PAPER

Increasing knowledge of the transmissivity field of a detrital aquifer by geostatistical merging of different sources of information

Eulogio Pardo-Igúzquiza[1] · Peter A. Dowd[2] · Juan A. Luque-Espinar[3] · Mario Chica-Olmo[4]

Abstract

Transmissivity is a significant hydrogeological parameter that affects the reliability of groundwater flow and transport models. This study demonstrates the improvement in the estimated transmissivity field of an unconfined detritic aquifer that can be obtained by using geostatistical methods to combine three types of data: hard transmissivity data obtained from pumping tests, soft transmissivity data obtained from lithological information from boreholes, and water head data. The piezometric data can be related to transmissivity by solving the hydrogeology inverse problem, i.e., including the observed water head to determine the unknown model parameters (log transmissivities). The geostatistical combination of all the available information is achieved by using three different geostatistical methodologies: ordinary kriging, ordinary co-kriging and inverse problem universal co-kriging. In addition, there are eight methodological cases to be compared according to which log-transmissivity data are considered as the primary variable in co-kriging and whether two or three variables are used in inverse-problem universal co-kriging. The results are validated by using the performance statistics of the direct modelling of the unconfined groundwater flow and comparing observed water heads with the modelled ones. Although the results show that the two sets of log-transmissivity data are incompatible, the set of log-transmissivity data from the lithofacies provides a good log-transmissivity image that can be improved by inverse modelling. The map provided by inverse-problem universal co-kriging provides the best results. Using three variables, rather than two in the inverse problem, gives worse results because of the incompatibility of the log-transmissivity data sets.

Keywords Geostatistics · Groundwater flow · Transmissivity · Lithofacies · Universal kriging

✉ Eulogio Pardo-Igúzquiza
e.pardo.iguzquiza@csic.es

Peter A. Dowd
peter.dowd@adelaide.edu.au

Juan A. Luque-Espinar
ja.luque@igme.es

Mario Chica-Olmo
mchica@ugr.es

[1] Instituto de Geociencias (IGEO, CSIC), Calle Severo Ochoa, 7, planta 4, 28040 Madrid, Spain

[2] Faculty of Sciences, Engineering and Technology, The University of Adelaide, Adelaide, Australia

[3] Instituto de Geológico y Minero de España (IGME, CSIC), Unidad de Granada, Urbanización Alcázar del Genil, 4. Edificio Zulema, 18006 Granada, Spain

[4] Departamento de Geodinámica, Universidad de Granada, Avenida Fuentenueva, 18071 Granada, Spain

Introduction

Unconfined detritic aquifers can be modelled as randomly heterogeneous porous media in which hydrogeological properties (e.g., transmissivity) vary spatially in a manner that is only partially understood (Kitanidis 1997). A reasonably accurate aquifer transmissivity field is essential for simulating the flow and transport of solutes (Anderson et al. 2015). Pumping tests are the main source of hard data for describing the transmissivity field of an aquifer (Renard 2005; Demir et al. 2017). However, in general, the number of pumping tests in a single aquifer is usually small (Christensen 1996; Teramoto et al. 2021); hence, transmissivity data have significant uncertainty, as will be shown by their spatial variability quantified by a variogram that shows a high nugget variance due to sampling error and/or short scales of spatial variability that is not accounted for by the available experimental data locations. Auxiliary information, however, can be used to

increase the accuracy of a transmissivity field estimated from only hard transmissivity data. Of the several types of secondary information, piezometric data (water head data for the unconfined aquifer case) have often been used to improve the estimation of the transmissivity field in inverse hydrogeology approaches (Dagan 1985; Zimmerman et al. 1998; Carrera et al. 2005; among many others). When there is a correlation between the water table and topographic elevations, a digital elevation model with full coverage of elevation data can be introduced as secondary information (Renard and Jeannée 2008). Other than inverse approaches, transmissivity fields have been estimated by geostatistical co-kriging using secondary variables such as specific capacity (Ahmed and de Marsily 1987; Razack and Huntley 1991; Richard et al. 2016), connectivity data (Freixas et al. 2017) or any auxiliary data correlated with transmissivity (Kupfersberger and Blöschol 1995). Correlation and regression analysis have also been used to improve transmissivity fields by using geophysical data such as magnetic resonance (Boucher et al. 2009) or electrical resistivity (Soupios et al. 2007). Hydrofacies simulation has also been used to estimate the spatial variability of hydraulic conductivity (Zhu et al. 2016; Xue et al. 2022).

This paper shows that different types of co-kriging (the multivariate geostatistical estimator) can be used to improve the knowledge of a transmissivity field of a detritic aquifer in southern Spain (the Vega de Granada aquifer). Different aspects of the geostatistical analysis of the Vega de Granada aquifer, such as the estimation of the variogram of transmissivities and the estimation of groundwater hydraulic gradients, have been reported in Pardo-Igúzquiza et al. (2009); Pardo-Igúzquiza and Chica-Olmo 2004, 2007 and Kuhlman and Pardo-Igúzquiza (2010). This paper introduces and demonstrates several new components of the methodology in a case study of the Vega de Granada aquifer. In the methodology three types of information are used: transmissivity data from pumping tests, transmissivity data from lithofacies measured in boreholes, and water head data. These three types of information are merged in order to improve the aquifer parameter field defined by spatially variable log-transmissivities. The geostatistical approach to the inverse model has been illustrated by applying it to the Vega de Granada aquifer and the performances of the procedures have been ranked. Rather than using cross-validation to rank performance, the paper shows the use of the direct solution of the groundwater equation which is derived in this paper. In this sense, this paper presents novel results and a novel application to the Vega de Granada as a case study. The data, the methodology, the results, and the discussion are presented in the following sections.

Materials and methods

Study area

The study area is the unconfined Vega de Granada detritic aquifer in southern Spain (Fig. 1a). The aquifer has a surface area of around 200 km^2 and is composed of Quaternary materials of fluvial origin (Castillo 1989; Luque-Espinar 2001; García-Soldado 2009; Pardo-Igúzquiza et al. 2011) which are part of the Granada intramountainous basin (Fig. 1b). The aquifer has an elongated shape in the E–W direction with a long axis of around 22 km and a perpendicular axis of around 8 km.

The aquifer is composed of various fluvial and colluvial detritic materials such as gravel, sand, silt and clay in spatially varying proportions. The spatial variations in the material proportions imply spatial variations in porosity and hence in hydraulic conductivity and transmissivity (taking into account the thickness of saturated aquifer). At places where gravel and sand dominate, the transmissivity is high, and where silt and clay dominate, the transmissivity is low.

The Genil River (Fig. 1c) crosses the elongated central section of the aquifer from east to west. However, this river is regulated by the Canales Dam and thus the water that it contributes to the aquifer can be considered negligible. For the same reason, the Cubillas River in the northwest contributes negligible amounts of water as it is regulated by the Cubillas Dam. The remaining rivers shown in Fig. 1c (Beiro, Darro, Monachil, Dilar and Salado rivers) are ephemeral streams that carry water only after significant rainfall events. Thus, the water input from rivers has no significant effect on the methodology used in the work presented here.

Hard transmissivity data

Figure 2a shows the spatial location of the 41 measurements of transmissivity obtained from pumping tests. Most of the measurements are located around the central part of the aquifer and this zone has the highest transmissivity. Pumping tests are rare in zones that are not water productive, which implies a bias in the transmissivity data set towards relatively high transmissivity values. It should be noted that the transmissivity data are not the product of a single acquisition but have been compiled from historical records available from the office of the Spanish Geological Survey in the city of Granada. The pumping tests were conducted by different teams and the details of every test are unknown (for example, whether or not a borehole fully penetrated the aquifer); thus, these measurements have an unknown uncertainty. In addition, they are not point support data as is usually assumed, but have a specific support that is generally unknown, which implies an additional source of uncertainty if they are used as point data. The original experimental data from the pumping test were measured in m^2/day
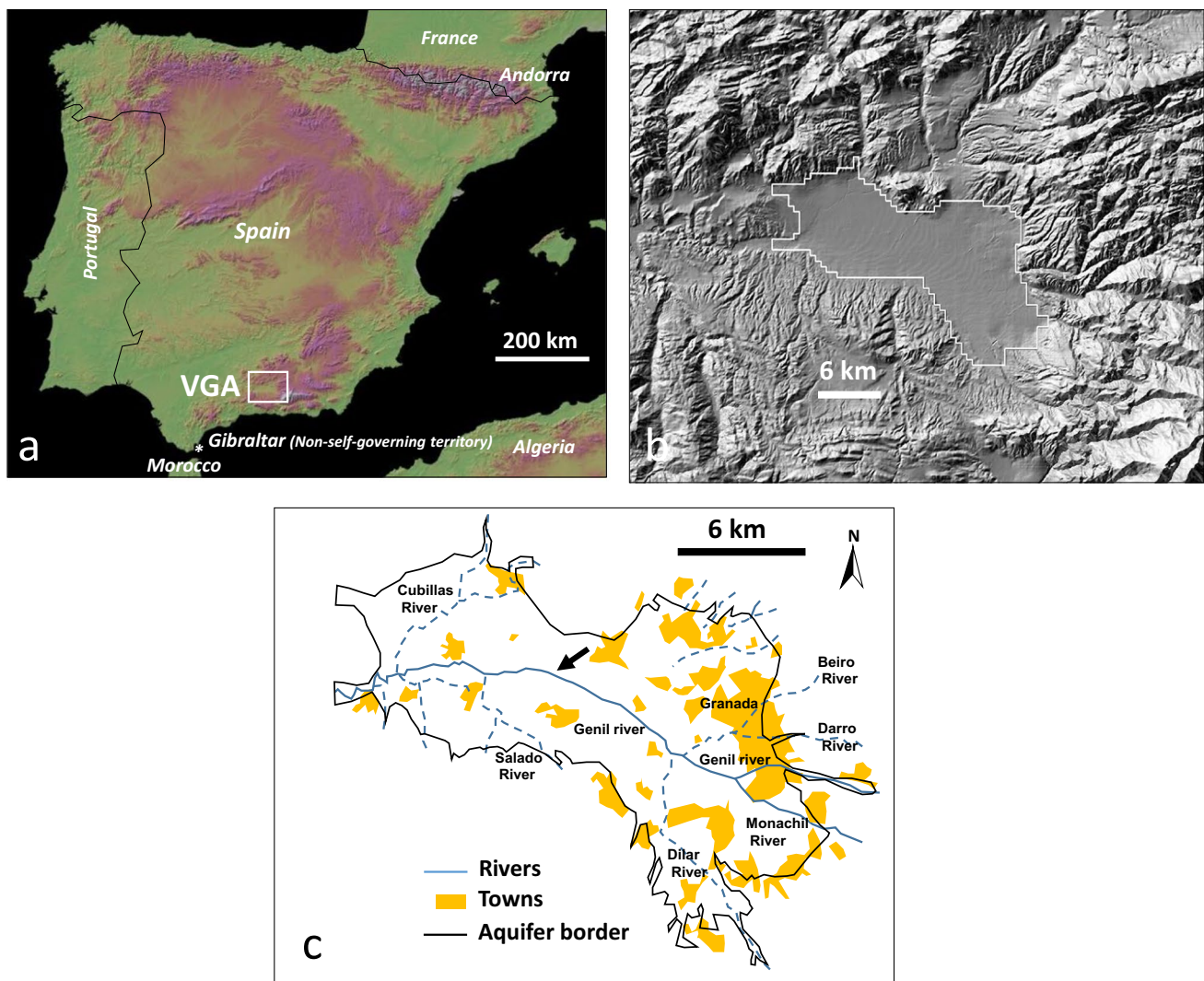
Fig. 1 a Physical map with the general geographical location of the Vega de Granada Aquifer (VGA), Spain, inside the white square located in the southern part of the Iberian Peninsula. The colour scheme in the figure depicts topographic relief, with red colours representing higher elevations. **b** Image showing terrain shadows where the Vega de Granada aquifer can be seen as an intramountain basin. The polygon defines the border of the aquifer used in the modelling and the geometry of the field used for geostatistical estimation. **c** Drainage network and towns in the Vega de Granada aquifer. The Genil River flows from east to west through the middle of the area, where the highest transmissivities are found. The black arrow indicates an area of relatively high transmissivity that is discussed in the main text.

and their natural logarithms (i.e. log-transmissivity data) are provided in Table 1. The histogram and the basic statistics are shown in Fig. 2b. The mean value of the log-transmissivity data is 6.96 ln(m^2/day) and the variance is 2.91 ln(m^2/day)^2. The median is 7.29 ln(m^2/day), which is close to the mean and the 25 percentile is 6.29 ln(m^2/day)—that is, 75% of the measurements are greater than 6.29 ln(m^2/day).

