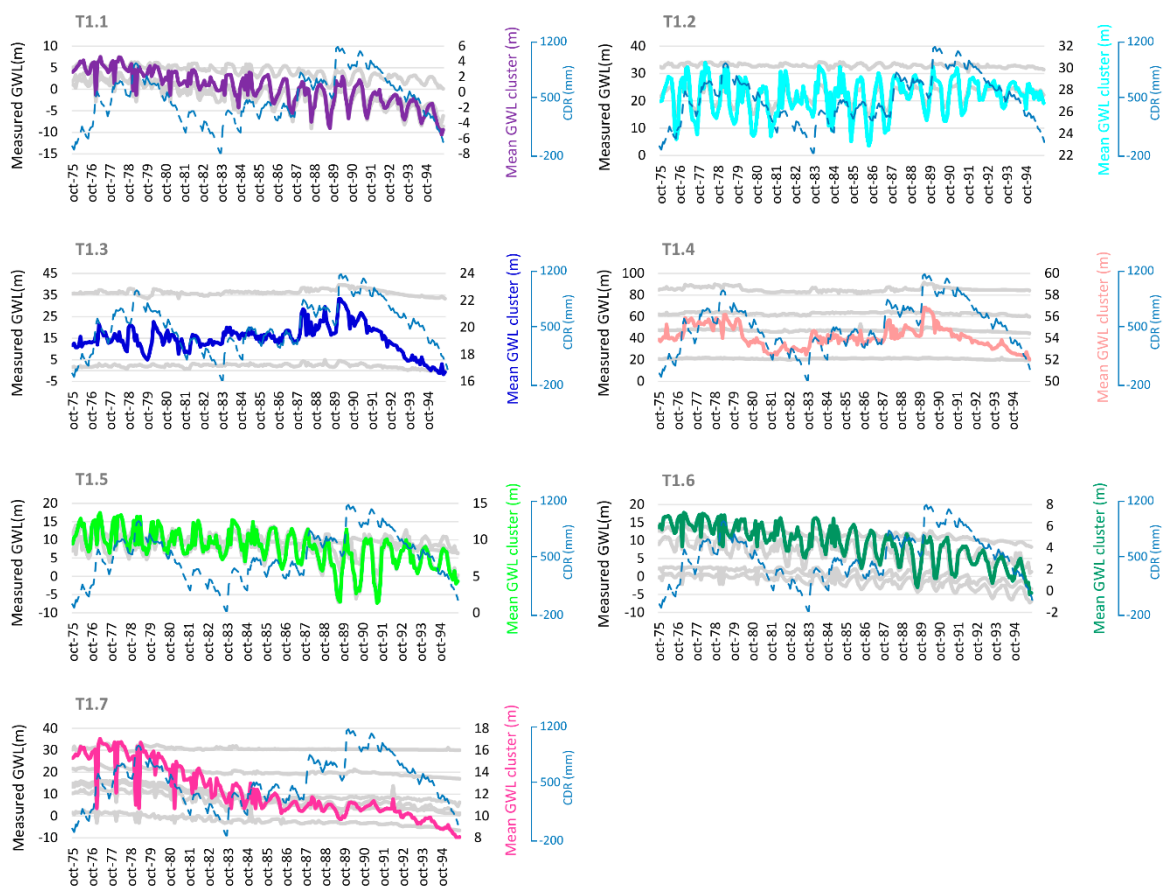


Figure 10. GWL series (grey lines), mean of the GWL cluster series (coloured line) from October 1975 to September 1995 (first sub-period), classified by TSC, and cumulative deviation from the mean monthly rainfall (CDR).

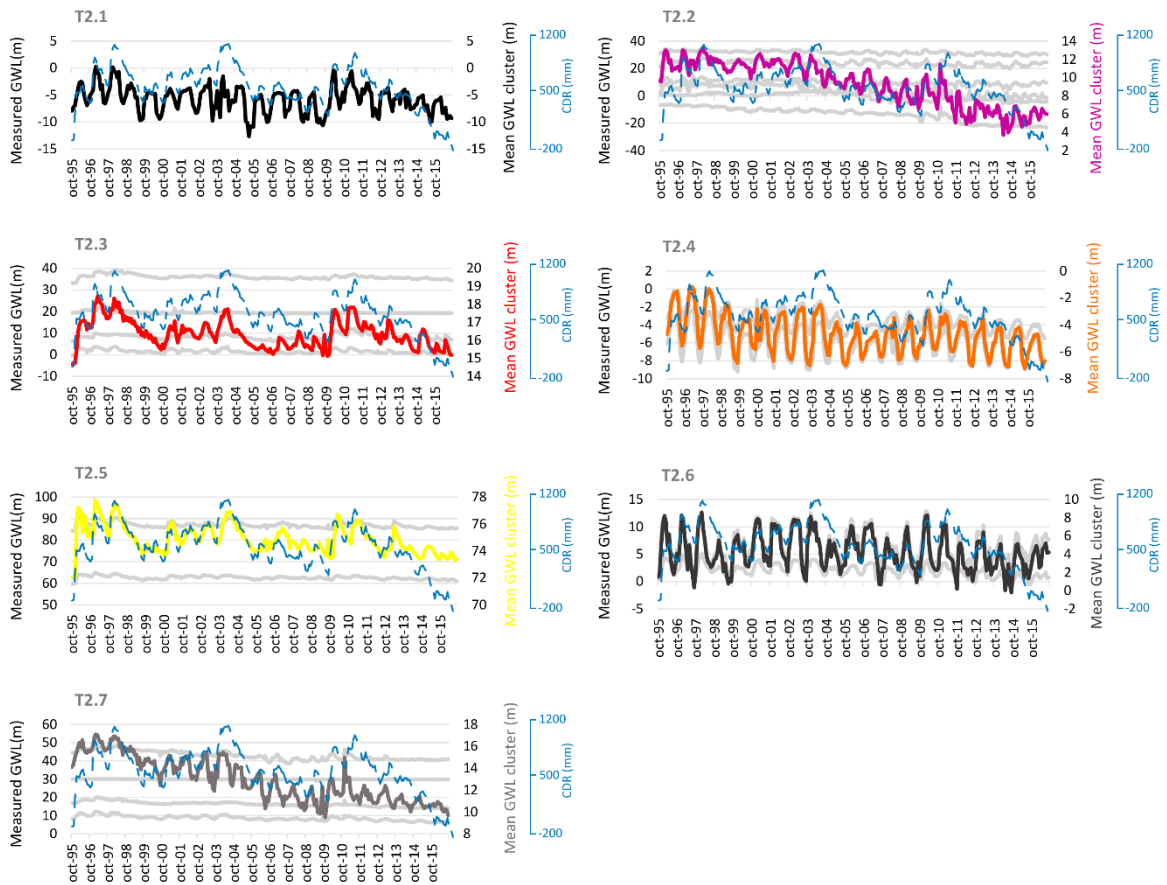


Figure 11. GWL series (grey lines), mean of the GWL cluster series (coloured line) from October 1995 to September 2016 (second sub-period), classified by TSC, and cumulative deviation from the mean monthly Rainfall (CDR).

Comparing k-means results with TSC, the clusters reveal some similarities. For example, for the first sub-period, the location and mean graphs of pairs K1.2-T1.2 and K1.5-T1.5 are equal but K1.6-T1.7 only share the mean series shape (Figures 3, 7 and 10). That is the case in the second sub-period for pairs K2.2-T2.2 and K2.3-T2.3 (Figures 8 and 11). For the second sub-period, the spatial distribution of the results is quite different when both methods are compared. K-means is based on the Euclidean distance between point coordinates in the feature space, thus distance is evaluated in terms of the difference between the GWL of the piezometers at the same time coordinate. Basically, k-means works by clustering together piezometers showing the least distant GWL in the complete time period vector but it does not take into account how the GWL change along the different time series. On the other hand, TSC uses the shape-based distance (SBD) that accounts for the cross-correlation between GWL series. This allows for the detecting of lagged responses in different piezometers which is, in turn, more closely related to the

hydrodynamic characteristics (Guijo-Rubio et al. 2018). TSC warps the series in order to fit the oscillation curves. This is confirmed by the results, for example K1.4 is composed of only one piezometer (104140047, Appendix 1-Table 2). This point is far away from the other 23 points and is located in the northern part of the aquifer with a higher topographic elevation (Figure 1c). On the other hand, the same piezometer is included with another 3 piezometers in the T1.4 group. The average mean of T1.4 GWL series is similar to K1.4 but smoothed out. This means that there are 4 points with similar patterns but k-means could not detect it (Figures 7 and 10). Furthermore, unlike most k-means clusters, that show relatively constant GWL trends (Figures 7 and 8), several TSC clusters (T1.1, T1.6, T1.7, T2.2, T2.4 and T2.7), present clearly decreasing patterns (Figures 10 and 11).

In the second sub-period, there are at least two groups with similar GWL behaviour to the CDR time series. As aquifer storage is depleted because of exploitation, the rainfall dependency of the GWL response increases (Lorenzo-Lacruz et al. 2017). For the second sub-period, the mean graphs of the cluster groups (Figure 11) are more similar to CDR than the first sub-period results (Figure 10)

TSC was also successfully applied to a 0.2 km² pre-alpine Swiss shallow aquifer to investigate subsurface hydrological connectivity (Rinderer et al. 2019). The authors found six groundwater response clusters over a 19 month study period. As is shown in the present work, they found that relative groundwater levels in the wells in each cluster were significantly different, in terms of amplitude and the rate of recession. Otherwise, GWL ranges in the Almonte-Marismas aquifer are quite dissimilar between piezometers belonging to the same cluster.

Rinderer et al. 2019 studied a small, shallow aquifer, which is difficult to compare with the high GWL ranges of the bigger and heterogeneous Almonte-Marismas aquifer. As these authors state, grouping the groundwater dynamics into a fewer number of clusters is a gross simplification as there are obviously more types of groundwater responses. However, this clustering is a necessary step in synthesizing site-specific groundwater dynamics to general groundwater behaviour. For example, TSC clusters could be used to define the over-exploitation zones in the Almonte-Marismas aquifer. Another application could be the selection of a number of key piezometers to focus the

calibration process of the existing Almonte-Marismas groundwater mathematical model (Guardiola-Albert et al. 2009). This will reduce the dimension of the optimisation problem and, consequently, the computation time.

The most important result from TSC is the six characterised types of hydrodynamic behaviour (T2.1 is a singularity) that are now present in the aquifer (Figure 11). Visually, the mean piezometric series of T2.2 is similar to that of T2.7 and the same happens with T2.4 and T2.6. The area around T2.4 is being excavated for the gravels that are under the marshlands, while gravels that are under the aeolian sands are being extracted from the area around T2.6. The different hydrogeological parameters in both cluster locations explain the differentiation in these GWL time series that cannot be seen with a qualitative comparison. In addition, the descending trends seen in T2.2 and T2.7 (from the beginning of the 2000s) are the result of uncontrolled increases in the berry fields in the northwestern part of the aquifer. This development has been denounced by WWF on several occasions (WWF 2016). On the other hand, clusters T2.3 and T2.5 are less anthropically influenced but are more affected by recharge fluctuations.

On a global scale, GWL time series clustering can contribute to the identification and separation of spatiotemporal processes associated with both climate and human disturbances. This separation is often tackled through complex transfer function-noise models (Shapoori et al. 2015; van Dick et al 2019). Similar aquifer exploitation regimes in different countries or regions could then be clustered together and provide the basis to build time series classification models based on labeled time series resulting from the clustering analysis. In turn, this would support the development of global actions for sustainable groundwater management.

4. Conclusions

The main goals of this work were to study the historical piezometry of the Almonte-Marismas Aquifer, to improve the knowledge about hydrodynamics and to prove that anthropic activities have changed GWL trends in the system. It was possible to reach these aims by analyzing the historic piezometric database (for a period of over 41 hydrological years), applying two robust clustering methods, comparing the results and highlighting the advantages and disadvantages of both. Temporal changes in the

variables that affect GWL (recharge, water exploitation, climate change, surface/groundwater interactions) could help to explain the visual, k-means and TSC results. The study period was divided in two in order to search for important piezometric changes due to the growing demand of irrigation fields. The differences observed between the GWL clusters for the first and the second sub-periods reflect the influence of the socioeconomic and anthropic development on the natural regime. The relationship between CDR and GWL decreases with anthropic exploitation, making aquifer storage more vulnerable to the effects of climate change but the response of GWL to rainfall events becomes quicker when the aquifer is over-exploited. The natural system exhibits complex hydrological behaviour, highlighting the need to adapt water policy and management to every single case.