Transmissivity data from lithofacies measured along boreholes

The Spanish Geological Survey database contains 167 boreholes, in relation to the Vega de Granada aquifer, for which the lithofacies intersected by the boreholes were recorded during drilling. It should be noted that the boreholes were drilled over the years by different teams and, most of the time, the lithofacies were recorded by members of the drilling team who used their experience but often lacked a geological background. Thus, there was no standard terminology used for the stratigraphic column and the information has a high uncertainty, and for this reason, the categorization was performed by the authors of this paper. The lithofacies were interpreted and a code was assigned ranging from 0 (clay) to 6 (clean gravel) covering the full range of the different grain sizes sorted into seven categories (Table 2). Figure 3a shows a three-dimensional (3D) view of the boreholes
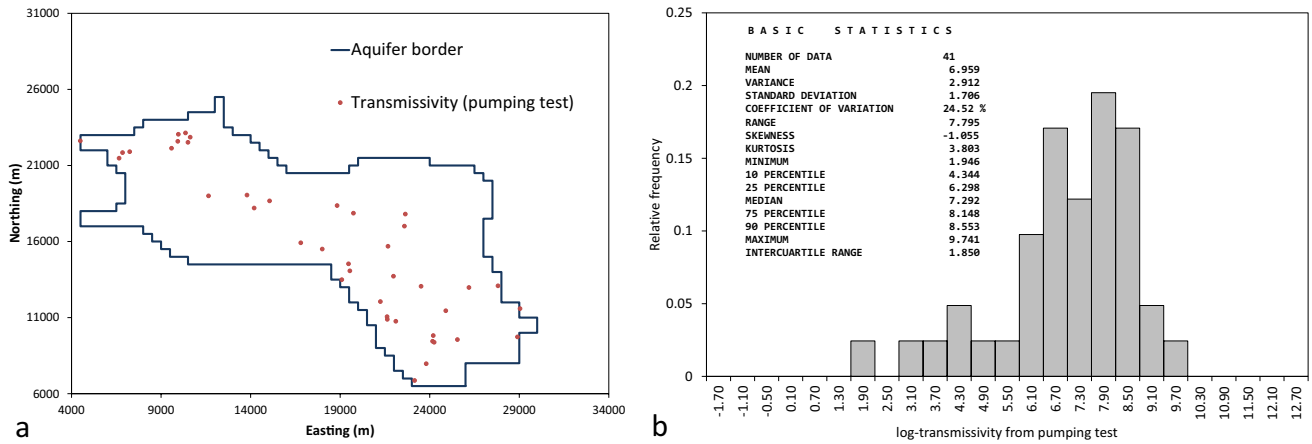
Fig. 2 a The Vega de Granada aquifer border (solid line) and the 41 locations at which log-transmissivity was measured from pumping tests. **b** Histogram and basic statistics of log-transmissivity data from pumping tests. The experimental values are provided in Table 1. The units of log-transmissivity are ln(m²/day)

and Fig. 3b shows the spatial locations of the 167 boreholes in plan within the borders of the aquifer. The log-transmissivity values in the boreholes were calculated by assigning standard values of the hydraulic conductivity, K, (Freeze and Cherry 1979) to each lithofacies and using an average aquifer saturated thickness (b) of 100 m that was determined from historical borehole and geophysics data:

$$\ln(T) = \ln(K) + \ln(b) \tag{1}$$

Next, for each borehole the mean value was calculated and 4.6 (i.e., ln100) was added to give the final resulting value for the two-dimensional (2D) locations shown in Fig. 3b. Obviously, these indirect values of transmissivity also have significant uncertainty. The experimental values of the log-transmissivity data from the lithofacies are provided in Table 3. The histogram and the basic statistics are shown in Fig. 3c. The mean value of the log-transmissivity data is 7.35 ln(m²/day) and the variance is 6.08 ln(m²/day)². The median is 7.61 ln(m²/day) and the 25-percentile is 5.58 ln(m²/day)—that is, despite similar means and medians of the log-transmissivities from the pumping test data and from the lithofacies data, there is a much larger dispersion of the log-transmissivity data from the boreholes and there is a larger proportion of relatively high values in the log-transmissivity data from the pumping test as shown, for example, by the 25% percentile. However, because of the higher dispersion of the log-transmissivity data from the boreholes, they contain more high values in absolute terms, as can be seen by comparing the 90% percentiles of both data sets.

Piezometric head measurements

Water head measurements vary in space and time. If water head measurements for a particular date can be considered

steady state, they can be related to hydraulic conductivity and transmissivity by solving the inverse problem in hydrogeology (Yeh 1986). Among the many procedures proposed for solving the inverse problem (Zimmerman et al. 1998), the geostatistical approach, explained later on, was chosen. Of the numerous water head data sets that are available for estimating the transmissivity field, the set collected in May 1970 has been chosen as it has the most (37) water head measurements (Fig. 4). The experimental values of the water head are provided in Table 4.

Co-kriging

Co-kriging is a geostatistical method for optimal multivariate spatial interpolation (Chilès and Delfiner 1999). In geostatistics, a spatial variable $Y(\mathbf{u})$ is modelled as a random variable. The set of all random variables $Y(\mathbf{u})$ in a region χ of the space, $\mathbf{u} \in \chi$, is considered a random function or random field $Y(\mathbf{u})$. With $\chi \subset \Re^d$ and $d = 2$ the problem is two dimensional as is the case in the work presented here. It is assumed that $Y(\mathbf{u})$ is second-order stationary with constant spatial mean

$$E\{Y(\mathbf{u})\} = m_Y \tag{2}$$

and the two-point statistics, the covariance and the variogram functions, depend only on the vector $\mathbf{h}$:

$$C_Y(\mathbf{u}, \mathbf{u} + \mathbf{h}) = C_Y(\mathbf{h}) = E\{Y(\mathbf{u})Y(\mathbf{u} + \mathbf{h})\} - m_Y^2 \tag{3}$$

$$\gamma_Y(\mathbf{h}) = C_Y(0) - C_Y(\mathbf{h}) \tag{4}$$

$$C_Y(0) = \sigma_Y^2 \tag{5}$$

where $m_Y, \sigma_Y^2, \gamma_Y(\mathbf{h})$ and $C_Y(\mathbf{h})$ are, respectively, the mean, variance, variogram and covariance of the random function $Y(\mathbf{u})$ and E{.} is the mathematical expectation operator.

Table 1 Experimental data (41 values) of log-transmissivity, with units ln(m^2/day), measured from pumping tests. A map of their spatial location is shown in Fig. 2

Easting (m)	Northing (m)	Log-T
4,491	22,618	6.622
6,850	21,843	6.249
9,588	22,128	5.293
9,936	22,587	8.147
9,965	23,050	7.454
10,503	22,521	7.791
10,629	22,859	8.707
11,650	19,000	9.198
13,800	19,050	8.975
14,200	18,200	8.006
16,800	15,915	5.886
21,250	12,050	7.261
21,620	11,060	6.401
21,633	10,880	8.243
22,587	17,006	8.265
22,642	17,806	3.583
23,160	6,862	1.945
23,522	13,054	6.536
23,809	7,968	7.167
24,165	9,445	8.220
24,200	9,820	9.740
24,250	9,375	6.345
24,900	11,450	6.684
25,550	9,550	8.058
28,904	9,733	7.860
29,043	11,589	4.836
27,814	13,088	7.892
18,832	18,360	7.292
19,741	17,860	6.761
17,998	15,500	6.656
19,541	14,071	8.147
15,063	18,666	8.553
21,673	15,689	8.265
19,470	14,534	7.292
6,657	21,478	4.454
7,252	21,904	2.833
10,367	23,135	4.234
21,980	13,715	8.553
26,193	12,978	8.147
19,093	13,489	6.068
22,109	10,758	6.684

In the simplest case for co-kriging, a variable of interest, or a primary variable (e.g.,log-transmissivity), is estimated from experimental values of the variable and experimental values of a secondary variable that is correlated with the primary variable. Obviously, the number of auxiliary variables can be increased.

Table 2 Lithofacies codes and lithology, and the range of hydraulic conductivity (K) values typically assigned to the lithology, plus the same information in natural logarithm form. The last column gives the value selected from the interval that was assigned to the lithofacies code. This value was selected in such a way that the final mean of log-transmissivities for this data set was similar to the mean of the log-transmissivities from pumping tests

Lithofacies code	Lithology description	K (m/day)	ln K (ln(m/day))	Selected ln K (ln(m/day))
6	Clean gravel	>1,000	>6.91	9.1
5	Clean coarse sand	1,000–10	6.91–2.3	6.2
4	Heterometric sand	10–5	2.3–1.61	1.9
3	Fine sand	5–1	1.61–0	1.3
2	Silty sand	2–0.1	0.69 to –2.3	–1.1
1	Silt	0.5–0.001	–0.69 to –6.9	–3.2
0	Clay	<0.001	<–6.9	–6.8

The co-kriging estimator of the log-transmissivity at a given geographical location $\mathbf{u}_0 = \{x_0, y_0\}$, where $\{x_0, y_0\}$ are the easting and northing coordinates respectively, can be expressed as:

$$Y^*(\mathbf{u}_0) = \sum_{i=1}^{n} \lambda_i^0 Y(\mathbf{u}_i) + \sum_{j=1}^{m} \beta_j^0 Z(\mathbf{u}_j) \qquad (6)$$

where $Y(\mathbf{u}) = \ln T(\mathbf{u})$ is the natural logarithm of transmissivity $T(\mathbf{u})$ and $Z(\mathbf{u})$ is the log-transmissivity measured from the lithofacies. Also, as an alternative that will be assessed further on, the roles of the primary variable and the secondary variable could be swapped between both types of log-transmissivity data. The secondary variable could also be $H(\mathbf{u})$, the variable that represents head data. However, the log-transmissivity data and the head data will be merged by using inverse problem co-kriging because both variables are related by the groundwater flow equations.

n and m are the number of values of the variables $Y(\mathbf{u})$ and $Z(\mathbf{u})$ respectively, used in the estimation in Eq. (6). Usually, these data are inside a neighbourhood centred on the estimation location $\mathbf{u}_0$.

The optimal weights for the linear estimation in Eq. (6) are obtained by solving the corresponding and well-known co-kriging system—see for example, Goovaerts (1997) or Wackernagel (2003). If only the primary variable is used, ordinary co-kriging collapses to ordinary kriging of the primary variable.

For co-kriging with only the two log-transmissivity data sets, the direct variograms of the two variables and the cross-variogram (or direct covariances and cross-covariance) between the two variables must be estimated from the experimental data. When the number of experimental data is small, as is the case with transmissivity
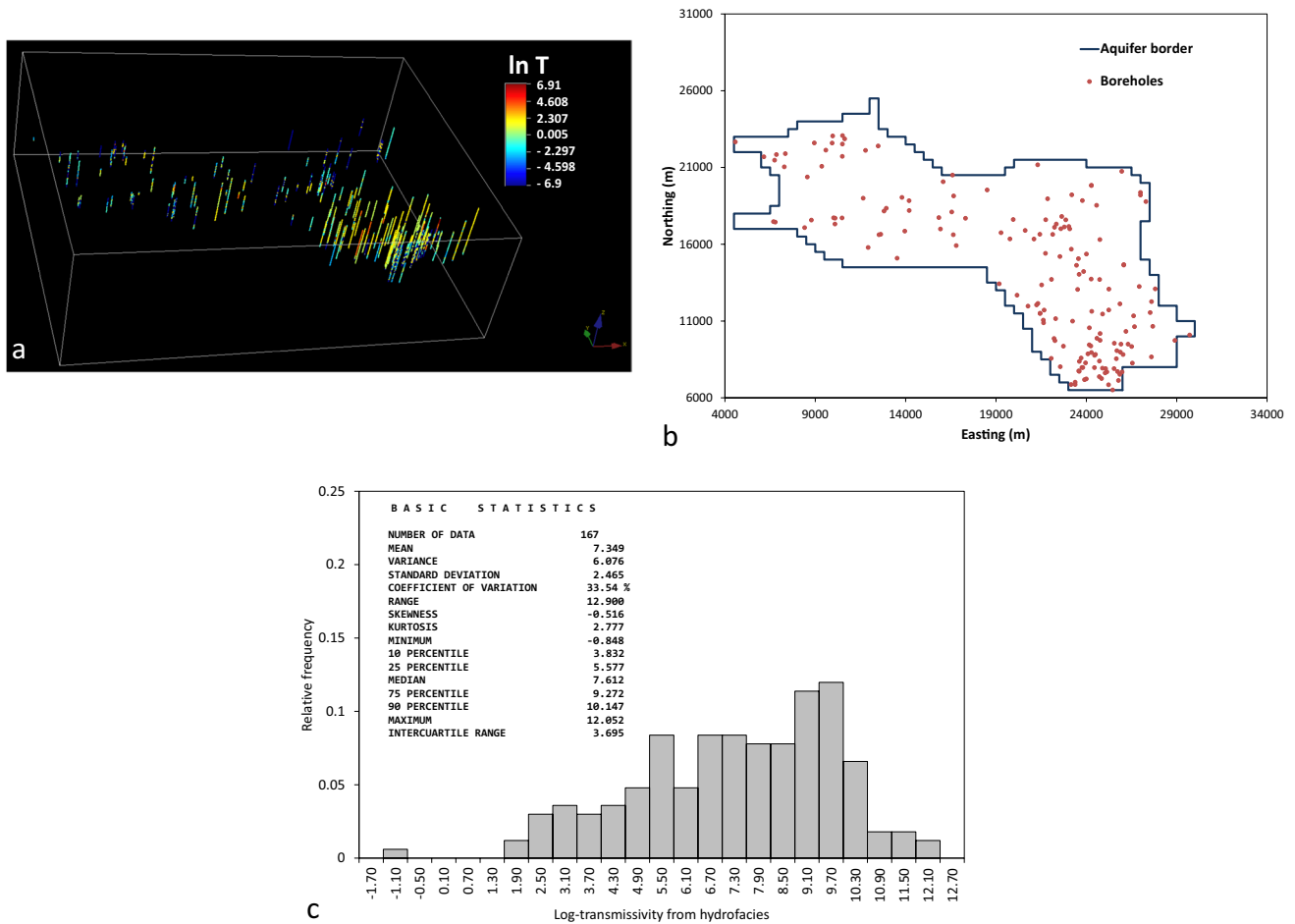
Fig. 3 **a** Three-dimensional view of the 167 boreholes for which the logarithm of transmissivity (*T*) was recorded (raw values in m²/day). **b** The Vega de Granada aquifer border (solid line) and the 167 locations at which the log-transmissivity was measured from lithofacies data from pumping tests, the estimation of the variogram in the boreholes. These points are the plan projection of the boreholes (**a**). **c** Histogram and basic statistics of log-transmissivity data from hydrofacies. The experimental values are provided in Table 2. The units of log-transmissivity are ln(m²/day)

data from pumping tests, the estimation of the variogram parameters by maximum likelihood (Kitanidis 1983, Mardia and Marshall 1984, Pardo-Igúzquiza 1997, 1998) is a good choice. A combination of maximum likelihood and visual fitting was used in this work to estimate the variogram model parameters.

The unbiasedness of the co-kriging estimator in Eq. (6) implies that the mean estimation error is zero:

$$E\{Y^*(\mathbf{u}_0) - Y(\mathbf{u}_0)\} = 0 \tag{7}$$

This is achieved by including the following conditions in the co-kriging system:

$$\sum_{i=1}^{n} \lambda_i^0 = 1 \tag{8}$$

and

$$\sum_{k=1}^{m} \beta_k^0 = 0 \tag{9}$$

The variance of the estimation error can be written as:

$$\text{Var}\{Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)\} = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i^0 \lambda_j^0 C_Y(\mathbf{h}_{ij}) + \sum_{i=1}^{m}\sum_{j=1}^{m} \beta_i^0 \beta_j^0 C_Z(\mathbf{h}_{ij})$$
$$+ \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_i^0 \beta_j^0 C_{YZ}(\mathbf{h}_{ij})$$
$$+ \sum_{j=1}^{m}\sum_{i=1}^{n} \beta_j^0 \lambda_i^0 C_{ZY}(\mathbf{h}_{ji})$$
$$- 2\sum_{i=1}^{n} \lambda_i^0 C_Y(\mathbf{h}_{i0}) - 2\sum_{j=1}^{m} \beta_j^0 C_Z(\mathbf{h}_{j0}) + C_Y(\mathbf{h}_{00}) \tag{10}$$

Co-kriging is a form of data merging (Dowd and Pardo-Igúzquiza 2006, 2012)—that is, co-kriging provides an efficient means of combining different types of data to provide a new layer of information. The co-kriging system is

Table 3 Experimental data (167 values) of log-transmissivity, T, in units $\ln(m^2/\text{day})$, calculated from the lithofacies data. A map of their spatial location is shown in Fig. 3b

Table 3 (continued)

Easting (m)	Northing (m)	Log-T
4,581	22,666	5.852
6,159	21,695	6.232
6,689	17,469	6.872
6,748	21,474	5.482
6,787	17,437	4.462
6,850	21,843	6.372
7,287	21,038	5.582
7,343	21,901	3.052
8,411	17,083	5.172
8,563	20,380	9.072
8,786	17,574	6.962
8,951	22,596	1.752
9,360	21,080	5.382
9,588	22,128	5.572
9,936	22,587	-0.847
9,965	23,050	5.312
10,019	17,718	7.732
10,089	17,316	2.562
10,100	17,700	5.602
10,487	17,713	9.122
10,496	21,720	9.722
10,503	22,521	5.932
10,508	23,076	4.572
10,629	22,859	8.392
11,650	19,000	9.282
11,781	22,110	7.712
11,927	15,791	3.312
12,480	22,400	6.932
12,525	16,618	5.822
12,624	16,648	3.862
12,809	18,157	7.342
12,933	18,341	10.452
13,530	15,090	5.412
13,800	19,050	4.552
13,957	16,854	7.412
14,190	18,839	10.512
14,200	18,200	7.532
15,841	17,733	9.872
15,930	16,992	7.372
16,078	20,072	6.152
16,550	18,100	8.282
16,599	20,500	6.402
16,642	16,617	11.952
16,661	19,143	6.542
16,800	15,915	8.542
17,316	17,690	9.832
18,513	19,531	2.952
19,183	13,424	4.922

Easting (m)	Northing (m)	Log-T
19,281	16,751	3.892
19,772	16,348	6.162
19,953	17,610	5.562
20,165	12,677	7.262
20,614	16,896	5.302
20,776	11,964	9.442
21,103	16,338	8.982
21,255	12,055	6.962
21,310	21,175	1.622
21,320	12,145	11.312
21,426	16,644	10.122
21,440	11,497	8.792
21,526	13,346	8.622
21,620	11,060	8.192
21,633	10,880	8.752
21,663	11,711	8.102
21,738	15,409	9.392
21,753	17,597	2.982
21,860	18,960	7.712
22,061	8,566	2.382
22,071	13,712	9.522
22,141	16,638	10.282
22,217	9,858	11.752
22,243	17,100	4.492
22,265	9,721	9.652
22,301	11,152	10.512
22,343	17,316	7.162
22,540	15,200	8.302
22,560	8,020	6.522
22,587	17,006	6.992
22,642	17,806	5.982
22,732	9,362	9.782
22,785	17,128	6.732
22,862	17,589	4.312
23,032	17,157	3.802
23,080	17,002	3.052
23,160	6,862	4.992
23170	15,676	7.862
23,194	19,220	3.932
23,237	10,992	10.882
23,382	6,862	5.602
23,383	7,016	5.702
23,459	14,626	7.652
23,522	13,054	10.692
23,561	15,058	9.162
23,585	7,754	9.032
23,603	14,041	9.722
23,610	7,723	9.202
23,614	8,369	9.732
23,714	8,585	8.502

Table 3 (continued)

Easting (m)	Northing (m)	Log-T
23,760	7,969	8.992
23,783	18,847	5.732
23,809	7,968	9.012
23,850	14,224	9.732
23,902	7,176	5.052
23,959	8,275	9.122
24,001	7,227	5.642
24,007	15,363	10.232
24,086	8,860	8.902
24,165	9,445	10.222
24,172	10,554	9.862
24,250	9,375	10.472
24,267	13,728	9.262
24,282	19,830	3.232
24,284	8,950	7.452
24,302	11,632	9.212
24,451	7,964	9.932
24,455	8,765	8.002
24,506	8,826	6.682
24,570	18,533	2.532
24,586	9,874	8.202
24,718	7,376	4.982
24,725	8,393	9.732
24,735	13,698	9.262
24,750	16,300	8.382
24,760	9,740	8.792
24,762	10,180	8.962
24,841	7,252	6.532
24,894	7,930	7.612
24,900	11,450	9.132
25,041	7,620	8.532
25,067	7,898	6.542
25,139	7,682	9.872
25,233	6,849	8.052
25,240	11,718	9.742
25,249	13,074	11.082
25,453	6,509	8.102
25,536	7,895	5.342
25,555	9,555	7.412
25,663	8,541	4.782
25,692	9,065	7.572
25,732	7,708	3.552
25,778	7,123	9.262
25,855	7,523	7.412
25,859	12,115	10.512
25,888	8,972	8.032
25,916	9,495	6.442
25,964	20,743	2.762
25,979	7,676	2.222
26,035	8,816	7.222