Visual clustering offers a first approach to characterising the behaviour of the hydrographs, considering expert judgement and providing a basis for the validation of mathematical methods. However, in the second study sub-period, it was observed that the combination of extractions (in terms of quantity and time of year) produces a merging of responses in which there is no clear spatial correspondence. More specifically, k-means methodology allows the distinguishing of singular GWL oscillations, assigning piezometers to classes with only one GWL point. This permits the identification of errors in data measurements or temporal changes in the external factors that does not occur in other areas of the aquifer. In areas not affected by groundwater extractions, k-means classification gives similar results to visual classification.

Our findings confirm that TSC offers better results for studying time series distribution because the members of the same group reproduce the same pattern. K-means considers time (average monthly GWL as attributes for each piezometer) but not the correlation structure of the time series. This is an important limitation to detecting changes in temporal patterns in different groups. TSC considers the time series distribution and the results show consistent distribution of GWL responses. Further studies will facilitate the introduction of cluster results, in order to optimise and reduce the amount of input data for mathematical groundwater flow models, contributing to speeding up and simplifying the modelling process. Also, the obtained results could offer a way to zone the effects of uncontrolled exploitation wells by estimating the extent of

the effects of water extractions. The new characterisation of the Doñana piezometry, based on a long GWL series and TSC objective criteria, proves the important impact that uncontrolled extraction has had over a period of decades.

Funding: This research has been funded by the CLIGRO project [MINECO, CGL2016-77473-C3-1-R] of the Spanish National Plan for Scientific and Technical Research and Innovation, and the grant to the execution of Industrial PhDs within the Autonomous Region of Madrid (Spain) IND2018/AMB-9553.

Acknowledgments: We are grateful for the support given by Fernando Ruíz Bermudo and Antonio Sánchez de la Nieta from the Spanish Geological Survey (IGME) for the acquisition and maintenance of the piezometer database; the Guadalquivir River Basin Authority (CHG) for the piezometric data supplied; and the Spanish State Meteorological Agency, the Biological Station of Doñana and Junta de Andalucía for the meteorological data provided.

Appendix A

Table 2. Statistical parameters of data points.

Code	N	Mean	Med.	Max.	Min.	St. Dev	% MV
104140047	377	86.34	86.27	91.64	81.33	2.02	23.37
104180012	234	20.57	20.81	22.11	18.62	0.90	52.44
104180021	422	18.07	17.88	23.63	13.88	2.36	14.23
104210004	252	46.41	46.78	53.04	38.56	2.08	48.78
104220006	253	36.30	36.13	39.74	33.19	1.42	48.58
104230003	261	62.94	62.97	65.05	59.80	1.09	46.95
104240033	321	10.28	10.39	15.47	0.39	2.64	34.76
104240058	431	23.10	23.71	27.92	14.46	2.78	12.40
104240066	335	32.08	32.26	34.21	28.88	1.18	31.91
104240082	431	5.01	5.41	12.09	-5.48	3.81	12.40
114150046	343	8.75	8.97	11.98	6.03	1.30	30.28
114150065	404	2.37	1.94	15.50	-5.36	5.10	17.89
114160012	392	30.27	30.02	33.10	28.83	0.68	20.33
114170034	350	-7.95	-5.54	4.73	-23.19	7.59	28.86
114170040	335	-6.51	-5.36	3.89	-21.29	5.75	31.91
114180059	327	-4.35	-4.56	3.55	-12.66	3.31	33.54

Code	N	Mean	Med.	Max.	Min.	St. Dev	% MV
114210031	235	5.45	7.87	16.44	-4.15	6.86	52.24
114210051	379	6.60	6.64	14.23	-7.06	3.87	22.97
114210076	356	2.67	2.47	6.60	-0.54	1.71	27.64
114210094	243	5.06	4.87	12.64	-2.82	3.54	50.61
114210114	338	9.65	9.50	14.20	5.98	1.98	31.30
114220007	390	1.66	1.65	5.00	-0.78	1.03	20.73
114220013	428	-1.49	-1.83	2.62	-5.52	2.50	13.01
114230024	359	-3.87	-3.88	0.86	-9.21	2.96	27.03

Capítulo 5: Discusión general

5.1. Nuevas metodologías de trabajo aplicadas a los datos de entrada para la mejora del modelo matemático del acuífero Almonte-Marismas

Para la consecución de los objetivos de esta tesis doctoral se aplicaron técnicas geoestadísticas y estadísticas en los datos de entrada al modelo de flujo del acuífero Almonte-Marismas.

5.1.1. Simulación estocástica condicionada para la realización de un modelo de hidrofacies en 3D

La primera de las aportaciones de este trabajo ha sido la mejora del modelo geológico conceptual. Especialmente, la redefinición del muro del acuífero obtenida mediante interpolación estadística previa, ha facilitado partir de una geometría del acuífero mejor definida que la utilizada en los modelos hidrogeológicos anteriores ya que está basada en los datos de los sondeos geológicos realizados. Se observó un aumento en el espesor del acuífero con respecto a la caracterización geológica de referencia (Salvany et al., 2011).

La proporción de presencia de las hidrofacies en los datos geológicos, estimada en base a los datos geológicos y geofísicos, fue: 25% de arenas eólicas (Hidrofacies 1), 2% de limos basales (Hidrofacies 2), 12% de arenas basales (Hidrofacies 3), 15% de arcillas de marisma (Hidrofacies 4), 24% de arcilla con arena y gravas (Hidrofacies 5) y 22% de arenas con gravas (Hidrofacies 6). Estos porcentajes fueron conservados en la simulación estocástica.

En la modelación de los variogramas, se obtuvieron rangos verticales menores que los horizontales, debido a que, en general, las hidrofacies definidas poseen mayor extensión horizontal que espesor en la vertical. Todos los modelos de variograma presentaron anisotropía excepto los limos basales, debido a la distribución homogénea que presentan este tipo de depósitos.

En cuanto a la simulación obtenida, al estar condicionada se ha logrado que los resultados reproduzcan la geología de una forma más cercana a la realidad. Las hidrofacies de arenas eólicas y arenas basales alcanzaron profundidades más elevadas

que en el modelo conceptual original. Sin embargo, las hidrofacies acuíferas (5 y 6) simuladas presentaron espesores que reducen a la mitad los espesores del modelo conceptual original, pasando de alcanzar 300 m de profundidad a 150 m.

Se compararon los dos modelos estacionarios (el modelo anterior disponible con las permeabilidades por zonas, y el modelo con hidrofacies), calibrados de manera independiente con la finalidad de discutir las implicaciones que conlleva esta mejora sobre los balances hídricos resultantes. El modelo estacionario con la implementación de la simulación estocástica desveló un comportamiento hídrico más coherente con la complejidad geológica del acuífero. Por un lado, permitió eliminar la impermeabilidad total en las celdas que representan la marisma y añadir celdas de baja permeabilidad para simular las arcillas plásticas intercaladas con las unidades de arenas y gravas. Por otro, el modelo estacionario pasó de ser un modelo bicapa a poder ejecutarse con siete capas diferentes. Además, se apreciaron grandes diferencias en la distribución de las isopiezas entre ambos modelos, sobre todo en las zonas norte y oeste del acuífero. En la zona norte, esta variación se atribuyó a la presencia de los limos basales que originaron la variación de las propiedades hidrogeológicas con respecto al modelo original. Se obtuvo una mayor altura piezométrica general en el modelo con hidrofacies. Se observaron diferencias significativas en las salidas del balance hídrico hacia el mar y los arroyos, infraestimadas un 18% y un 23% respectivamente en el modelo original.

En términos de gestión, se detectó un incremento de hasta el 30% en el intercambio de flujo entre el arroyo de la Rocina y el acuífero en el modelo con hidrofacies. Estas elevadas diferencias porcentuales entre ambos balances ponen de manifiesto la importancia de realizar un estudio en profundidad de los parámetros hidrogeológicos que pueden afectar a los resultados del modelo matemático de flujo.

5.1.2. Interpolación espacio-temporal de la precipitación y su influencia sobre la recarga

A través de la aplicación del krigeado espacio-temporal, se interpolaron los datos de precipitación, obteniendo una mejor resolución espacial y temporal para los años hidrológicos comprendidos entre 1975 y 2016. Se obtuvieron valores de precipitación a distintas escalas temporales (mensual y diaria) y a distintas escalas espaciales (píxeles

de 500x500 m, zonas de recarga y masas de agua subterránea). Los resultados mostraron una elevada variabilidad espacio-temporal tanto en años secos como en años húmedos. Se pone así de manifiesto que el uso de un valor de lluvia media o de una única estación climática para el cálculo de los recursos hídricos disponibles puede generar un error muy elevado, sobre todo en las épocas más húmedas. Las principales desventajas del método fueron la detección de datos completados con 1 mm aproximadamente, que a pesar de que no contabilizan en el volumen de recarga, pueden generar una mala interpretación de la cantidad de días con precipitación registrados. También, cabe destacar el elevado tiempo de cálculo que implica interpolar la precipitación a escala diaria con el krigado espacio-temporal. La principal ventaja fue la capacidad de calcular datos de precipitación de manera consistente, incluso con un gran número de datos faltantes.

Conseguir un cálculo preciso de la precipitación y aplicarlo al cálculo de la recarga, puede tener un gran impacto en las decisiones de los futuros Planes de Gestión. Frente a la metodología habitual del cálculo de la recarga llevada a cabo por la Confederación Hidrográfica del Guadalquivir (CHG, 2015) que se realiza con paso de tiempo mensual, el uso de los datos a escala diaria que se obtuvo mediante el krigado espacio-temporal puso de manifiesto la importancia de interpretar con sumo cuidado los valores de recarga que se obtienen cuando los cálculos se hacen de forma mensual, pudiendo infraestimar o sobreestimar los recursos. Esto podría dar lugar a decisiones de gestión erróneas que incluso puedan acrecentar la sobreexplotación del acuífero en el futuro. Se determinó que los valores de recarga mensuales, calculados a partir de los valores diarios, oscilan entre el 21 y el 91 % de la precipitación máxima. Como consecuencia de la variabilidad espacial detectada, se destacó la importancia de trabajar con subdivisiones o zonas de recarga definidas con criterios hidrogeológicos y no administrativos, como es el caso de las masas de agua subterránea (Guardiola-Albert et al., 2016), con las que se trabaja actualmente.

En un contexto de cambio climático, se observó un descenso tanto en la media anual como en el número de días al año que registran recarga, presentando una tendencia descendente desde 1995. Este fenómeno se traduce en un aumento en la intensidad de las precipitaciones, un descenso de la recarga efectiva, un detrimento del

almacenamiento neto en el acuífero y, por último, en una tendencia cada vez más elevada a episodios de sequía que será cada vez más complicado suplir mediante la extracción de aguas subterráneas debido a su agotamiento.

5.1.3. Agrupamiento estadístico de series temporales de piezometría

Con este último trabajo, se logró mejorar la interpretación global de los datos de piezometría mediante la realización de un agrupamiento de series temporales. Se aplicó para el mismo periodo de tiempo que el estudio de la precipitación, de octubre de 1975 a septiembre de 2016. En este caso, y con el fin de analizar la variación dentro de este largo periodo, el agrupamiento se realizó en dos subperiodos de 20 y 21 años hidrológicos. Se compararon los resultados obtenidos mediante k-medias y agrupamiento de series temporales. El agrupamiento de series temporales resultó ser la técnica más adecuada ya que tiene en cuenta la correlación temporal de los datos.

El método de k-medias permitió distinguir las oscilaciones de los niveles de agua subterránea y separar incluso grupos singulares con un solo piezómetro, lo cual permite encontrar errores en las medidas de los datos o cambios temporales en factores externos que no ocurren en otras áreas del acuífero. En zonas donde no se producen elevadas extracciones de agua subterránea, la clasificación obtenida por k-medias fue muy similar a la realizada visualmente.