Table 3 (continued)

Easting (m)	Northing (m)	Log-T
26,073	14,671	10.102
26,073	14,640	9.882
26,193	10,325	7.222
26,311	9,492	7.762
26,508	9,337	4.742
26,550	8,258	6.572
26,619	11,339	9.332
26,664	10,630	11.372
26,927	13,249	9.562
27,000	19,375	4.652
27,000	19,200	4.432
27,300	18,780	7.072
27,534	11,550	8.332
27,600	8,660	7.392
27,612	12,258	9.962
27,676	10,655	12.052
27,814	13,088	9.642
28,904	9,733	9.792
29,720	10,087	10.172

obtained by minimizing the estimation variance in Eq. (1) subject to the unbiasedness conditions and, in matrix form (Isaaks and Srivastava 1989), is:

$$\mathbf{C\lambda} = \mathbf{B} \tag{11}$$

with:

$$
\mathbf{C} =
\begin{bmatrix}
C_Y(\mathbf{h}_{11}) & \cdots & C_Y(\mathbf{h}_{1n}) & C_{YZ}(\mathbf{h}_{11}) & \cdots & C_{YZ}(\mathbf{h}_{1m}) & 1 & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
C_Y(\mathbf{h}_{n1}) & \cdots & C_Y(\mathbf{h}_{nn}) & C_{YZ}(\mathbf{h}_{n1}) & \cdots & C_{YZ}(\mathbf{h}_{nm}) & 1 & 0 \\
C_{ZY}(\mathbf{h}_{11}) & \cdots & C_{ZY}(\mathbf{h}_{1n}) & C_Z(\mathbf{h}_{11}) & \cdots & C_Z(\mathbf{h}_{1m}) & 0 & 1 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
C_{ZY}(\mathbf{h}_{m1}) & \cdots & C_{ZY}(\mathbf{h}_{mn}) & C_Z(\mathbf{h}_{m1}) & \cdots & C_Z(\mathbf{h}_{mm}) & 0 & 1 \\
1 & \cdots & 1 & 0 & \cdots & 0 & 0 & 0 \\
0 & \cdots & 0 & 1 & \cdots & 1 & 0 & 0
\end{bmatrix}
\tag{12}
$$

$$
\mathbf{\lambda} =
\begin{bmatrix}
\lambda_1^0 \\
\vdots \\
\lambda_n^0 \\
\beta_1^0 \\
\vdots \\
\beta_m^0 \\
\mu_1 \\
\mu_2
\end{bmatrix},
\mathbf{B} =
\begin{bmatrix}
C_Y(\mathbf{h}_{10}) \\
\vdots \\
C_Y(\mathbf{h}_{n0}) \\
C_{YZ}(\mathbf{h}_{01}) \\
\vdots \\
C_{YZ}(\mathbf{h}_{0m}) \\
1 \\
0
\end{bmatrix}
\tag{13}
$$

where μ_1 and μ_2 are Lagrange multipliers or parameters that are used to include the constraints given in Eqs. (8) and (9).

The solution of the co-kriging system:

$$\lambda = \mathbf{C}^{-1}\mathbf{B} \tag{14}$$

provides the weights required in the estimator in Eq. (6).

In the geostatistical literature, the form of co-kriging summarised previously is known as "ordinary co-kriging" to distinguish it from "simple co-kriging" in which the mean m_Y in Eq. (2) is known (Journel and Huijbregts 1978; Goovaerts 1997; Chilès and Delfiner 1999).

Inverse problem co-kriging

In an aquifer, the head and log-transmissivity data are linked by the physical equations that describe the groundwater flow, the type of boundary conditions at the border of the aquifer and by other inputs (e.g., recharge by rainfall) and outputs (e.g., extraction of water by pumping wells or by evapotranspiration). For the case of an unconfined aquifer, the groundwater flow model with Dupuit assumptions gives the Boussinesq equation that can be written as (Cordano and Rigon 2013):

$$\frac{\partial}{\partial x}\left(K(\mathbf{u})H(\mathbf{u})\frac{\partial H(\mathbf{u})}{\partial x} \right) + \frac{\partial}{\partial y}\left(K(\mathbf{u})H(\mathbf{u})\frac{\partial H(\mathbf{u})}{\partial y} \right) = S\frac{\partial H(\mathbf{u})}{\partial t} - R(\mathbf{u}, t) \tag{15}$$

where $K(\mathbf{u})$ and $H(\mathbf{u})$ are the hydraulic conductivity and the water head respectively, at the spatial location $\mathbf{u} = \{x, y\}$. $R(\mathbf{u}, t)$ represents the recharge or extraction term of water from the aquifer. S and t are, respectively, the specific yield (drainable porosity) and time.

For the steady-state case:

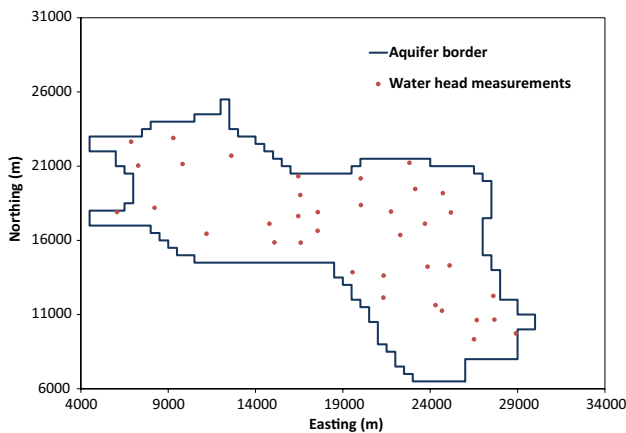$$\frac{\partial H(\mathbf{u})}{\partial t} = 0 \tag{16}$$

Fig. 4 The Vega de Granada aquifer border (solid line) and the locations of the 37 water head measurements surveyed in May 1970. The experimental values are provided in Table 3

and, in the absence of recharge and extraction (pumping wells or evapo-transpiration, for example):

$$R(\mathbf{u}, t) = 0 \tag{17}$$

and

$$\frac{\partial}{\partial x}\left(K(\mathbf{u})H(\mathbf{u})\frac{\partial H(\mathbf{u})}{\partial x} \right) + \frac{\partial}{\partial y}\left(K(\mathbf{u})H(\mathbf{u})\frac{\partial H(\mathbf{u})}{\partial y} \right) = 0 \tag{18}$$

and considering the following relationships:

$$T(\mathbf{u}) = K(\mathbf{u})b \tag{19}$$

$$\frac{\partial \ln T(\mathbf{u})}{\partial x} = \frac{1}{T(\mathbf{u})}\frac{\partial T(\mathbf{u})}{\partial x} \tag{20}$$

$$\frac{\partial (H(\mathbf{u}))^2}{\partial x} = 2H(\mathbf{u})\frac{\partial H(\mathbf{u})}{\partial x} \tag{21}$$

and

$$Y(\mathbf{u}) = \ln(T(\mathbf{u})) \tag{22}$$

After operating with the derivative of a product of functions and some minor calculus operations, the resulting partial differential equation is:

$$\frac{\partial Y(\mathbf{u})}{\partial x}\frac{\partial (H(\mathbf{u}))^2}{\partial x} + \frac{\partial Y(\mathbf{u})}{\partial y}\frac{\partial (H(\mathbf{u}))^2}{\partial y} + \frac{\partial^2 (H(\mathbf{u}))^2}{\partial x^2} + \frac{\partial^2 (H(\mathbf{u}))^2}{\partial y^2} = 0 \tag{23}$$

where $T(\mathbf{u})$, and $Y(\mathbf{u})$ are transmissivity and log-transmissivity respectively, at the spatial location $\mathbf{u} = \{x, y\}$. b is the mean saturated thickness of the aquifer, which is used because it is common practice to consider time-invariant "transmissivity" values defined by the hydraulic conductivity multiplied by a time-invariant representative thickness (Pulido-Velazquez et al. 2007).

As there will be given boundary conditions at the border of the aquifer, it is assumed that any other source of recharge or extraction in the aquifer is zero. Given the boundary of the aquifer Ω and the unitary vector $\mathbf{n}$ normal to the boundary, the no-flow boundary condition implies:

$$\left.\frac{\partial H(\mathbf{u})}{\partial \mathbf{n}}\right|_{\Omega} = 0 \tag{24}$$

In the same way, in the parts of the boundary Ω where the boundary condition is a fixed head, the boundary condition is:

$$H(\mathbf{u}_i) = H_i = \text{constant}, \qquad \mathbf{u}_i \in \Omega \tag{25}$$

Obviously, this steady-state condition is only realistic if it is considered for a particular period of time (e.g., a day) as groundwater flow and other recharge processes are slow and it is assumed that there are no pumping wells running.

Table 4 Experimental water head data (37 values). A map of their spatial locations is shown in Fig. 4

Easting (m)	Northing (m)	Water head (m)
6,076	17,905	523.33
8,223	18,195	531.88
9,826	21,140	540.64
9,299	22,897	544.64
7,287	21,038	548.07
12,615	21,703	549.16
6,882	22,645	552.84
11,192	16,445	559.11
16,449	20,317	565.34
14,798	17,124	567.69
15,083	15,859	570.01
16,562	19,052	570.64
16,453	17,635	572.31
17,564	17,904	575.88
16,588	15,847	578.93
17,555	16,641	581.41
22,287	16,360	595.49
20,033	18,379	596.61
21,331	13,624	598.21
23,850	14,224	600.14
19,556	13,853	602.08
25,109	14,307	602.31
21,320	12,145	602.94
24,302	11,632	602.99
24,670	11,260	603.20
27,612	12,258	605.34
27,676	10,655	606.29
26,664	10,630	609.72
20,021	20,167	610.45
21,755	17,936	610.80
23,698	17,121	613.14
28,904	9,733	614.33
26,508	9,337	629.34
23,146	19,467	645.20
22,813	21,226	660.70
25,182	17,882	671.52
24,721	19,180	676.13

Thus, the steady-state conditions can be considered realistic for the day on which the water head data were measured but a transient model should be used for longer periods of time (e.g., a week, a month, or more).

Ordinary co-kriging could be applied to the log-transmissivity and water head and ignoring Eq. (23) to estimate the covariance of log-transmissivity, the covariance of the water head and the cross-covariance between log-transmissivity and water head. However, there are several inference problems with this approach.

Firstly, the hypothesis of a constant mean water head, such as that in Eq. (2), is not realistic for an aquifer water head. Water head generally shows a spatial trend. That trend, or drift, is the difference in potential energy (for an unconfined aquifer) that results in the groundwater flowing from areas of higher potential to areas of lower potential. Thus, a spatially variable mean must be considered:

$$E\{H(\mathbf{u})\} = m_H(\mathbf{u}) \tag{26}$$

The trend is usually modelled as a low order polynomial:

$$m_H(\mathbf{u}) = \sum_{i=0}^{p} b_i f_i(\mathbf{u}) \tag{27}$$

where $\{b_i; i = 0, …, p\}$ is a set of coefficients that must be estimated and

$\{f_i(\mathbf{u}); i = 0, …, p\}$ is a set of known base functions, usually monomial, of the co-ordinates—for example, $\{1, x, y, xy, x^2, y^2\}$ are the six base functions for the case of a quadratic polynomial trend on a plane.

A trend requires new unbiasedness conditions for the co-kriging estimator—for example, for a linear trend in the secondary variable only, the co-kriging system must take into account the following unbiased conditions:

$$\sum_{i=1}^{n} \lambda_i^0 = 1 \tag{28}$$

$$\sum_{j=1}^{m} \beta_j^0 = 0 \tag{29}$$

$$\sum_{i=1}^{n} \beta_i^0 x_i = 0 \tag{30}$$

$$\sum_{i=1}^{n} \beta_i^0 y_i = 0 \tag{31}$$

Firstly, the ordinary co-kriging system that takes these unbiasedness conditions into account is universal co-kriging (Journel and Huijbregts 1978; Goovaerts 1997; Chilès and Delfiner 1999). Secondly, it can be shown (Hoeksema and Kitanidis 1984; Dagan 1985; Ahmed and de Marsily 1993; Kitanidis 1997, among others) that the groundwater flow equation that physically links the log-transmissivity and water head implies that their cross-covariance is anisotropic and antisymmetric. As it is difficult to provide a good estimate of this convoluted cross-covariance from small sets of experimental log-transmissivity and water head data, a theoretical derivation of the cross-covariance may be preferable. This problem has received significant attention in the scientific literature. Basically, there are two approaches—use various numerical methods to provide an analytical solution (Rubin 2003) or use a Monte Carlo method (Kitanidis 1997). In this sense, Dagan (1985) derived

an analytical expression that considered an aquifer of infinite extent and a water head with a linear trend along one of its principal axes together with an exponential covariance for the log-transmissivity and found that the cross-covariance between log-transmissivity and water head, $C_{YH}(\mathbf{h})$, is given by:

$$C_{YH}(\mathbf{h}) = \sigma_Y^2 \beta_1 \ell \frac{h_x}{h^2} \left[1 - (1 + h) \exp(-h) \right] \tag{32}$$

where σ_Y^2 : is the variance, or total sill, of the log-transmissivity $Y(\mathbf{u})$ and ℓ is the range of the exponential covariance of log-transmissivity $Y(\mathbf{u})$. The covariance of the log-transmissivity is:

$$C_Y(\mathbf{h}) = C_0 + C_1 \exp(-h/\ell) \tag{33}$$

$$\sigma_Y^2 = C_0 + C_1 \tag{34}$$

where:

C_0 is the nugget variance;

C_1 is the partial sill variance;

$\mathbf{h} = (h_x, h_y) = \left(\frac{x_1 - x_2}{\ell}, \frac{y_1 - y_2}{\ell} \right)$ is the distance vector standardized by the range of log-transmissivity;

$h = \sqrt{h_x^2 + h_y^2}$ is the magnitude of the standardized distance vector and $m_H(\mathbf{u}) = m(x, y) = \beta_0 + \beta_1 x$ is the linear trend of the water head, which, in this case, is an east–west trend.

The universal co-kriging estimator can then be written as:

$$Y^*(\mathbf{u}_0) = \sum_{i=1}^n \lambda_i^0 Y(\mathbf{u}_i) + \sum_{j=1}^m \beta_j^0 H(\mathbf{u}_j) \tag{35}$$

for which the optimal weights are obtained by solving the universal co-kriging system (in matrix form) such as given in Eq. (11) but with the following matrix elements:

$$\mathbf{C} = \begin{bmatrix} C_Y(\mathbf{h}_{11}) & \cdots & C_Y(\mathbf{h}_{1n}) & C_{YH}(\mathbf{h}_{11}) & \cdots & C_{YH}(\mathbf{h}_{1m}) & 1 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_Y(\mathbf{h}_{n1}) & \cdots & C_Y(\mathbf{h}_{nn}) & C_{YH}(\mathbf{h}_{n1}) & \cdots & C_{YH}(\mathbf{h}_{nm}) & 1 & 0 & 0 & 0 \\ C_{HY}(\mathbf{h}_{11}) & \cdots & C_{HY}(\mathbf{h}_{1n}) & C_H(\mathbf{h}_{11}) & \cdots & C_H(\mathbf{h}_{11}) & 0 & 1 & x_1 & y_1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{HY}(\mathbf{h}_{m1}) & \cdots & C_{HY}(\mathbf{h}_{mn}) & C_H(\mathbf{h}_{m1}) & \cdots & C_H(\mathbf{h}_{mm}) & 0 & 1 & x_m & y_m \\ 1 & \cdots & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & x_1 & \cdots & x_m & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & y_1 & \cdots & y_m & 0 & 0 & 0 & 0 \end{bmatrix} \tag{36}$$

$$\lambda = \begin{bmatrix} \lambda_1^0 \\ \vdots \\ \lambda_n^0 \\ \beta_1^0 \\ \vdots \\ \beta_m^0 \\ \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} C_Y(\mathbf{h}_{10}) \\ \vdots \\ C_Y(\mathbf{h}_{n0}) \\ C_{YH}(\mathbf{h}_{01}) \\ \vdots \\ C_{YH}(\mathbf{h}_{0m}) \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{37}$$

where μ_1, μ_2 μ_3 and μ_4 are Lagrange multipliers or parameters that are included in the minimization of the estimation variance subject to the constraints given in Eqs. (12) and (13). The solution of the co-kriging system is given by Eq. (14). In the geostatistical literature the co-kriging system in Eqs. (12) and (13) is known as universal co-kriging. It should be noted that in Eqs. (12) and (13) the trend is considered only in the secondary variable (i.e., water head) as there is no trend in the primary variable (i.e., log-transmissivity). Furthermore, the universal co-kriging system in which the cross-covariances between log-transmissivity and water head take into account the groundwater flow equation (i.e., the physical link between log-transmissivity and water head) is the geostatistical solution to the inverse problem in hydrogeology (Kitanidis 1996; Zimmerman et al. 1998; Rubin 2003) and hence the descriptive name of "inverse problem universal co-kriging" given to the estimator.

Results

The first result is the log-transmissivity obtained by ordinary kriging using the log-transmissivity data from the pumping tests. The estimated variogram and the fitted model for the log-transmissivity data from the pumping tests are shown in Fig. 5a. The estimated log-transmissivity field is shown in Fig. 6a. The model fitted to the estimated covariance is:

$$C_Y(h) = 1.5 + 1.5 \exp(-h/3373) \tag{38}$$

i.e., an isotropic exponential covariance with a nugget variance of 1.5, a partial sill of 1.5 and a range of 3,373 m. A study of the uncertainty of the covariance parameters for log-transmissivity can be found in Pardo-Igúzquiza et al. (2009).

The second result is the log-transmissivity obtained by ordinary kriging of the log-transmissivity data from the hydrofacies data. The estimated variogram and the model fitted for the log-transmissivity data from pumping tests are shown in Fig. 5b; the estimated log-transmissivity field is shown in Fig. 6b. The model fitted to the estimated covariance is:

$$C_Z(h) = 2.14 + 4.98 \exp(-h/3373) \tag{39}$$

i.e., an isotropic exponential covariance with a nugget variance 2.14, a partial sill of 4.98 and a range of 3,373 m. Both covariances were modelled with the same range because the same range must be used in the linear co-regionalization model to generate the co-kriged estimates.

Another result of estimated log-transmissivity maps is provided by ordinary co-kriging with one unbiasedness condition. In this case the estimated cross-variogram and the fitted model are shown in Fig. 5c and the fitted model is
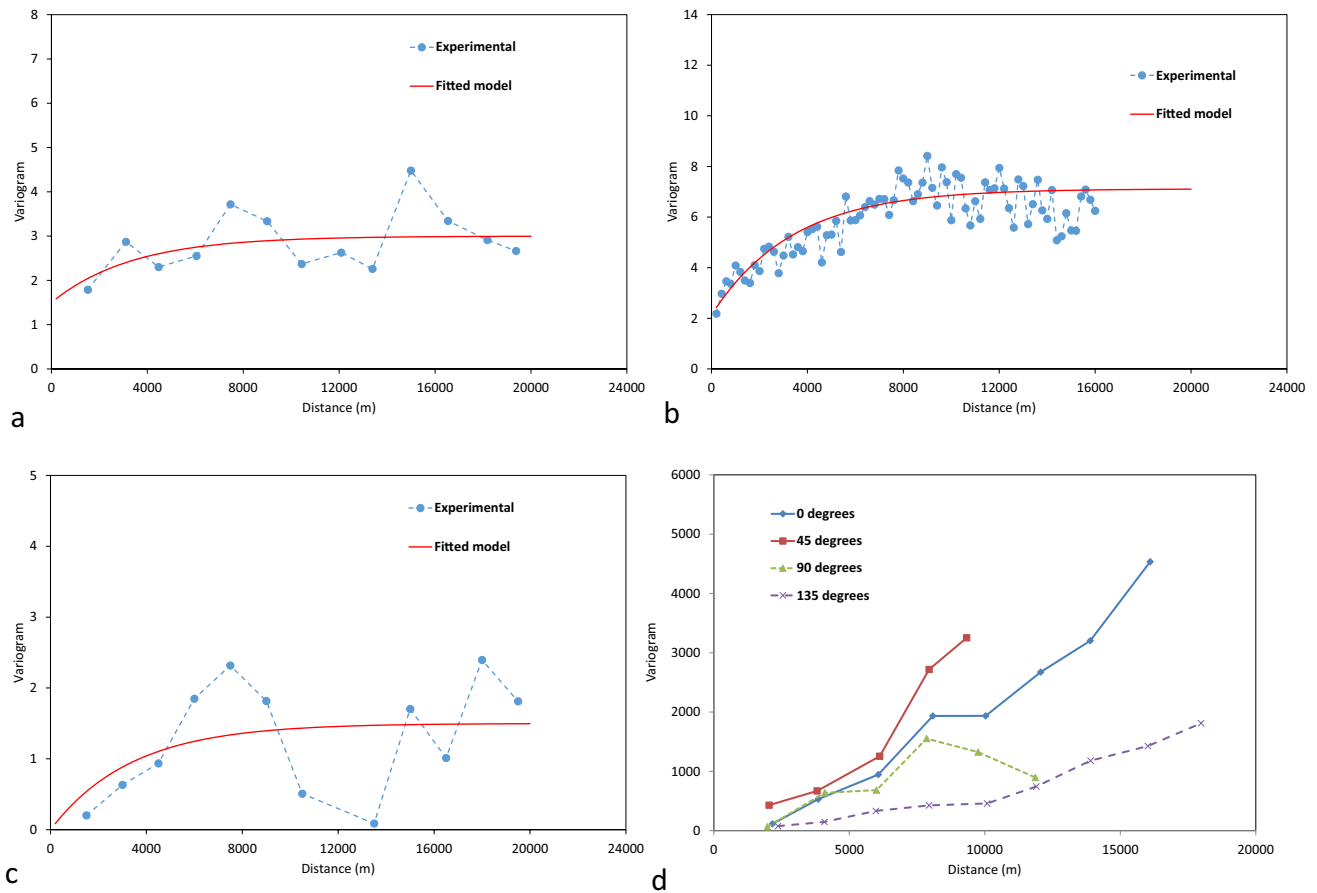
Fig. 5 a Omnidirectional experimental variogram (solid blue dots) and fitted model (solid red line) for log-transmissivity estimated from pumping tests. **b** Omnidirectional experimental variogram (solid blue dots) and fitted model (solid red line) for the log-transmissivity estimated from borehole lithofacies. **c** Omni-directional experimental cross-variogram (solid blue dots) and fitted model (solid red line) between log-transmissivity estimated from pumping tests and log-transmissivity estimated from borehole lithofacies. **d** Experimental variogram of water head for the four main geographical directions

$$C_Y(h) = 1.5 \exp(-h/3373) \qquad (40)$$

The cross-covariance between the two types of log-transmissivity data was difficult to estimate from the data as the experimental sampling plan was heterotopic (Wackernagel 2003) with few sites at which both the primary and secondary variables were sampled.