Con la aplicación del agrupamiento de series temporales se obtuvieron siete grupos de respuestas hidrológicas de los piezómetros en los que se detectó cuáles corresponden a zonas de aprovechamiento de las aguas subterráneas para uso agrícola, y cuáles presentan elevada correspondencia con la curva de desviaciones acumuladas de la precipitación, es decir, no se encuentran tan afectados por las extracciones de agua subterránea (zona norte). Además, se detectaron tendencias descendentes claras en los piezómetros de la zona central y centro-oeste del acuífero desde el principio de los 2000, coincidiendo con el incremento de explotaciones de frutos rojos en el noroeste del acuífero.

Comparando la piezometría media de cada grupo obtenido con la desviación de la lluvia acumulada, se detectó que en el segundo subperiodo (1995-2016) había mayor similitud entre ambos parámetros. Debido a esto, se pudo determinar que la respuesta de los

niveles piezométricos es más dependiente de la precipitación cuanto más disminuyen las reservas del acuífero por la sobreexplotación.

La aplicación de las técnicas de agrupamiento estadístico ha contribuido a la detección y separación de procesos asociados tanto a cambios climáticos como a cambios en la intervención humana. Aplicarlo a acuíferos con regímenes similares de explotación, de manera global, podría ayudar en el desarrollo de acciones resilientes de gestión sostenible de las aguas subterráneas.

5.2. Aportaciones a la investigación sobre modelos matemáticos y sus consecuencias para la mejora de la gestión de los recursos hídricos que se apoyan en ellos

En el caso concreto del acuífero Almonte-Marismas, la aplicación del modelo matemático es una herramienta irremplazable mediante la cual se obtienen los datos de balance hidrológico para gestionar las dotaciones futuras (Figura 11). Además, se debe priorizar el buen estado y la sostenibilidad a largo plazo (Carmona et al., 2013), ya que el agua subterránea es la garantía de suministro en esta región. La dificultad de gestión de este acuífero reside en llegar a un equilibrio entre el buen estado ecológico y la actividad socioeconómica, sin generar un impacto ambiental que pueda ser irreversible en el futuro. Además, el hecho de que el acuífero esté subdividido en masas de agua subterránea que no se corresponden con límites hidrogeológicos (Guardiola-Albert et al., 2016) sino administrativos, pone aún más en valor la necesidad de gestionar los recursos hídricos mediante el modelo del acuífero completo.

El conocimiento generado mediante el desarrollo de esta tesis doctoral tiene una repercusión directa en la posibilidad de mejora de la gestión que realiza la Confederación Hidrográfica del Guadalquivir, lo que afectará en el futuro a las decisiones que se tomen en la redacción de los Planes de Gestión. Adicionalmente, ha aportado la generación de nuevos conocimientos sobre sus variables, permitiendo una mejor comprensión de funcionamiento hidrogeológico y pudiendo paliar muchos de los déficits que en los modelos ya realizados se detectaban.

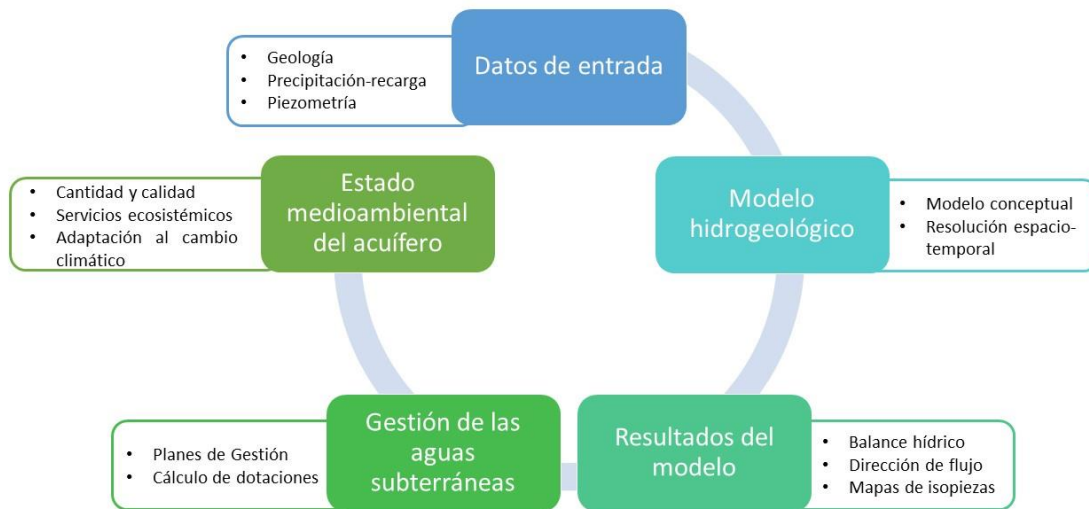


Figura 11. Esquema de flujo de las ventajas que proporciona la utilización de datos de entrada mejorados en el modelo matemático de flujo.

La consecución de la presente tesis ha aportado:

- Un modelo geológico digital de hidrofacies que proporciona información sobre la distribución espacial de las diferentes unidades, acercando el modelo a un comportamiento que reproduce mejor la dinámica del flujo real.
- Un mejor conocimiento de la precipitación y su variabilidad, lo cual incide en la implementación de datos más precisos de la recarga, fuente principal de entradas en el balance hídrico, por lo que, afectará a los nuevos cálculos del balance hídrico obtenido.
- Un agrupamiento de piezómetros que mejora el entendimiento sobre la respuesta de los niveles del agua subterránea ante la explotación antrópica. Aplicado al modelo hidrológico, podrá optimizarse el número de piezómetros representativos, lo que reducirá el tiempo de calibración automática, ya que podría calibrarse únicamente un piezómetro característico por zona agrupada.

El desarrollo de modelos matemáticos en Hidrogeología es un trabajo en constante evolución (Figura 12). Cualquier mejora en uno de sus componentes conlleva necesariamente una modificación y un avance en los restantes. En el caso de esta tesis doctoral, se ha aplicado una optimización del análisis de los datos monitoreados, lo cual

mejora la calidad del modelo conceptual y de la modelización matemática del modelo estacionario, aportando nuevos conocimientos a la investigación hidrogeológica.

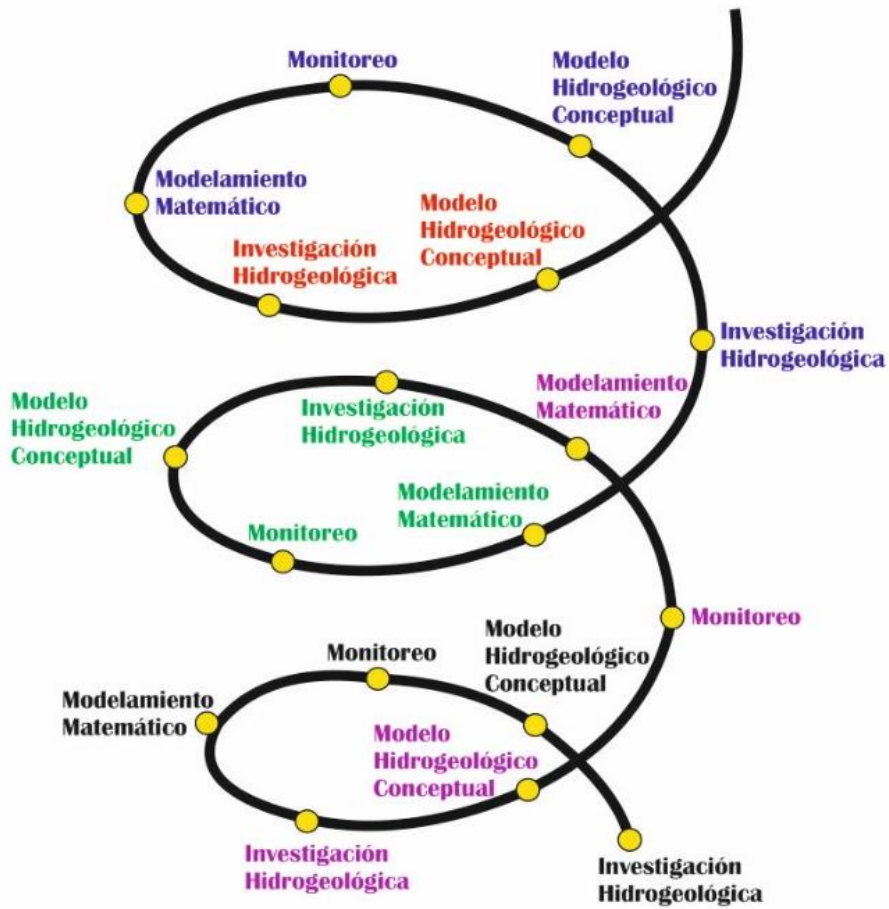


Figura 12. Esquema de trabajo en modelización hidrogeológica.

Capítulo 6: Conclusiones y líneas de trabajo futuras

El objetivo principal de este trabajo ha sido profundizar en el estudio estadístico de tres de las propiedades del acuífero Almonte-Marismas con el fin de mejorar su caracterización. En este sentido, se ha aportado una mejora en el conocimiento sobre la geología, el binomio precipitación-recarga y la piezometría a partir de este estudio. Los tres artículos que componen esta tesis doctoral se basan en una serie de 41 años, lo que posibilita dar una visión global sobre el estado hidrogeológico del acuífero desde los primeros años de su explotación.

Los objetivos parciales descritos en el Capítulo 2 de esta tesis, han sido satisfactoriamente alcanzados:

- Se ha utilizado por primera vez una simulación estocástica para definir la variabilidad espacial de la conductividad hidráulica y se ha logrado incorporar el resultado al modelo numérico de flujo regional del acuífero Almonte-Marismas. El modelo se construyó mediante la aplicación de simulación geoestadística con la modelización previa de los variogramas obtenidos y la caracterización de las heterogeneidades geológicas. Los resultados han confirmado que el balance de agua puede variar hasta un 23% al introducir la información espacial que se obtiene de la simulación. Por ejemplo, las variaciones en el balance hidrológico permiten estimar con mayor precisión los caudales en las zonas de descarga hacia los cursos superficiales, ya que son de vital importancia para mantener los ecosistemas ribereños y el aporte superficial que alimenta a las marismas de Doñana.
- Se ha aplicado por primera vez en un acuífero español el krigeado espacio-temporal para la interpolación de la precipitación a nivel diario para conseguir mejorar la estimación de la recarga e introducirla en un modelo numérico de flujo. Se ha confirmado que, a pesar de ser un área de estudio sin muchas variaciones de elevación topográfica, las estimaciones de la recarga tienen una elevada variabilidad espacial, relacionada con la heterogeneidad espacial tanto de la precipitación como de las propiedades y usos del suelo. La elección de la escala temporal en el cálculo de la recarga es un factor clave en el contexto del cambio

climático en regiones áridas, en las cuales se está produciendo un incremento de las temperaturas diarias y de las precipitaciones extremas. Comparando los valores medios de dos periodos diferentes, 1975-1995 y 1996-2016, que cubren 41 años hidrológicos, se observó un decrecimiento anual en la recarga y en el número de días con recarga efectiva. Se demostró, además, que las pequeñas cantidades de lluvia diaria no contribuyen a la recarga cuando ésta se calcula con paso de tiempo mensual. Como consecuencia de este fenómeno, puede existir una sobreestimación de los recursos hídricos disponibles. Teniendo en cuenta las conclusiones obtenidas, en los nuevos planes de gestión de las aguas de Doñana se deberían tomar escalas espacio-temporales más detalladas, para obtener una estimación más exacta del balance de agua real del acuífero Almonte-Marismas.

- Se ha concluido con el estudio detallado de la piezometría histórica del acuífero Almonte-Marismas. Se ha mejorado el conocimiento de la hidrodinámica y de cómo las actividades antrópicas la han transformado la piezometría durante los 41 años de estudio. Las variaciones en los niveles piezométricos obtenidos están relacionadas con variables como la recarga, la explotación de las aguas subterráneas, el cambio climático y la interacción aguas superficiales-aguas subterráneas. Los resultados del agrupamiento de series temporales han desvelado que esta metodología es más apropiada para el estudio de la distribución temporal de la piezometría, ya que los miembros del mismo grupo reproducen el mismo patrón de correlación temporal. TSC considera la distribución temporal de la piezometría y los resultados muestran una distribución consistente de las respuestas de los niveles piezométricos. Estudios como éste pueden optimizar y reducir la cantidad de datos de entrada en los modelos matemáticos de flujo, contribuyendo a acelerar y simplificar los procesos de modelación. Además, los resultados obtenidos podrían ofrecer la posibilidad de conocer la extensión de los efectos de las explotaciones incontroladas en los niveles de agua subterránea. Por otra parte, cabe destacar que es el primer ejemplo de aplicación del agrupamiento de series temporales a datos piezométricos en el ámbito nacional y el segundo a nivel internacional.