The linear model of co-regionalization was used to model the direct and cross covariances. The estimated maps are provided in Fig. 7a,b for ordinary co-kriging using both types of log-transmissivity data and considering the pumping test data and the lithofacies data as primary variables respectively.

A fourth result is the log-transmissivity estimated by universal co-kriging using the log-transmissivity data from the pumping test and the head data. The experimental variograms of the water head for the four main geographical directions is shown in Fig. 5d. The water head clearly has

a trend (Fig. 8) that provides the difference in potential energy (in an unconfined aquifer) that drives the water flow from the northern and eastern parts of the aquifer towards the south-east where the water exits the aquifer. As there is a small number (37) of water head data together with a trend and the parameters of the variogram of the residual must be estimated, the maximum likelihood estimation method was used to infer these parameters (Pardo-Igúzquiza 1997). The estimated water head trend is:

$$m_H(x,y) = 423.3874 + 0.006125x + 00.3368y \qquad (41)$$

where

$$E\{H(\mathbf{u})\} = m_H(\mathbf{u}) = m_H(x,y) \qquad (42)$$

The estimated isotropic variogram of the residual is:

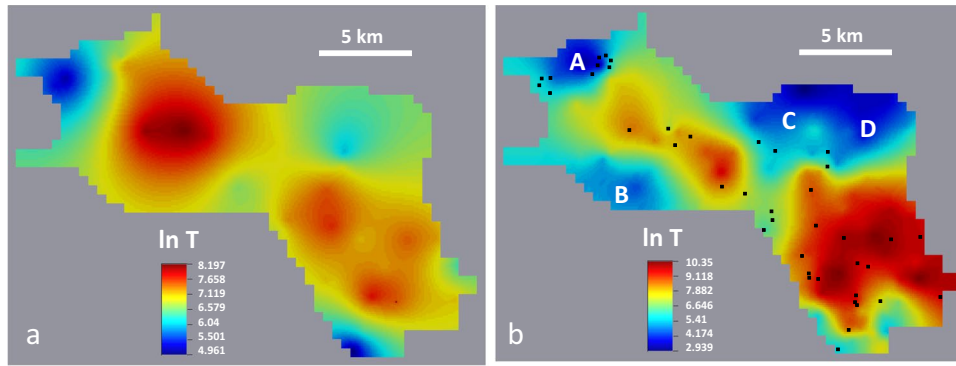$$\gamma_R(h) = 241.4\mathrm{Sph}(h;7147) \qquad (43)$$

Fig. 6 Log-transmissivity field estimated by ordinary kriging. **a** Using log-transmissivity data from pumping tests (OK-1 in Table 5). **b** Using log-transmissivity data from borehole lithofacies (OK-2 in Table 5). The black squares represent the 41 locations at which log-transmissivity was measured from pumping tests. The biased spatial locations of the pumping tests are clear because the low transmissivity regions (*A*, *B*, *C* and *D*) have no pumping tests. The units of log-transmissivity are ln(m²/day)

i.e., a spherical model with no nugget effect, a sill of 241.4 and a range of 7,147 m. The water head field, estimated by using universal kriging with a linear trend, is shown in Fig. 8.

The cross-covariance between log-transmissivity and water head is required for the application of inverse model universal co-kriging. The estimated cross-covariance between log-transmissivity data from the lithofacies and the water head data is shown in Fig. 9a. This cross-covariance is anisotropic, antisymmetric and shows a high statistical variability because of the small amount of data in the data sets, especially in the water head data set. Consequently, instead of trying to fit a convoluted model to the experimental variogram, Dagan's (1985) analytical solution was used. Figure 9b shows the analytical cross-covariance provided by Dagan's solution which can be compared with the experimental cross-covariance in Fig. 9a. A comparison of Figs. 9a and 9b shows that the neat cross-covariance model in Fig. 9b can be intuited in the experimental cross-covariance in Fig. 9a. The cross-covariance is anisotropic, with higher variability in the flow direction and vanishing in a direction perpendicular to the water flow. In addition, the cross-covariance is antisymmetric with:

$$C_{YH}(h) = -C_{YH}(-h) \tag{44}$$

However, the co-kriging covariance matrix is still symmetric because for a given single entry in this matrix in relation to the cross-covariance:

$$C_{YH}\left(\mathbf{u}_i - \mathbf{u}_j\right) = C_{HY}\left(\mathbf{u}_j - \mathbf{u}_i\right) \tag{45}$$

The cross-covariance between log-transmissivity from the pumping tests and the water head is the imposed cross-covariance given in Eq. (7) with 3.0, 0.0061, and 3,373 m for σ_Y^2, β_1 and ℓ, respectively. The same is valid for the cross-covariance between log-transmissivity data from lithofacies

and water head but taking into account that σ_Y^2 is equal to 7.12. Four cases of the inverse problem by universal co-kriging have been considered:

- Inverse problem universal co-kriging using log-transmissivity data from pumping tests and water head data. The estimated map of log-transmissivity is shown in Fig. 10a.
- Inverse problem universal co-kriging using log-transmissivity data from lithofacies and water head data. The estimated map of log-transmissivity is shown in Fig. 10b.
- Inverse problem universal co-kriging using log-transmissivity data from the pumping test, log-transmissivity data from lithofacies and water head data. The log-transmissivity data from the pumping test is the primary variable, the other two are secondary variables. The estimated map of log-transmissivity is shown in Fig. 11a.
- Inverse problem universal co-kriging using log-transmissivity data from pumping tests, log-transmissivity data from lithofacies and water head data. The log-transmissivity data from lithofacies is the primary variable, the other two are secondary variables. The estimated map of log-transmissivity is shown in Fig. 11b.

The implications of these results are discussed in the following section.

Discussion

Although the different geostatistical estimation methods provide different transmissivity fields (Figs. 6, 7, 10 and 11), they are significantly coherent. The log-transmissivity field obtained by ordinary kriging of the log-transmissivity data (Fig. 6a) from the pumping test provides a transmissivity
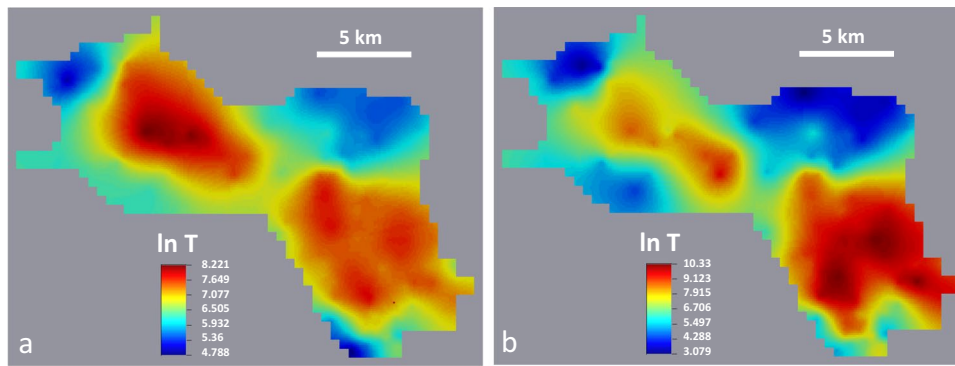
Fig. 7 Log-transmissivity field estimated by ordinary co-kriging with one unbiasedness condition and using log-transmissivity data from the pumping test and the borehole log-transmissivity lithofacies data. **a** Using log-transmissivity data from pumping tests as the primary variable (OCOK-1-2 in Table 5). **b** Using log-transmissivity data from borehole lithofacies as the primary variable (OCOK-2-1 in Table 5). The units of log-transmissivity are $\ln(m^2/day)$

map with the coarsest detail and it is based on biased data in the sense of preferential sampling in areas of high transmissivity as may be seen in Fig. 6b, where low transmissivity zones A, B, C and D, in the figure, have no pumping test. The transmissivity field estimated by ordinary kriging of the log-transmissivity data from the hydrofacies data generates a map (Fig. 7b) that agrees with the general trend provided by the previous transmissivity field. It also incorporates more detail because there are more experimental data that were sampled more homogenously over the whole area rather than only in high transmissivity zones.

An improved map could be generated by ordinary co-kriging using the additional indirect transmissivity data from the hydrofacies or by using inverse problem co-kriging and head data. The estimated log-transmissivity maps for ordinary co-kriging are provided in Fig. 7a,b for the cases that

consider as primary variables the experimental data from pumping tests or from lithofacies. A comparison of the ordinary kriging maps and the ordinary co-kriging maps shows that they are very similar because the cross-correlation between the two sets of log-transmissivity data is low and thus the information provided by the secondary information is low. This is especially clear by comparing Figs. 6b and 7b, which are almost identical.

With respect to the inverse problem, universal co-kriging of the four previous possibilities resulted in the log-transmissivity maps in Figs. 10 and 11. Figure 10 is a log-transmissivity map estimated by inverse problem universal co-kriging using two variables. Figure 10a uses log-transmissivity data from pumping tests and water head data while Fig. 10b uses log-transmissivity data from lithofacies and water head data. Figure 11 is the log-transmissivity map
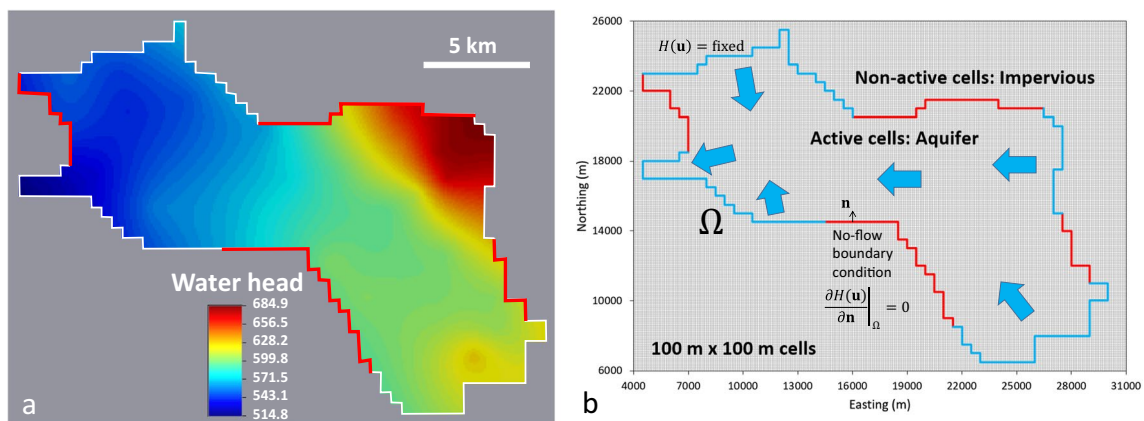

Fig. 8 **a** Water head map (in metres) estimated by universal kriging with a linear drift and the variogram of the residuals estimated by maximum likelihood. **b** Numerical model of the aquifer using a discretization of 100 m × 100 m cells. The cells of the aquifer are active cells while the cells outside of the aquifer are impervious cells

that remain inactive. The coloured line (blue and red) represents the boundary Ω. The different colours represent the different boundary conditions: the blue line represents the border with a fixed head boundary condition and the red line represents the no-flow boundary condition
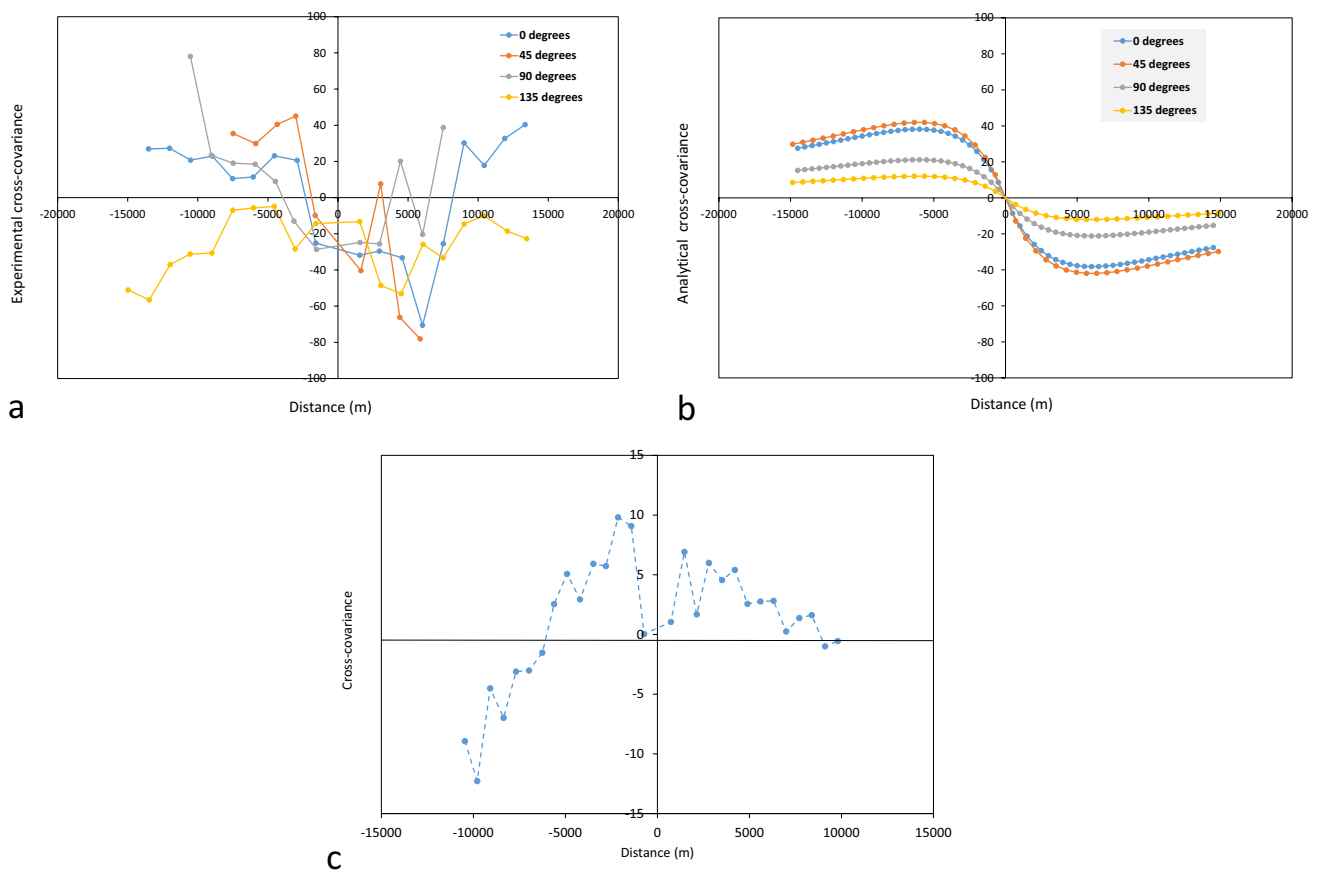
a

b

c

Fig. 9 a Experimental cross-covariances between log-transmissivity and water head. The different colours represent different spatial directions. Experimental variogram (solid blue dots) and fitted model (solid red line) for log-transmissivity estimated from pumping tests. **b** Theoretical cross-covariances between log-transmissivity and water head. Experimental variogram (solid blue dots) and fitted model (solid red line) for the log-transmissivity estimated from borehole

lithofacies. **c** Experimental cross-covariance between log-transmissivity data from pumping tests and log-transmissivity data from lithofacies. Note the low value of the cross-covariance close to the origin indicating that the maximum correlation is reached for a distance for which one data set should be shifted with respect to the other. This explains the low correlation between the two data sets indicated by the variogram in Fig. 5c

estimated by inverse problem universal co-kriging using three variables: the two log-transmissivity data sets and water head data. Figure 11a uses log-transmissivity data from pumping tests as the primary variable and in Fig. 11b the primary variable used was log-transmissivity data from the lithofacies.

The different methods could also be compared by their estimation variance for the different methodologies; however, these results depend significantly on the parameters estimated for the covariances of the different variables and a more independent comparison would be better. For this, the transmissivity fields can be used as input to the direct problem of using the groundwater flow model to estimate the May 1970 head field. The direct model solves the 2D steady-state saturated groundwater flow model in an unconfined aquifer and can be written in terms of log-transmissivities as given in Eq. (22) with the boundary conditions of no-flow and prescribed head given in Fig. 8.

The area in Fig. 8 was discretized into 265×200 cells of $100\,\text{m} \times 100\,\text{m}$ and a finite difference scheme was used to solve Eq. (22) taking into account two types of boundary conditions—no-flow and prescribed head, of which the no-flow boundary condition implies no-flow in the perpendicular direction.

The boundary conditions for the steady-state model were chosen as the specified heads as provided by the universal kriging map of water head shown in Fig. 8 and the no-flow boundaries in the borders are those determined from previous studies—Pardo-Igúzquiza et al. (2009); Pardo-Igúzquiza and Chica-Olmo 2004, 2007 and Kuhlman and Pardo-Igúzquiza (2010).

Taking each estimated log-transmissivity field as an input it is possible to estimate the water head data by solving Eq. (17) using the boundary conditions of Fig. 8. From the modelled water head field it is possible to calculate the corresponding water head values at the experimental locations where water
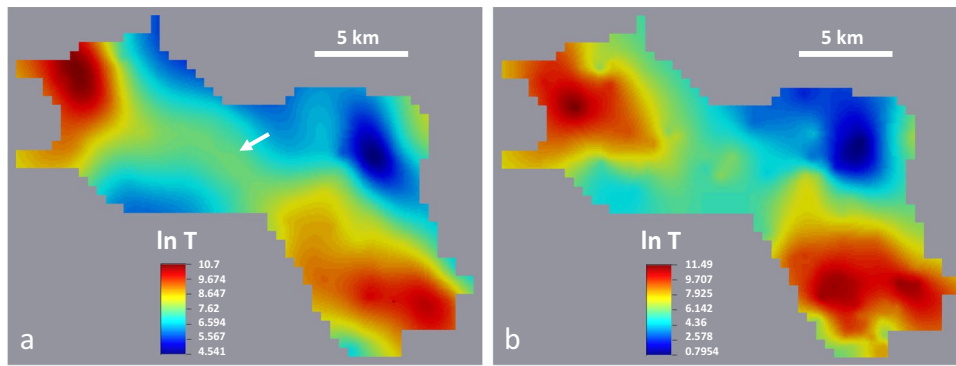
Fig. 10 Log-transmissivity (*T*) field estimated by inverse problem ordinary universal co-kriging with two variables: **a** Using log-transmissivity data from pumping tests and water head data (IP-UCOK-1 in Table 5). The white arrow shows a channel of relatively high trans-missivity that is discussed in the main text. **b** Using log-transmissivity data from borehole lithofacies and water head data (IP-UCOK-1 in Table 5). The units of log-transmissivity are ln(m²/day)

head is known and different validation statistics can be calculated. The validation statistics considered in this work are mean error (ME), mean squared error (MSE), mean normalized squared error (MNSE), the mean absolute error (MAE), the Nash and Sutcliffe (1970) efficiency (NSE) and the Kling-Gupta Efficiency criteria (KGE) (Knoben et al. 2019):

$$\text{ME} = \frac{1}{N} \sum_{i=1}^{N} \left(h_i^{\text{o}} - h_i^{\text{m}} \right) \tag{46}$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left(h_i^{\text{o}} - h_i^{\text{m}} \right)^2 \tag{47}$$

$$\text{MNSE} = \frac{1}{N} \sum_{i=1}^{N} \frac{\left(h_i^{\text{o}} - h_i^{\text{m}} \right)^2}{h_i^{\text{o}}} \tag{48}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| h_i^{\text{o}} - h_i^{\text{m}} \right| \tag{49}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N} \left(h_i^{\text{o}} - h_i^{\text{m}} \right)^2}{\sum_{i=1}^{N} \left(h_i^{\text{o}} - \overline{h_i^{\text{o}}} \right)^2} \tag{50}$$

$$\text{KEG} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{\text{m}}}{\sigma_{\text{o}}} - 1 \right)^2 + \left(\frac{\mu_{\text{m}}}{\mu_{\text{o}}} - 1 \right)^2} \tag{51}$$

h_i^{o} : Head datum observed at the *i*th location

h_i^{m} : Head datum modelled by the groundwater flow Eq. (17) at the *i*th location

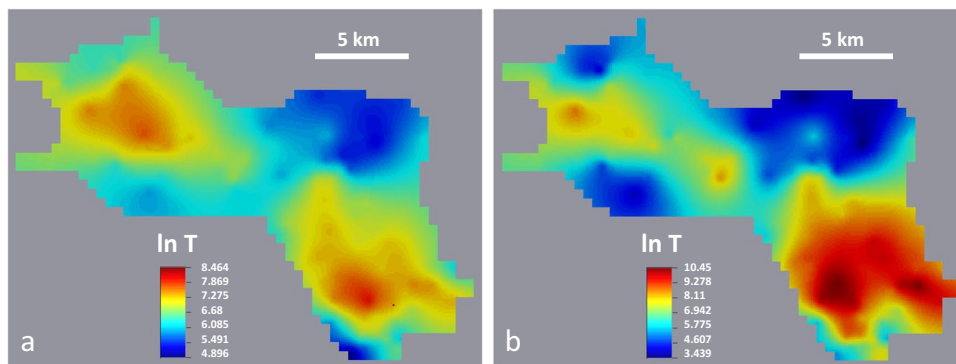$\overline{h_i^{\text{o}}}$: Mean of the observed values

Fig. 11 Log-transmissivity (*T*) field estimated by inverse problem ordinary universal co-kriging with three variables: log-transmissivity data from pumping tests, log-transmissivity data from borehole lithofacies and water head data. **a** Using log-transmissivity data from pumping tests as the primary variable (IP-UCOK-1-2 in Table 5). **b** Using log-transmissivity data from borehole hydrofacies as the primary variable (IP-UCOK-2-1 in Table 5). The units of log-transmissivity are ln(m²/day)

σ_o Standard deviation of observed values
σ_m Standard deviation of modelled values
μ_o Mean of observed values, i.e., equal to $\overline{h_i^o}$
μ_m Mean of modelled values
N Number of spatial locations at which the water head was measured

The ME should be zero for an unbiased model. The MSE, MNSE and MAE should be as small as possible, whereas NSE and KGE should be as close to 1 as possible. The results obtained for the different validation statistics for the different geostatistical methodologies are shown in Table 5.