Las líneas de trabajo futuro que se relacionan con esta tesis doctoral son:

- Monitoreo de caudales de los cursos de agua superficial que vierten al Espacio Natural de Doñana (Arroyo Madre de las Marismas, Arroyo de la Rocina, Caño del Guadamar) para establecer las relaciones aguas superficiales-aguas subterráneas.
- Aplicación de técnicas geoestadísticas y de *machine learning* a una combinación de datos de pluviometría recogidos en campo y por radar meteorológico para la interpolación de la precipitación. Estas técnicas que han demostrado ser muy útiles para esta aplicación, pero nunca han sido utilizadas en Doñana.
- La detección de zonas con concentración de extracciones ilegales, a través de la interpolación de alta resolución de la piezometría, apoyado por técnicas de teledetección InSar y GRACE.
- La simulación de la recuperación de las reservas hídricas de las masas de agua subterránea declaradas en riesgo de no alcanzar el buen estado cuantitativo mediante la aplicación de distintos escenarios de gestión: sustitución de aguas subterráneas por superficiales en las zonas de El Fresno y El Rocío y estudio de reubicación de sondeos o instalación de desaladoras en Matalascañas.
- Mejora del estudio estadístico de la sequía a través de la caracterización de la incertidumbre climática para la definición de escenarios futuros de gestión.
- La realización de un modelo integrado SWAT-MODFLOW que permita conocer mejor la relación río-acuífero y la contaminación por nitratos en la zona, apoyada en relaciones físicas robustas que representan la infiltración a través de la zona no saturada.

Referencias

- Aguilera, H., Guardiola-Albert, C., Naranjo-Fernández, N., Kohfahl, C. (2019) Towards flexible groundwater level prediction for adaptive water management: using Facebook's Prophet forecasting approach. *Hydrological Sciences Journal*, 64, 1504-1518, DOI: 10.1080/02626667.2019.1651933.
- Aguilera, H., Guardiola-Albert, C., Serrano-Hidalgo, C. (2020) Estimating extremely large amounts of missing precipitation data. *Journal of Hydroinformatics*, DOI: 10.21667/hydro.2020.127.
- Ahmadaali, K., Damaneh, H.E., Jalbarezi, B. (2018) Application of Geostatistical methods to estimate groundwater level fluctuations. *International Journal of Advanced Biological and Biomedical Research*, 6(1), 11-22.
- Alabert, F. (1987) Stochastic imaging of spatial distributions using hard and soft information. Master's thesis, Department of Applied Earth Sciences, Stanford University, Stanford, 1987.
- Almeida, J.A. (2010) Stochastic simulation methods for characterization of lithoclasses in carbonate reservoirs. *Earth-Science Reviews*, 101, 250-270, DOI: 10.1016/j.earscirev.2010.05.002.
- Álvarez-Cobelas, M., Cirujano, S., Sanchez-Carrillo, S. (2001) Hydrological and botanical man-made changes in the Spanish wetland of Las Tablas de Daimiel. *Biological Conservation*, 78(1), 89-98, DOI: 10.1016/S0006-3207(00)00102-6.
- Álvarez-Rodríguez, J., Sanchez, A., Quintas, L. (2005) SIMPA, a GRASS based Tool for Hydrological Studies. *International Journal of Geoinformatics*, 1, 1-14.
- Amini, A., Homayounfar, V. (2017) The groundwater balance in alluvial plain aquifer at Dehgolan, Kurdistan, Iran. *Applied Water Science*, 7, 3113-3123.
- Anderson, M.P. (1989) Hydrogeological facies models to delineate large scale spatial trends in glacial and glacio fluvial sediments. *Geological Society of America Bulletin*, 101, 501-511.
- Ardabali, S.F., Mosavi, A., Dehghani, M. (2019) Deep learning and machine learning in hydrological processes, climate change and earth systems: a systematic review. *Preprints*, 2019080166, DOI: 10.209447preprints201908.0166.v1.
- Armstrong, M. (1998) Basic linear geostatistics. Springer-Verlag Berlin Heidelberg, 155p. DOI: 10.1007/978-3-642-58727-6.
- Arun, M., Sananda, K., Anirban, M. (2012) Rainfall Trend Analysis by Mann-Kendall Test: A Case Study of North-Eastern Part of Cuttack District, Orissa. *International Journal of Geology, Earth and Environmental Sciences*, 2, 70-78.
- Ashgarinia, S.; Petroselli, A. (2020) A comparison of statistical methods for evaluating missing data of monitoring wells in the Kazeroun Plain, Fars Province, Iran. *Groundwater for Sustainable Development*, 10, 100294, DOI: 10.1016/j.gsd.2019.100294.

- Baalousha, H. (2009) Fundamentals of groundwater modelling. In *Groundwater: Modelling, Management and Contamination*. Konig, L.F., Weiss, J.L., Eds.; Nova Science Publishers, Inc.: New York, NY, USA, 2009, pp. 149–166. ISBN 9781604568325.
- Babovic, V. (2005) Data mining in hydrology. *Hydrological Processes*, 19, 1511-1515, DOI: 10.1002/hyp.5862.
- Balek, J. (1988) Groundwater recharge concepts. Estimation of Natural Groundwater Recharge. Boston. Ed. Reidel, NATO ASI Series, 1988. p. 3-9.
- Bennett, B., Thyer, M., Leonard, M., Lambert, M., Bates, B. (2018) A comprehensive and systematic evaluation framework for a parsimonious daily rainfall field model. *Journal of Hydrology*, 556, 1123-1138.
- Beucher, H., Renard, D. (2016) Truncated Gaussian and derived methods. *C R Geosciences*, 348, 510–519.
- Bloomfield, J.P., Marchant, B.P., Bricker, S.H. (2015) Morgan, R.B. Regional analysis of groundwater droughts using hydrograph classification. *Hydrological Earth Systems Science*, 19, 4327–4344, DOI: 10.5194/hess-19-4327-2015.
- BOE- Boletín Oficial del Estado – nº141, Real Decreto 712/2006, de 9 de junio de 2006.
- BOE- Boletín Oficial del Estado - nº 16, Real Decreto 1/2016, de 19 de enero de 2016.
- BOE- Boletín Oficial del Estado -nº 225, Anuncio 39064, de 19 de septiembre de 2019.
- Bottou, L., Bengio, Y. (1994) Convergence properties of the k-mean algorithms. *Neural Information Processing Systems 7 (NIPS 1994)*, 585-592, MIT Press, Denver.
- Breiman, L. (2001) Random Forests. *Machine Learning*, 45(1), 5-32, DOI: 10.1023/A:1010933404324.
- Burgueño, A., Martínez, M.D., Lana, X., Serra de Larrocha, C. (2005) Statistical distributions of the daily rainfall regime in Catalonia (Northeastern Spain) for the years 1950-2000. *International Journal of Climatology*, 25(10), 1381-1403, DOI: 10.1002/joc.1197.
- Buytaert, W., Celleri, R., Willems, P., De Bievre, B., Wyseure, G. (2006) Spatial and temporal rainfall variability in mountainous areas: a case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329, 413-421, DOI:10.1016/j.jhydrol.2006.02.031.
- Cáceres Puro, L.M., Rodríguez Vidal, J. Alteraciones por clima de tipo tropical al final de Plioceno en Bonares. En: Olías Álvarez, M., et al.: "Geomorfología del Parque Nacional de Doñana". 2ª ed. Huelva: Universidad de Huelva, 2008. págs. 150-155
- Caers, J. (2000) Adding Local Accuracy to Direct Sequential Simulation. *Mathematical Geosciences*, 32, 815-50, DOI: 10.1023/A:1007596423578.
- Calinski, T., Harabasz, J. (1974) A dendrite method for cluster analysis. *Communications in Statistics 3*, 1-27.
- Carmona, G., Varela-Ortega, C., Bromley, J. (2013) Supporting decision making under uncertainty: Development of a participatory integrated model for water management in

the middle Guadiana river basin. *Environmental Modelling and Software*, 50, 144–157, DOI:10.1016/j.envsoft.2013.09.007.

Carnicer i Cols, J. (2007) Species richness, interaction networks, and diversification in bird communities: a synthetic ecological and evolutionary perspective. Bellaterra: Universitat Autònoma de Barcelona, 2008. ISBN 9788469139042. Tesis doctoral - Universitat Autònoma de Barcelona, Facultat de Ciències, Departament de Biologia Animal, Biologia Vegetal i Ecologia, 2007 <<https://ddd.uab.cat/record/36713>> [Visited: 10th December 2019].

Castellazzi, P., Martel, R., Rivera, A., Huang, J., Pavlic, G., Calderhead, A. I., Chaussard, E., Garfias, J., Salas, J. (2016) Groundwater depletion in Central Mexico: Use of GRACE and InSAR to support water resources management. *Water Resources Research*, 52, 5985–6003, DOI:10.1002/2015WR018211.

Castro, A., Vazquez-Sune, E., Carrera, J., Jaen, M., Salvany, J.M. (1999) Calibración del modelo regional de flujo subterráneo en la zona de Aznalcollar, España: ajuste de las extracciones. *Hidrología Subterránea* II, 13. Ed A. Tineo. Congreso Argentino de Hidrogeología y IV Seminario Hispano Argentino sobre temas actuales de la Hidrogeología.

Cazenave, G., Vives, L. (2004) Modelo de transporte de solutos en aguas subterráneas de la ciudad de azul, provincia de Buenos Aires, Argentina. Cuadernos del CURIHAM, Volumen 10, pp 33-43. ISSN 1514-2906.

Chappell, A., Renzullo, L.J., Raupach, T.H., Haylock, M. (2013) Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. *Journal of Hydrology*, 493, 105-114.

Chen, Y., Han, D. (2014) Big data and hydroinformatics. *Journal of Hydroinformatics*, 18(4), 599-614, DOI: 10.2166/hydro.2016.180.

CHG. Confederación Hidrográfica del Guadalquivir (2015) Plan Hidrológico de la Demarcación Hidrográfica del Guadalquivir. Segundo ciclo de planificación: 2016-2021.

CHG. (2017) Plan especial de sequía. Demarcación hidrográfica del Guadalquivir. Borrador para consulta pública. Diciembre de 2017. Memoria, 441 pp.

CHG. (2019) Informe de seguimiento del plan hidrológico de la Demarcación Hidrográfica del Guadalquivir. Ciclo de Planificación 2016-2021. Ministerio de Transición Ecológica. 2019.

CHG. (2019b) Propuesta para la declaración de la masa de agua subterránea de La Rocina en riesgo de no alcanzar un buen estado cuantitativo y químico. Ministerio de Transición Ecológica. 2019.

CHG. (2020) Esquema provisional de Temas Importantes de la Demarcación Hidrográfica del Guadalquivir. Tercer ciclo de planificación hidrológica (2021-2027). Ministerio de Transición Ecológica. 2020.

Chiles, J., Delfiner, P. (1999) Geostatistics: modeling spatial uncertainty, 2nd Ed. John Wiley & Sons, New York, 1999; 734 p, ISBN: 978-0-470-18315-1.

- Christakos G. (2000) Modern Spatiotemporal Geostatistics. Oxford University Press, IAMG #6. ISBN 0-19-513895-3.
- Comunian, A., De Micheli, L., Lazzati, C., Felletti, F., Giacobbo, F., Giudici, M., Bersezio, R. (2016) Hierarchical simulation of aquifer heterogeneity: implications of different simulation settings on solute-transport modeling. *Hydrogeology Journal*, 24, 319-334.
- Custodio, E. (1997) Evaluación de la recarga por lluvia mediante métodos ambientales químicos, isotópicos y térmicos. Le evaluación de la recarga a los acuíferos en la planificación hidrológica. TEXTOS DEL SEMINARIO CELEBRADO EN LAS PALMAS DE GRAN CANARIA. Instituto Tecnológico Geominero de España. Madrid. 1997. p. 83-198.
- Custodio, E. (2001) Aguas subterráneas y humedales. Papeles del Proyecto Aguas Subterráneas, Serie C, Aguas Subterráneas y Medio Ambiente. Fundación Marcelino Botín. Madrid. p 33–72.
- Custodio, E. (2002) Aquifer Overexploitation: What does it mean?. *Hydrogeology Journal*, 10, 254-277, DOI: 10.1007/s10040-002-0188-6.
- Custodio, E. (2020) Recarga natural a los acuíferos, metodología y soporte de la isotopía del agua. Aplicación a la planificación hidrológica y conocimiento de las aguas subterráneas en España. RAEMIA 2019: 1-1206. <http://hdl.handle.net/2117/182282>.
- Custodio, E., Manzano, M., Montes, C. (2008) Perspectiva general del papel y gestión de las aguas subterráneas en el área de Doñana, Sudoeste de España. *Boletín Geológico y Minero*, 119(1):81-92.
- Custodio, E., Manzano, M., Montes, C. (2009) Las aguas subterráneas en Doñana. Aspectos ecológicos y sociales. Junta de Andalucía. Sevilla, España, 2009, 244 pp, ISBN: 978-84-92807-19-2.
- Davies, D.L., Bouldin, D.W. (1979) A clustering separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1, 224-227.
- De Iaco, S., Posa, D. (2012) Predicting spatio-temporal random fields: Some computational aspects. *Computers & Geosciences*, 41,12–24, DOI: 10.1016/j.cageo.2011.11.014.
- De Iaco, S., Maggio, M., Palma, M., Posa, D. (2012) Towards an Automatic Procedure for Modeling Multivariate Space-Time Data. *Computers & Geosciences*, 41(4), 1-11.
- Dell’Arciprete, D., Bersezio, R., Felletti, F., Giudici, M., Comunian, A., Renard, P. (2012) Comparison of three geostatistical methods for hydrofacies simulation: a test on alluvial sediments. *Hydrogeology Journal*, 20, 299-311.
- De Marsily, G., Delay, F., Gonçalvès, J., Renard, P., Teles, V., Violette, S. (2005) Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13, 161-183.
- Deng, S., Chen, T., Yang, N., Qu, L., Li Mand Chen, D. (2018) Spatial and temporal distribution of rainfall and drought characteristics across the Pearl River basin. *Science Total Environment*, 619, 28-41, DOI: 10.1016/j.scitotenv.2017.10.339.

- De Silva, R.D. (2010) The effect of using different time steps in a soil water balance model to estimate groundwater recharge in the dry zone of Sri Lanka. *Journal of Environmental Hydrology*, 18.
- De Silva, C.S., Rushton, K.R. (1996) Interpretation of the behavior of agro-well systems in Sri Lanka using radial flow models. *Journal of Hydrological Sciences*, 41(6), 825-835.
- Deutsch, C.V. (1998) Cleaning categorical variable (lithofacies) realizations with maximum a posteriori selection. *Computer & Geosciences*, 24, 551-562, DOI: 10.1016/S0098-3004(98)00016.
- Deutsch, C.V. (2006) A sequential indicator simulation program for categorical variables with point and block data: BlockSIS. *Computer & Geosciences*, 32, 1669-1681, DOI:10.1016/j.cageo.2006.03.005.
- Deutsch, C.V., Tran, T.T. (2002) FLUVSIM: A program for object-based stochastic modelling of fluvial depositional systems. *Computer & Geosciences*, 28, 525–535.
- Díaz-Delgado, R., Aragonés, D., Afán, I., Bustamante, J. (2016) Long-term monitoring of the flooding regime and hydroperiod of Doñana marshes with landsat time series (1974-2014). *Remote Sensing*, 8(9).
- Díaz-Paniagua, C., Aragonés, D. (2015) Permanent and temporary ponds in Doñana National Park (SW Spain) are threatened by desiccation. *Limnetica*, 34(2),407–424.
- Díaz-Paniagua, C., Fernández-Zamudio, R., Florencio, M., García-Murillo, P., Gómez-Rodríguez, J.M., Siljeström, P., Serrano, L. (2008) The temporary ponds of Doñana: conservation value and present trends. *European Pond Conservation Network Newsletter*, 1, 5- 685.
- Díaz-Paniagua, C., Zamudio, R.F., Florencio, M., Murillo, P.G., Rodríguez, C.G., Porthault, A., Martín, L.S., Siljeström, P. (2010) Temporary ponds from Doñana National Park: A system of natural habitats for the preservation of aquatic flora and fauna. *Limnetica*, 29, 41–58.
- Diersch, H.J.G. (2013) FEFLOW: Finite Element Modeling of flow, mass and heat transport in porous and fractured media. Springer Science & Business Media, Berlin.
- Dimitriou, E., Moussoulis, E., Díaz-Paniagua, C., Serrano, L. (2017) Hydrodynamic numerical modelling of the water level decline in four temporary ponds of the Doñana National Park (SW Spain). *Journal of Arid Environments*, 147, 90-102, DOI:10.1016/j.jaridenv.2017.09.004.
- Dogrul, E.C., Brush, C.F. (2016) Groundwater modeling in support of water resources management and planning under complex climate, regulatory and economic stresses. *Water*, 8, 592, DOI: 10.3390/w8120592.
- Dourte, D., Shukla, S., Singh, P., Haman, D. (2013) Rainfall Intensity-Duration-Frequency Relationships for Andhra Pradesh, India: Changing Rainfall Patterns and Implications for Runoff and Groundwater Recharge. *Journal of Hydrologic Engineering*, 18(3), 324-330, DOI:10.1061/(ASCE)HE.1943-5584.0000625.
- Du, X., Lu, X., Hou, J., Ye, X. (2018) Improving the reliability of numerical groundwater modeling in a data-sparse region. *Water*, 10(3), 289, DOI: 10.3390/w10030289.

- Dunn, J.C. (1973) A fuzzy relative of the ISODARA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics* 3, 32-57.
- Erostate, M., Huneau, F., Garel, E., Ghiotti, S., Vystavna, Y., Garrido, M., Pasqualini, V. (2020) Groundwater dependent ecosystems in coastal Mediterranean Regions: Characterization, challenges and management for their protection. *Water Research*, DOI:10.1016/j.watres.2019.115461.
- Fernández-Ayuso, A., Aguilera, H., Guardiola-Albert, C., Rodríguez-Rodríguez, M., Heredia, J., Naranjo-Fernández, N. (2019) Unraveling the hydrological behavior of a coastal pond in Doñana National Park (southwest Spain). *Groundwater*, 57(6), 895-906, DOI:10.1111/gwat.12906.
- Fiori, A., Cvetkovic, V., Dagan, G., Attinger, S., Bellin, A., Dietrich, P., Zech, A., Teutsch, G. (2016) Debates—Stochastic subsurface hydrology from theory to practice: The relevance of stochastic subsurface hydrology to practical problems of contaminant transport and remediation. What is characterization and stochastic theory good for? *Water Resources Research*, 52, 9228–9234, DOI:10.1002/2015WR017525.
- Freeze, R.A. (1971) Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. *Water Resources Research*, 7(2), 347-366.
- Freeze, R.A., Cherry, J.A. (1979) *Groundwater*. Prentice-Hall Inc., Englewood Cliffs, Vol. 7632,604.
- Fundación Doñana 21. Bases estratégicas para una agricultura sostenible en Doñana. Área de agricultura.2003.<http://donana.es/source/BASES%20ESTRATEGICAS%20PARA%20UNA%20AGRICULTURA%20SOSTENIBLE%20EN%20DO%3%91ANA.pdf>.
- García Murillo, P., Sousa Martín, A. (1999) El paisaje vegetal de la zona oeste del Parque Natural de Doñana (Huelva). *Lagascalia*, 21 (1), 111-131.
- Ghazavi, R., Ebrahimi, H. (2019) Predicting the impacts of climate change on groundwater recharge in an arid environment using modeling approach. *International Journal of Climate Change Strategies and Management*, 11(1), 88-99, DOI: 10.1108/IJCCSM-04-2017-0085.
- Geóvariances. (2009) *Isatis Technical Ref.*, ver. 2012.4. Geovariances & Ecole Des Mines De Paris. Avon Cedex, France, 2009.
- Gómez-Hernández, J.J., Srivastava, R. M. (1990) ISIM3D: An ANSI-C three-dimensional multiple indicator conditional simulation program. *Computer & Geosciences*, 16, 395-440.
- Goovaerts, P. (1996) Stochastic simulation of categorical variables using a classification algorithm and simulating annealing. *Mathematical Geology*, 28, 909-921, DOI:10.1007/BF02066008.
- Gräler, E., Pebesma, E., Heuvelink, G. (2016) Spatio-Temporal Interpolation using gstat. *R Journal*, 8(1), 204-2018.
- Green, A.J., Alcorlo, P., Peeters, E.T., Morris, E.P., Espinar, J.L., Bravo-Utrera, M.A., Bustamante, J., Díaz-Delgado, R., Koelmans, A., Mateo, R., Mooij, W., Rodríguez Rodríguez, M., Van Nes,

- E., Scheffer, M. (2017) Creating a safe operating space for wetlands in a changing climate. *Frontiers in Ecology and the Environment*, 15(2), 99–107, DOI: 10.1002/fee.1459.
- Gringarten, E., Deutsch, C.V. (1999) Methodology for Improved Variogram Interpretation and Modeling for Petroleum Reservoir Characterization. SPE Annual Technical Conference and Exhibition, 3-6 October, Houston, Texas. Society of Petroleum Engineers, 1999, DOI:10.2118/56654-MS.
- Gringarten, E., Deutsch, C.V. (2001) Teacher's Aide. Variogram Interpretation and Modeling. *Mathematical Geology*, 3.
- Guardiola-Albert, C., García-Bravo, N., Mediavilla, C., Martín Machuca, M. (2009) Gestión de los recursos hídricos subterráneos en el entorno de Doñana con el apoyo del modelo matemático del acuífero Almonte-Marismas. *Boletín Geológico y Minero*, 120(3), 361-376, ISSN: 0366-0176.
- Guardiola-Albert, C., Jackson, C.R. (2011) Potential impacts of climate change on groundwater supplies to the Doñana wetland, Spain. *Wetlands*, 31, 907–920.
- Guardiola-Albert, C., Mediavilla-Laso, C., Aguilera, H., Fernández-Naranjo, N., Ruiz-Bermudo, F., García-Bravo, N. (2016) Recurso natural o recarga en la gestión del sistema acuífero Almonte-Marismas (Doñana) según la revisión del Plan Hidrológico (2016-2021) de la Demarcación Hidrográfica del Guadalquivir. Las aguas subterráneas y la planificación hidrológica. Congreso hispano-luso. AIH-GE. Madrid. ISBN: 978-84-938046-5-7: 193-199.
- Guijo-Rubio, D., Durán-Rosal, A.M., Gutiérrez, P.A., Troncoso, A., Hervás-Martínez, C. (2018) Time series clustering based on the characterisation of segment typologies. arXiv:1810.11624v1.
- Gutjahr, A.L., Gelhar, L.W., Bakr, A.A., MacMillan, J.R. (1978) Stochastic analysis of spatial variability in subsurface flows. 2. Evaluation and application. *Water Resources Research*, 14(5), 953-959, DOI: 10.1029/WR014i005p00953.
- Harbaugh, A.W., Banta, E.R., Hill, M.C., McDonald, M.G. (2000) The U.S. Geological Survey Modular Ground-Water Model – User Guide to Modularization Concepts and the Ground-Water Flow Process. U.S Geological Survey Research Open File Report 00-92, Reston, Virginia.
- Harbaugh, A.W. (2005) MODFLOW-2005, The US Geological Survey Modular Ground-Water Model: The Ground-WaterFlow Process; Department of the Interior, Techniques and Methods 6–A16. U.S. Geological Survey. Healy RW (2010) Estimating Groundwater Recharge. Cambridge University Press. 245 pp, DOI: .1017/CBO9780511780745.
- Hargreaves, G.H. (1989) Accuracy of estimated reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 115(6),1000-1007.
- Hartigan, J.A., Wong, M.A. (1979) Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 28, 100-108, DOI: 10.2307/2346830.
- Hendricks, J.M., Walcker, G.R. (1997) Recharge from precipitation. Recharge of Phreatic Aquifers in (Semi-) Arid Areas. Róterdam: A.A. Balkema, 1997. p. 19-111.

- Herrera, S., Gutiérrez, J.M., Ancell, R., Pons, M.R., Frías, M.D., Fernández, J. (2012) Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). *International Journal of Climatology*, 32,74–85.
- Hu, L.Y., Chugunova, T. (2008) Multiple-point geostatistics for modeling subsurface heterogeneity: a comprehensive review. *Water Resources Research*, 44(11), DOI:10.1029/2008WR006993.
- IGME. (1976) Informe preliminar sobre la realización y resultados del modelo matemático Almonte Marismas. 41 pp + anexos.
- IGME. (1982) Modelo matemático bidimensional del sistema acuífero nº 27: Unidad Almonte-Marismas. 93 pp+ anexos.
- IGME. (1983) Hidrogeología del Parque Nacional de Doñana y su entorno. Colección Informe, Madrid, 120 p.
- IGME. (1992) Hidrogeología del Parque Nacional de Doñana y su entorno. Colección Informes Aguas Subterráneas y Geotecnia. Ed. IGME. 64 pp + 2 mapas.
- IGME. (1994) Modelo matemático del acuífero de Almonte-Marismas (Parque Nacional de Doñana).
- IGME. (1999) Modelo regional de flujo subterráneo del sistema acuífero Almonte-Marismas y su entorno.
- IGME. (2005) Modelo matemático del acuífero Almonte—Marismas: actualización y realización en Visual Modflow. Available online: https://www.researchgate.net/publication/264229860_Modelo_matematico_revisado_d_el_acuifero_Almonte-Marismas_aplicacion_a_distintas_hipotesis_de_gestion (Accessed 25 September 2018).
- IGME. (2007) Revisión de la información geofísica existente en el acuífero Almonte-Marismas (Doñana). Available online: http://info.igme.es/SidPDF/137000/833/137833_0000001.pdf and http://info.igme.es/SidPDF/137000/834/137834_0000001.pdf (Accessed on 22 May 2018).
- IGME. (2009) Mejora del modelo matemático del acuífero Almonte-Marismas como apoyo a la gestión de los recursos hídricos: estimación de la recarga, modelo estocástico y actualización. Informe inédito.
- IGME. (2009) Informe técnico sobre los sondeos mecánicos de reconocimiento geológico e hidrogeológico realizados en el acuífero Almonte-Marismas.
- IGME. (2011) Proyecto para la actualización del modelo numérico (Visual Modflow 4.3, IGME 2007) del sistema acuífero Almonte-Marismas, como apoyo para el desarrollo de un modelo de gestión y uso sostenible del acuífero en el Ámbito de Doñana (Huelva-Sevilla).
- IHOBE. (2006) Guía técnica para la aplicación de modelos informáticos para el transporte y el flujo de contaminantes en el agua subterránea. Ley I/2005 para la prevención y corrección de la contaminación del suelo. Sociedad pública de Gestión ambiental. Disponible online:

<https://www.ihobe.eus/publicaciones/guia-tecnica-para-aplicacion-modelos-informaticos-para-transporte-y-flujo-contaminantes-en-agua-subterranea>.

- Illman, W.A., Zhu, J., Craig, A.J., Danting, Y. (2010) Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study. *Water Resources Research*, 46, DOI: 10.1029/2009WR007745.
- IPCC. International Panel on Climate Change. (2007) Climate Change 2007: Synthesis Report. Summary for Policymakers. An Assessment of the Intergovernmental Panel on Climate Change.
- Irawan, D.E., Puradimaja, D.J., Notosiswoyo, S., Soemintadiredja, P. (2009) Hydrogeochemistry of volcanic hydrogeology based on cluster analysis of Mount Ciremai, West Java, Indonesia. *Journal of Hydrology*, 376, 221-324, DOI: 10.1016/j.jhydrol.2009.07.033.
- Jankowski, J., Beck, P. (2000) Aquifer heterogeneity: hydrogeological and hydrochemical properties of the Botany Sands aquifer and their impact on contaminant transport. *Australian Journal of Earth Sciences*, 47, 45-64.
- Johnson, N.M., Dreiss, S. J. (1989) Hydrostratigraphic interpretation using indicator geostatistics. *Water Resources Research*, 25(12), 2501-2510.
- Jones, N.L., Lemon, A., Kennard, M.J. (2005) Efficient data management strategies for large MODFLOW models. World Water and Environmental Resources Congress, DOI: 10.1061/40792(173)365.
- Kassambara, A., Mundt, F. (2017) Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. <http://www.sthda.com/english/rpkgs/factoextra>.
- Ketchen, D.J., Shook, C.L. (1996) The application of cluster analysis in Strategic Management Research: An analysis and critique. *Strategic Management Journal*, 17(6), 441-458.
- Kilibarda, M., Hengl, T., Heuvelink, G.B.M., Gräler, B., Pebesma, E., Percec Tadic, M., Bajat, B. (2014) Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres*, 119, 2294-2313, DOI:10.1002/2013JD020803.
- Klingbeil, R., Kleineidam, S., Asprion, U., Aigner, T., Teutsch, G. (1991) Relating lithofacies to hydrofacies: outcrop-based hydrogeological characterisation of Quaternary gravel deposits. *Sedimentary Geology*, 129, 299-310.
- Kohfahl, C., Molano-Leno, L., Guardiola-Albert, C., Manzano, M., Ruiz, F., Naranjo-Fernández, N., Rebollo, A., Martínez, D., Martínez, A., Heredia, J., Custodio, E., Vanderlinden, K., Horacio Higuera J., Giráldez, V. (2019) Monitoring and current research of groundwater resources and recharge in the Doñana Natural area (SE Spain). *Boletín Geológico y Minero*, 130 (4), 661-690, ISSN: 0366-017. DOI: 10.21701/bolgeomin.130.4.002.
- Kohfahl, C., Molano-Leno, L., Martínez, G., Vanderlinden, K., Guardiola-Albert, C., Moreno, L. Determining groundwater recharge and vapor flow in dune sediments using a weighable precision meteor lysimeter. *Science of the Total Environment*, 656, 550-557, DOI: 10.1016/j.scitotenv.2018.11.415.

- Konikow, L.F., Bredehoeft, J.D. (1992) Ground-water models cannot be validated. *Advances in Water Resources*, 15(1),75-83.
- Krige, D.G. (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of Chemincal Metal Mining Society of South Africa*, 55, 119-139.
- Kumar, C.P. (2019) An overview of commonly used groundwater modelling software. *International Journal of Advanced Research in Science, Engineering and Technology*, 6(1).
- Kumar, S.H., Nagaraj, M.K. (2018) Asessment of interactions between river and aquifer in the Gowri-hole sub-cathment. *Journal of the Geological Society of India*, 92, 435-440.
- Kryszczuk K., Hurley P. (2010) Estimation of the Number of Clusters Using Multiple Clustering Validity Indices. In: El Gayar N., Kittler J., Roli F. (eds) Multiple Classifier Systems. MCS 2010. Lecture Notes in Computer Science, vol 5997. Springer, Berlin, Heidelberg 2010.
- Lario, J., Zazo, A.J., Plater, A.J., Goy, J.L., Dabrio, C.J., Borja, F., Sierro, F., Luque, L. (2001). Particle size and magnetic properties of Holocele estuarine deposits from the Doñana National Park (SW Iberia): Evidence of gradual and abrupt coastal sedimentation. *Z. Geomorphology*, 45, 33-45.
- Lerner, D.N. (1990) Techniques. Groundwater recharge. A guide to understanding the natural recharge. Hannover. Ed. R. van Acken GmbH, 1990. p. 99-229.
- Lerner, D.N., Issar, A.S., Simmers, I. (1990) Groundwater recharge. A guide to understanding natural recharge. Hannover. Ed. R. van Acken GmbH. 345p.
- Li, L., Srinivasan, S., Zhou, H., Gomez-Hernandez, J. (2015) Two-point or multiple-point statistics? A comparison between the ensemble Kalman filtering and the ensemble pattern matching inverse methods. *Advances in Water Resources*, 297-310.
- Li, L., Zhou, H., Gómez-Hernández, J.J. (2011) Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media. *Advances in Water Resources*, 478-489.
- Li, L., Zhou, H., Gómez-Herández, J.J. (2011) A comparative study of three-dimensional hydraulic conductivity upscaling at the macro-dispersion experiment (MADE) site, Columbus Air Force Base, Mississipi (USA). *Journal of Hydrology*, 278-293.
- Llamas, M.R. (1988) Conflicts between wetland conservation and groundwater exploitation: Two case histories in Spain. *Environmental Geology and Water Sciences*, 11, 241–251, DOI: 10.1007/BF02574814.
- Llamas, R. (2005) Lecciones aprendidas en tres décadas de gestión de las aguas subterráneas en España y su relación con los ecosistemas acuáticos. Universidad Autónoma de Madrid.
- Lorenzo-Lacruz, J., García, C., Morán-Tejada, E. (2017) Groundwater level responses to precipitation variability in Mediterranean insular aquifers. *Journal of Hydrology*, 552, 516-531, DOI: 10.1016/j.hydrol.2017.07.011.
- Lubis, R.F. (2017) Urban hydrogeology in Indonesia. IOP Conf. Series: Earth and Environmental Science, 118 (2018) 012022, DOI:10.1088/1755-1315/118/1/012022.

- Lukjan, A., Chalermyanont, T. (2017) Assesment of alluvial aquifer heterogeneity and development os stochastic hydrofacies models for the Hat Yai Basin in Sothern Thailand. *Environmental Earth Sciences*, 76.
- Ly, S., Charles, C., Degré, A. (2011) Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Science*, 15, 2259–2274, DOI:10.5194/hess-15-2259-2011.
- Ly, S., Charles, C., Degré, A. (2013) Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. *Biotechnology, Agronomy, Society and Environment*, 17(2), 392-406.
- Lu, S., Molz, F.J., Fogg, G.E., Castle, J.W (2002) Combining stochastic facies and fractal models for representing ntural heterogeneity. *Hydrogeology Journal*, 10, 475-482.
- MacCormack, R.W., Paullay, A.J. (1972) Computational efficiency achieved by time splitting of finite difference operators. AIAA paper 72-154, San Diego.
- MacQueen, J. (1967) Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1967, Volume 1: Statistics, 281–97. Berkeley, Calif.: University of California Press. <http://projecteuclid.org:443/euclid.bsm/1200512992>.
- Manzano, M. (2001) Clasificación de los humedales de Doñana atendiendo a su funcionamiento hidrológico. *Hidrogeología y Recursos Hidráulicos*, 24, 57–75.
- Manzano, M., Custodio, E., Higuera, H., Puig, R., Soler, A. (2009) Influencia de la gestión del acuífero sobre los humedales del manto eólico de Doñana. *Boletín Geológico y Minero*, 120 (3), 377-392, ISSN: 0366-0176.
- Manzano-Arellano, M.S., Custodio-Gimena, E., Montes, C., Mediavilla, C. (2009) Groundwater quality and quantity assessment through a dedicated monitoring network: the Doñana aquifer experience (SW Spain). Groundwater monitoring. John Wiley & Sons, 2009, p-273-287.
- Manzano, M., Borja, F., Montes, C. (2002) Metodología de tipificación hidrología de los humedales españoles con vistas a su valoración funcional y a su gestión. Aplicación a los humedales de Doñana. *Boletín Geológico y Minero*, 113(3), 313-330, ISSN: 0366-0176.
- Manzione R.L., Wedland, E., Tanikawa, D.H. (2012) stochastic simulation of time-series models combined with geostatistics to predict water-table scenarios in a Guarani Aquifer System outcrop area, Brazil. *Hydrogeology Journal*, 20, 1239-1249.
- Marín Celestino, A.E., Martínez Cruz, D.A., Otazo Sánchez, E.M., Gavi Reyes, F., Vásquez Soto, D. (2018) Groundwater Quality Assessment: An Improved Approach to K-Means Clustering, Principal Component Analysis and Spatial Analysis: A Case Study. *Water*, 10, 437, DOI: 10.3390/w10040437.

- Martín Machuca, M., López Geta, J.A., López Vílches, L., Mantecón Gómez, R., Cantos Robles, R., Coleto Fiaño, I. (1992) Hidrogeología del Parque Nacional de Doñana y su entorno. Instituto Tecnológico Geominero de España. IGME 1992.
- Marini, M., Felletti, F., Bereta G.P., Terrenghi, J. (2018) Three Geostatistical Methods for Hydrofacies Simulation Ranked Using a Large Borehole Lithology Dataset from Venice Hinterland (NE Italy). *Water*, 10, 844, DOI:10.3390/w10070844.
- Matheron, G. (1965) Les variables régionalisées et leur estimation. Une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson, Paris, 1965, 306 p.
- Maulik, U., Bandyopadhyay, S. (2002) Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12), 1650–1654.
- McDonald, P.W. (1971) The computation of transonic flow through two-dimensional gas turbine cascades. ASME paper 71-GT-89.
- McDonald, M.G., Harbaugh, A.W. (1998) A modular three-dimensional finite difference groundwater flow model: U.S. Geological Survey, Techniques of Water-Resources Investigations, 6.
- Ministerio para la transición ecológica y el reto demográfico (MITECO). Esquemas de Temas Importantes del Tercer Ciclo de Planificación (2021-2027). https://www.miteco.gob.es/es/agua/temas/planificacion-hidrologica/planificacion-hidrologica/ETI_tercer_ciclo.aspx [Consultada: 14-04-2020]
- Ministerio para la transición ecológica y el reto demográfico (MITECO). <https://www.miteco.gob.es/es/red-parques-nacionales/nuestros-parques/donana/valores-naturales/ecosistemas.aspx>. [Consultada: 24-03-2020]
- Monjo, R., Martín-Vide, J. (2016) Daily precipitation concentration around the world according to several indices. *International Journal of Climatology*, 36(11), DOI:10.1002/joc.4596.
- Moschou, E.C., Batelis, S.C., Dimakos, Y., Fountoulakis, I., Markonis, Y., Papalexiou, S.M., Mamassis, N., Koutsoyiannis, D. (2013) Spatial and temporal rainfall variability over Greece. 5th EGU Leonardo Conference, Hydrofractals 2013. STAHY '13, 2013.
- Mumtaz, R., Baig, S., Kazmi, S.S.A., Ahmad, F., Fatima, I., Ghauri, B. (2018) Delineation of groundwater prospective resources by exploiting geospatial decision-making techniques for the Kingdom of Saudi Arabia. *Neural Computing and Applications*, DOI:10.1007/s00521-018-3370-z.
- Muñoz-Reinoso, J.C. (2001) Vegetation changes and groundwater abstraction in SW Doñana, Spain. *Journal of Hydrology*, 242 (3-4), 197-209.
- Muthusamy, M., Schellart, A., Tait, S., Heuvelink, G.B.M. (2016) Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models. *Hydrology and Earth Systems Sciences*, 21, 1077-1091, DOI:10.5194/hess-2016-279.

- Nakagawa, K., Yu, Z.Q., Berndtsson, R., Kagabu, M. (2019) Analysis of earthquake-induced groundwater level change using self-organizing maps. *Environmental Earth Sciences*, 78, DOI: 10.1007/s12665-019-8473-z.
- Nandargi, S., Mulye, S.S. (2012) Relationships between rainy days, mean daily intensity, and seasonal rainfall over the Koyna catchment during 1961-2005. *Scientific World Journal*, DOI:10.1100/2012/894313.
- Naranjo-Fernández, N., Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Rodríguez-Rodríguez, M., Fernández-Ayuso, A., Ruíz-Bermudo, F., Montero-González, E. (2020) Relevance of spatio-temporal rainfall variability regarding groundwater management challenges under global change: case study in Doñana (SW Spain). *Stochastic Environmental Research and Risk Assessment*, DOI: 10.1007/s00477-020-01771-7.
- Narjary, B., Kumar, S., Kamra, S.K., Bundela, D.S., Sharma, D.K. (2014) Impact of rainfall variability on groundwater resources and opportunities of artificial recharge structure to reduce its exploitation in fresh groundwater zones of Haryana. *Current Science*, 107(8), 1305-1312.
- Negreiros, J., Painho, M., Aguilar, F., Aguilar, M. (2010) Geographical Information Systems Principles of Ordinary Kriging Interpolator. *Journal of Applied Sciences*, 10(11), 852-867, DOI:10.3923/jas.2010.852.867.
- Nguyen, T.T., Kawamura, A., Tong, T. N., Nakagawa, N., Amaguchi, H., Gilbuena, R. (2015) Clustering spatio-seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. *Journal of Hydrology*, 522, 661–673, DOI: 10.1016/j.hydrol.2015.01.023.
- Noëtinger, B., Artus, V., Zargar, G. (2005) The future of stochastic and upscaling methods in hydrogeology. *Hydrogeology Journal*, 13, 184-201.
- Nuñez Muñoz, J.L. (2001) Aplicación de los modelos matemáticos de flujo y transporte de contaminantes para el diseño de sistemas de remediación. Investigación, gestión y recuperación de acuíferos contaminados. Alicante. Octubre 2001.
- Oficial Journal of the European Commission. Case C-559/19. Accion brought on 22 July 2019 (2019/C34/10) European Commission vs Kingdom of Spain. Disponible online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62019CN0559&from=GA> [Acceso 10 de abril de 2020]
- Olías Álvarez, M., Rodríguez Rodríguez, M. (2013) Evolución de los niveles en la red de control piezométrica del acuífero Almonte-Marismas (periodo 1994-2012). *Hidrogeología y Recursos Hidráulicos* (30),1.121-1.130. Madrid: Instituto Geológico y Minero de España. X Simposio de Hidrogeología, Granada 2013.
- Paparrizos, J., Gravano, L. (2015) k-Shape: Efficient and Accurate Clustering of Time Series." In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, SIGMOD 2015, pp. 1855–1870. ACM, New York, NY, USA. ISBN 978-1-4503-2758-9, DOI:10.1145/2723372.2737793.
- Padilla, A., Delgado, J. (2013) Tratamiento y gestión de series temporales hidrológicas: manual del usuario: aplicación Trasero 2.0. Departamento de Ciclo Hídrico, Diputación de Alicante, 87 p.

- Parkhurst, D.L. (1995) User's guide to PHREEQC--A computer program for speciation, reaction-path, advective-transport, and inverse geochemical calculations: U.S. Geological Survey Water-Resources Investigations Report 95-4227, 143 p.
- Pebesma, E. (2012) Spacetime: Spatio-temporal data in R. *Journal of Statistical Software*, 51(7), 1-30.
- Pholkern, K., Saraphirom, P., Cloutier, V., Srisuk, K. (2019) Use of Alternative Hydrogeological Conceptual Models to Assess the Potential Impact of Climate Change on Groundwater Sustainable Yield in Central Huai Luang Basin, Northeast Thailand. *Water*, 11, 241.
- Pinder, G.F., Gray, W.G. (1977) Finite Element simulation in surface and subsurface hydrology. Academic Press, New York. 310 pp.
- Priyanka, B.N., Mohan Kumar, M.S. (2019) Three-dimensional modelling of heterogeneous coastal aquifer: upscaling from local scale. *Water*, 11, 421, DOI: 10.3390/w11030421.
- Pyrz, M.J., Deutsch, C.V. (2014) In: Geostatistical Reservoir Modeling, 2nd Ed. New York: Oxford University Press, 2014, p. 433.
- Rahman, M.M., Thompson, J.R., Flower, R.J. (2016) An enhanced SWAT wetland module to quantify hydraulic interactions between riparian depressional wetlands, rivers and aquifers. *Environmental Modelling & Software*, 84, 263-289, DOI: 10.1016/j.envsoft.2016.07.003.
- Rajaei, T., Ebrahimi, H., Nourani, V. (2019) A review of the artificial methods in groundwater level modeling. *Journal of Hydrology*, 572, DOI: 10.1016/j.hydrol.2018.12.037.
- Remy, N., Boucher, A., Wu, J. (2009) Applied geostatistics with SGeMS: a user's guide. Cambridge University Press, New York, 2009.
- Renard, P. (2007) Stochastic Hydrogeology: What Professionals Really Need?. *Groundwater*, 45, 531-541.
- Rinderer, M., Van Meerveld, H.J., McGlynn, B.L. (2019) From points to patterns: Using groundwater time series clustering to investigate subsurface hydrological connectivity and runoff source area dynamics. *Water Resources Research*, 55, 5784-5806, DOI:10.1029/2018WR023886.
- Rodríguez-Ramírez, A., Yáñez-Camacho, C., Gascó, C., Clemente-Salas, L., Antón, M.P. (2005) Colmatación natural y antrópica de las marismas del Parque Nacional de Doñana: Implicaciones para su manejo y conservación. *Cuaternario y Geomorfología*, 19, 37-48.
- Rodríguez-Rodríguez, M., Fernández-Ayuso, A., Moral, F. (2017) Cuantificación de los aportes hídricos subterráneos a la laguna de Santa Olalla a partir de balances hídricos diarios (Parque Nacional de Doñana, Huelva). *Geogaceta*, 61, 47-50.
- Roldán, F.J., Rodríguez-Fernández, J., Villalobos, M., Lastra, J., Díaz-Pinto, G., Pérez Rodríguez, A.B. Mapa Geológico Digital Continuo E. 1:50.000, Zonas Subbético, Cuenca del Guadalquivir y Campo de Gibraltar. In GEODE. Mapa Geológico Digital continuo de España. Available online:

<http://info.igme.es/cartografiadigital/geologica/geodezona.aspx?Id=Z2600> (Accessed on 22 May 2018).

- Ruiz, F., Rodríguez-Ramírez, A., Cáceres, L.M., Rodríguez, J., Carretero, M.I., Clemente, L., Muñoz, J.M., Yáñez, C., Abad, M. (2004). Late Holocene evolution of the southwestern Doñana National Park (Gudalquivir Estuary, SW Spain): a multivariate approach. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 204, 47-64.
- Ruiz-Villanueva, V., Bodoque, J.M., Díez-Herrero, A., Eguibar, M.A., Pardo-Igúzquiza, E. (2013) Reconstruction of a flash flood with large wood transport and its influence on hazard patterns in an ungauged mountain basin. *Hydrological Processes*, 27, 3424–3437, DOI:10.1002/hyp.9433.
- Rushton, K. (1998) Numerical and conceptual models for recharge estimation in arid and semi-arid zones. Estimation of Natural Groundwater Recharge. Boston. Ed Reidel, NATO ASI Series, 1998. P. 223-237.
- Ruybal, C.J., Hogue, T.S., McCray, J.E. (2019) Evaluation of Groundwater Levels in the Arapahoe Aquifer Using Spatiotemporal Regression Kriging. *Water Resources Research*, 55, DOI:10.1029/2018wr023437.
- Saaltink, M.W., Kohfahl, K., Molano-Leno, L. (2020) Analysis of water vapor absorption in soils by means of a lysimeter and numerical modeling. *Vadose Zone Journal*, 19(1), DOI: 10.1002/vzj2.20012.
- Sacks, L.A., Herman, J.S., Konikow, L.F., Vela, A.L. (1992) Seasonal Dynamics of Groundwater-Lake Interactions of Doñana National Park, Spain. *Journal of Hydrology*, 136, 123-154, DOI:10.1016/0022-1694(92)90008-J.
- Saha, G.C., Li, J., Thring, R.W. (2017) Understanding the effects of parameter uncertainty on temporal dynamics of groundwater-surface interaction. *Hydrology*, 4, 28.
- Sahoo, S., Russo, T.A., Elliot, J., Foster, I. (2017) Machine learning algorithms for modelling groundwater level changes in agricultural regions of the U.S. *Water Resources Research*, 53(5), 3878-3895, DOI: 10.1002/2016WE019933.
- Salvany, J.M., Custodio, E. (1995) Características litoestratigráficas de los depósitos pliocuaternarios del bajo Guadalquivir en el área de Doñana: implicaciones hidrogeológicas. *Revista de la Sociedad Geológica de España*, 8, 21–31.
- Salvany, J.M., Mediavilla, C., Rebollo, A. (2010) Las formaciones Plio-Cuaternarias de El Abalario en el litoral de la provincia de Huelva (España). *Estudios Geológicos*, 66, DOI: 10.3989/egeol.39953.072.
- Salvany, J.M., Larrasoana, J.C., Mediavilla, C., Rebollo, A. (2011) Chronology and tectono-sedimentary evolution of the Upper Pliocene to Quaternary deposits of the lower Guadalquivir foreland basin, SW Spain. *Sedimentary Geology*, 241, 22-39.
- Samdi, M.M., Zghoul, A. (2006) A sudden change in rainfall characteristics in Amman, Jordan during the mid. *The American Journal of Environmental Sciences*, 2,84-91.

- Samper, F.J. (1997) Métodos de evaluación de la recarga por la lluvia por balance de agua: utilización, calibración y errores. La evaluación de la recarga a los acuíferos en la planificación hidrológica. TEXTOS DEL SEMINARIO CELEBRADO EN LAS PALMAS DE GRAN CANARIA. Instituto Tecnológico Geominero de España. Madrid. 1997. p. 41-79.
- Sánchez-Vila, X., Fernández-García, D. (2007) Gestión de los recursos hídricos: los modelos hidrogeológicos como herramienta auxiliar. *Enseñanza de las ciencias de la tierra*, 15, 250-256.
- Sapriza-Azuri, G., Jódar, J., Carrera, J., & Gupta, H. V. (2013) Stochastic simulation of nonstationary rainfall fields, accounting for seasonality and atmospheric circulation pattern evolution. *Mathematical geosciences*, 45(5), 621-645. [DOI:10.1007/s11004-013-9467-0](https://doi.org/10.1007/s11004-013-9467-0)
- Sapriza-Azuri, G., Jódar, J., Navarro, V., Slooten, L. J., Carrera, J., & Gupta, H. V. (2015) Impacts of rainfall spatial variability on hydrogeological response. *Water Resources Research*, 51(2), 1300-1314. DOI: 10.1002/2014WR016168
- Sardá-Espinosa, A. (2019) Time-Series Clustering in R Using the dtwclust Package. *The R Journal*, 11, ISSN 2073-4859.
- Scanlon, B.R., Healy, R.W., Cook, P.G. (2002) Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10(1), 18-39.
- Scheffer, M., Barrett, S., Carpenter, S.R., Folke, C., Green, A.J., Holmgren, M., Hughes, T.P., Kosten, S., Van de Leemput, I.A., Nepstad, D.C., Van Nes, E.H., Peeters, E.T.H.M., Walker, B. (2015) Creating a safe operating space for iconic ecosystems. *Science*, 347, 1317-1319.
- SCS. Soil Conservation Service. (1975) Urban hydrology for small watersheds. Technical realize. No.55 U.S. Department of Agriculture, January 1975.
- Seifert, D., Jensen, J.L. (1991) Using sequential indicator simulation as a tool in reservoir description: issues and uncertainties. *Mathematical Geology*, 31, 527-550, DOI:10.1023/A:1007563907124.
- Seifert, D., Jensen, J.L. (1991) Using sequential indicator simulation as a tool in reservoir description: issues and uncertainties. *Mathematical Geology*, 31, 527-550, DOI:10.1023/A:1007563907124.
- Serrano, L., Serrano, L. (1996) Influence of Groundwater Exploitation for Urban Water Supply on Temporary Ponds from the Doñana National Park (SW Spain). *Journal of Environment Management*, 46(3), 229-238, DOI:10.1006/jema.1996.0018.
- Serrano, L., Reina, M., Martín, G., Reyes, I., Arechedera, A., León, D., Toja, J. (2006) The aquatic systems of Doñana (SW Spain): watersheds and frontiers. *Limnetica*, 25, 11-32.
- Serrano, L., Zunzunegui, M. (2008) The relevance of preserving temporary ponds during drought: Hydrological and vegetation changes over a 16-year period in the Doñana National Park (south-west Spain). *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18(3), 261-279, DOI: 10.1002/aqc.830.

- Shan, X., Tian, F., Cheng, F., Yang, C., Xin, W. (2019) Spectral decomposition and a waveform cluster to characterize strongly heterogeneous paleokarst reservoirs in the Tarim Basin, China. *Water*, 11, 256, DOI: 10.3390/w11020256.
- Shapoori, V., Peterson, T.J., Western, A.W., Costelloe, J.F. (2015) Top-down groundwater hydrograph time-series modeling for climate-pumping decomposition. *Hydrogeology Journal*, DOI: 10.1007/s10040-014-1223-0.
- Sherif, M., Singh, V.P. (1999) Effect of climate change on sea water intrusion in coastal aquifers. *Hydrological Processes*, 13, 1277-1287, DOI:10.1002/(SICI)1099-1085(19990615)13:8<1277::AID-HYP765>3.0.CO;2-W.
- Shi, P., Quiao, X., Chen, X., Zhou, M., Qu, S., Ma, X., Zhang, Z. (2014) Spatial distribution and temporal trends in daily and monthly precipitation concentration indices in the upper reaches of the Huai River, China. *Stochastic Environmental Research and Risk Assessment*, 28(2), 201-212, DOI:10.1007/s00477-013-0740-z.
- Silberstein, R.P. (2004) Hydrological models are so good, do we still need data?. *Environmental Modelling & Software*, 21(6), 1340-1352.
- Soares, A. (1998) Sequential indicator simulation with correction for local probabilities. *Mathematical Geology*, 30, 761-765, DOI:10.1023/A:1022451504120.
- Sousa, A., García-Murillo, P. (2003) Changes in the wetlands of Andalusia (Doñana Natural Park, SW Spain) at the end of the Little Ice Age. *Climatic Change*, 58(1-2), 193-217.
- Stekhoven, D.J. (2013) Nonparametric Missing Value Imputation using Random Forest. 1.4. version. CRAN repository. <http://www.r-project.org>, <https://github.com/stekhoven/missForest>.
- Stekhoven, D.J., Bühlmann, P. (2012) Miss-Forest? Non parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112-8, DOI: 10.1093/bioinformatics/btr597.
- Tang F, Ishwaran H (2017) Random forest missing data algorithms. *Statistical Analysis and Data Mining*, 10(6), 363-377, DOI:10.1002/sam.11348.
- Thorntwaite, C.W. (1948) An approach toward a rational classification of climate. *Geographical Review*, 38(1), 55-94, DOI:10.2307/210739.
- Trevisani, S., Fabbri, P. (2010) Geostatistical modeling of heterogeneous site bordering the Venice Lagoon, Italy. *Groundwater*, 48, 614-623, DOI: 10.1111/j.1745-6584.2009.00632.x.
- UPC. (1999) Regional groundwater Flow in the Almonte-Marismas aquifer. Groundwater Hydrology Group of the Technical University of Catalonia and Geological Institute of Spain, Madrid: 114pp.
- USGS. (2008) Estimates of Ground-Water Recharge in Minesota. https://www.eqo.state.mn.us/sites/default/files/documents/USGS_GW_Recharge_06_09_08.ppt. Accessed 20 October 2017.
- Van Dick, W.M.; Densmore, L.A.; Jackson, C.R.; Mackay, J.D.; Joshi, S.K.; Sinha, R.; Shekhar, S.; Gupta, S. Spatial variation of groundwater response to multiple drivers in a depleting

- alluvial aquifer system, northwestern India. *Progress in Physical Geography*, 1-26, DOI: 10.1177/0309133319871941.
- Van Zoest, V., Osei, F.B., Hoek, G. Stein, A. (2019) Spatio-temporal regression kriging for modelling urban NO₂ concentrations. *International Journal of Geographical Information Science*, DOI:10.1080/13658816.2019.1667501.
- Varouchakis, E.A., Corzo, G.A., Karatzas, G.P., Kotsopoulou, A. (2018) Spatio-temporal analysis of annual rainfall in Crete, Greece. *Acta Geophysica*, 66(3),319-328, DOI:10.1007/s11600-018-0128-z.
- Varouchakis, E.A., Hristopulos, D.T. (2019) Comparison of spatiotemporal variogram functions based on a sparse dataset of groundwater level variations. *Spatial Statistics*, 34, 100245, DOI:10.1016/j.spasta.2017.07.003.
- Vela, A. (1984) Estudio preliminar de la hidrogeología e hidrogeoquímica del sistema de dunas móviles y flecha litoral del Parque Nacional de Doñana. Tesina de Licenciatura. Universidad Complutense de Madrid. 221 p.
- Virgos, L., Martínez Alfaro, P.E., López Vilchez, L., Martín, M. (1983) Estudio del funcionamiento hidrogeológico del acuífero Almonte-Marismas (sistema no. 27) mediante un modelo digital bidimensional: Hidrogeología y Recursos Hidráulicos, vol. 9, p. 103–124.
- Vistrand, P. (2005) Comparison of upscaling methods to estimate hydraulic conductivity. *Groundwater*, 39(3), 401-407, DOI: 10.1111/j.1745-6584.2001.tb02324.x.
- Wackernagel, H. (1998) Multivariate geostatistics: an introduction with applications, 2nd Ed. Springer-Verlag, 1998.
- Watson, C., Richardson, J., Wood, B., Jackson, C., Hughes, A. (2015) Improving geological and process model integration through TIN to 3D grid conversion. *Computer & Geosciences*, 82, 45-54.
- Wikle, C.K., Zammit-Mangion, A., Cressie, N. (2019) Spatio-temporal Statistics with R. CRC Press, Taylor & Francis Group, 380 p.
- Winston, R.B. (1999) ModelMuse—A graphical user interface for MODFLOW–2005 and PHAST: U.S. Geological Survey Techniques and Methods 6–A29, 1999, 52 p. Available online: <http://pubs.usgs.gov/tm/tm6A29> (Accesed on 22 May 2018).
- WMO. World Meteorological Organization. (2003) Hydrological data management: present state and trend. WMO nº 964.[https://library.wmo.int/doc_num.php?explnum_id=4851]
- WMO. World Meteorological Organization (2014). Guidelines for Hydrological Data Rescue. WMO nº 1146.[http://www.wmo.int/pages/prog/hwrrp/publications/guidelines_hydrological_DR/wmo_1146_en.pdf]
- WWF. (2016) Salvemos Doñana. Del peligro a la prosperidad. [Visited 15th January 2020] http://awsassets.wwf.es/downloads/wwf_informe_salvemos_donana__2016.pdf?_ga=2.266321487.1975888914.1579078039-1333755382.1573218666.

- WWF. (2019) Doñana bajo plástico: Avanza la invasión de los frutos rojos. Seguimiento del incremento de la superficie de cultivos bajo plásticos en Doñana tras la entrada en vigor del Plan Especial de la Corona Forestal (2015-2019). https://d80g3k8vowjyp.cloudfront.net/downloads/donana_bajo_plastico.pdf
- Yang, X., Xie, X., Liu, D.L., Ji, F., Wang, L. (2015) Spatial Interpolation of Daily Rainfall Data for Local Climate Impact Assessment over Greater Sydney Region. *Advances in Meteorology*, article ID 563629, DOI:10.1155/2015/563629.
- Yuan, R.; Wang, M.; Wang, S.; Song, X. (2020) Water transfer imposes hydrochemical impacts on groundwater by altering the interaction of groundwater and surface water. *Journal of Hydrology*, 583, 124617, DOI: 10.1016/j.jhydrol.2020.124617.
- Yue, S., Wang, C.Y. (2002) Applicability of prewhitening to eliminate the influence of serial correlation on the Mann- Kendall test. *Water Resources Research*, 38(6), 41–47.
- Zeng, C. (1990) MT3D, A modular three-dimensional transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems. Rockville, Maryland. S.S. Papadopulos & Associates.
- Zhou, H., Li, L., Gómez-Hernández, J.J. (2010) Three-dimensional hydraulic conductivity upscaling in groundwater modeling. *Computer & Geosciences*, 1224-1235.
- Zimmerman, D.W. (1997) Teacher's corner: A note on interpretation of the paired-samples t test. *Journal of Educational and Behavioral Statistics*, 22(3), 349-360.
- Zurbietta, R., Saavedra, M., Silva, Y., Giráldez, L. (2017) Spatial analysis and temporal trends of daily precipitation concentration in the Mantaro River basin: central Andes of Peru. *Stochastic Environmental Research and Risk Assessment*, 31(6), 1305-1318., DOI:10.1007/s00477-016-1235-5.