A field of constant log-transmissivity was used as a reference for comparison and the validation results are also shown in Table 5. The ordinary kriging map of the log-transmissivity data from pumping tests provides a map that produces worse results (MSE = 5.4197, in Table 5) than considering a field of constant log-transmissivity (MSE = 5.2864). This implies that the image in Fig. 6a is not a good representation of the true unknown underlying log-transmissivity field and even a constant image—for example, the mean log-transmissivity value, gives better performance results. However, the ordinary kriging map of log-transmissivity obtained by using the lithofacies and shown in Fig. 6b produces the third-best result (MSE = 1.7418) of all the methods that have been tested. A comparison of Figs. 6a and 6b shows that Fig. 6b has significantly more spatial detail than Fig. 6a, especially in the elongated high log-transmissivity values in the centre of the western part of the aquifer. The use of ordinary co-kriging with the log-transmissivity data as the primary variable makes very little difference (MSE = 5.2764) when compared with ordinary kriging (MSE = 5.4197). Using ordinary co-kriging with log-transmissivity data from lithofacies gives worse results (MSE = 2.7325) than ordinary kriging (MSE = 1.7418). In fact, it seems that, despite the coherence of the basic statistics and histograms (Figs. 2b and 3c) of both log-transmissivity data sets, they have no spatial correlation for a distance of zero as may be seen in Fig. 9c, which shows the spatial cross-covariance between the two data sets. Thus, the cross-variogram (Fig. 5c) used in ordinary co-kriging does not capture the spatial shift between the two data sets and the sill of the cross-variogram indicates a low correlation of around 0.3 between the two data sets. As a result, merging the two data sets is not useful as the cross-variogram cannot capture the complexity of the cross-covariance in Fig. 9c; however, each data set can be used independently of the other. When applying inverse problem universal co-kriging, Fig. 10a is obtained by using the log-transmissivity data from pumping tests and the water head data provide the best image as measured by the validation statistics in Table 5 (MSE = 0.6305). This image is even better than the solution given by inverse problem universal co-kriging using the log-transmissivity data from the lithofacies and the water head data (MSE = 1.3799), even though the log-transmissivity from the lithofacies data were better in ordinary kriging and ordinary co-kriging. The images of the log-transmissivity fields (Fig. 10a,b) are clearly improved by including the water head data irrespective of the log-transmissivity data set used.

Finally, if the three variables (the two log-transmissivity data sets and the water head data) are used in the inverse problem universal co-kriging, the results deteriorate because of the spatial incompatibility of the two log-transmissivity data sets. When the pumping test data are used as the primary variable the results are bad (MSE = 5.9882). They are also bad when the primary variable is the lithofacies data (MSE = 2.9084). The estimated log-transmissivity fields are given in Fig. 11.

To put these results into perspective it must be said that the aquifer boundary conditions shown in Fig. 8 have a strong influence on the output of the direct modelling and even using a constant log-transmissivity field gives good results as can be seen in Fig. 12a where the scatterplot between observed versus modelled water head data can be compared with Fig. 12b, which shows the scatterplot with the best results obtained by inverse problem universal kriging using the pumping test log-transmissivities and water head data in Fig. 10a. In terms of the NSE and KGE validation statistics, the solutions provided by the different methods are very similar as can be seen in Table 5. However, the best estimated log-transmissivity field is the one in Fig. 10a, which was obtained by inverse problem universal co-kriging using the pumping test data and water head data. This figure shows remarkably well the high log-transmissivity channel that runs through the central part of the aquifer (highlighted with a white arrow in Fig. 11a) and follows the course of the Genil River (black arrow in Fig. 1c), thus confirming the conceptual model proposed by the hydrogeologists who have worked in this area over an extended period (Castillo 1986; Luque-Espinar 2001; García-Soldado 2009). This conceptual model consists of a channel of high transmissivity running through the middle of the aquifer with transmissivity decreasing towards the borders; however, Fig. 10a provides a quantitative and more complete picture than the simple conceptual scheme proposed by the hydrogeologists.

It is well-known that even the best estimated image, such as the one in Fig. 10a, is a smoothed version of the true but unknown log-transmissivity field. Since Delhomme (1979), geostatistical simulation has been used to generate simulated images of the transmissivity field of an aquifer in which each image reproduces the real variability of the log-transmissivity as expressed by the variogram. A representative set of these images includes the uncertainty introduced by the spatial variability of the transmissivity between the experimental data. It is possible that one of these simulated images is very close to

Table 5 Mean error (*ME*), mean square error (*MSE*), mean normalized square error (*MNSE*), mean absolute error (*MAE*), Nash and Sutcliffe efficiency (*NSE*) and the Kling-Gupta Efficiency (*KGE*) validation statistics for the different geostatistical methods. The geostatistical methods are ordinary kriging with log-transmissivity from pumping tests (*OK-1*), ordinary kriging with log-transmissivity data from boreholes (*OK-2*), ordinary co-kriging with log-transmissivity data from pumping tests and from hydrofacies (*COK-1-2*), inverse problem universal co-kriging with log-transmissivity data from pumping tests and head data (*IP-UCOK-1*) and inverse problem universal co-kriging with log-transmissivity data from hydrofacies and head data (*IP-UCOK-2*)

Geostatistical method	ME	MSE	MNSE	MAE	NSE	KGE
Constant ln T	0.0027	5.2864	0.0085	1.6880	0.9959	0.9800
OK-1	0.0070	5.4197	0.0087	1.7366	0.9958	0.9797
OK-2	0.0738	1.7418	0.0028	0.9529	0.9986	0.9867
OCOK-1-2	0.1374	5.2764	0.0084	1.6870	0.9960	0.9782
OCOK-2-1	0.1289	2.7325	0.0043	1.1788	0.9979	0.9831
IP-UCOK-1	-0.0041	0.6305	0.0010	0.6087	0.9995	0.9925
IP-UCOK-2	0.0449	1.3799	0.0022	0.8891	0.9989	0.9889
IP-UCOK-1-2	0.1121	5.9882	0.0096	1.7969	0.9954	0.9776
IP-UCOK-2-1	0.1100	2.9084	0.0046	1.2241	0.9977	0.9829

the reality but all of them must be used because the reality is unknown. Obviously, the uncertainty of the parameters of the variogram model should also be taken into account when generating the simulated images. By using a large number of simulated images, it is, for example, possible to transfer the uncertainty in the transmissivity field into the uncertainty of the modelled water head. Thus, at each location, instead of having a single modelled water head, there is an ensemble of water heads the histogram of which provides the probability of the water head having a given interval value or being higher than a given threshold at each location. However, this is not the purpose of the work presented here, which is to identify smooth log-transmissivity images that contain as much detail as possible, and that can be used, *inter alia*, for modelling, optimization of observation grids and planning future geophysical campaigns.

Conclusions

Co-kriging, in its various forms, provides a means of combining different sources of hydrogeological information, which is particularly useful in estimating transmissivity fields. A case study of a detritic unconfined aquifer in southern Spain has been used as a benchmark for comparing methodologies. The transmissivity field generated from the lithofacies measured at boreholes demonstrates the importance of the indirect measures of log-transmissivity obtained from indirect data. This is particularly important because log-transmissivity data from pumping tests are biased towards high values as pumping tests are usually conducted to validate boreholes for productive water extraction. Nevertheless, the validation results show the incompatibility of the two types of information—the hard log-transmissivity data from the pumping tests and
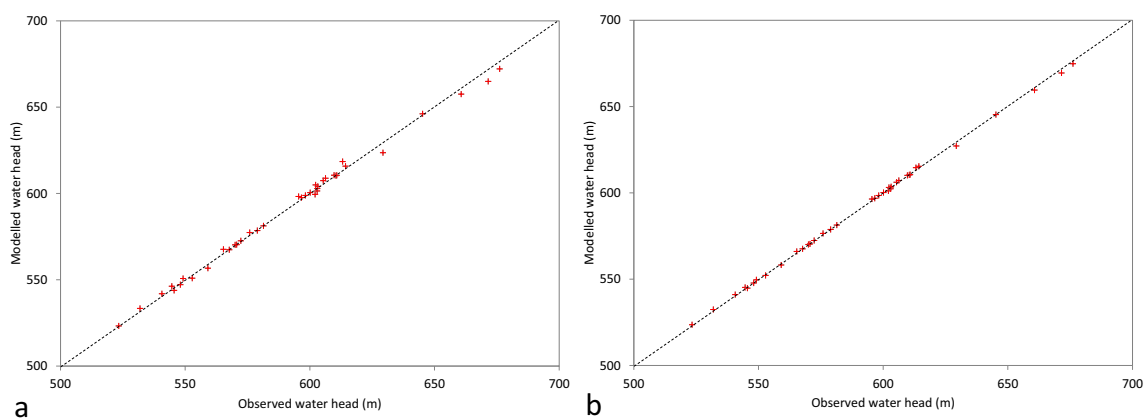
Fig. 12 Observed water head versus modelled water head for two extreme cases of the estimated log-transmissivity field. **a** Constant log-transmissivity. **b** The method with the best performance in Table 5, that is, inverse problem universal co-kriging using log-transmissivity data from pumping tests and water head data (IP-UCOK-1)

the soft log-transmissivity data from the lithofacies. Each log-transmissivity datum from the pumping tests encapsulates the permeability value (or hydraulic conductivity) and the saturated thickness at the location of the test. On the other hand, the transmissivity data from the lithofacies is proportional to the transmissivity from pumping tests (Table 3) and does not contain information on saturated thickness. A mean saturated thickness of 100 m for the aquifer is assumed in order to transform the permeability data into transmissivity data or, more exactly, data with "transmissivity units". The incompatibility between the two log-transmissivity data sets is due to the spatial variability of the saturated thickness. The log-transmissivity data from the pumping test incorporates that spatial variability (at least in theory), whereas the log-transmissivity data from the lithofacies does not.

However, even though the log-transmissivity data sets are incompatible, if there were no pumping test data, the log-transmissivity data from the lithofacies would provide reasonable images of the log-transmissivity field that could be further improved by incorporating the water head data by inverse modelling with universal co-kriging. Nevertheless, the best results were obtained by inverse problem universal co-kriging using the log-transmissivity data from pumping tests and the water head data (Fig. 10a; Table 5). In addition, the ability of analytical solutions to obtain the cross-covariance between log-transmissivity and water head data has been demonstrated. In particular, the analytical solution of Dagan (1985) overcomes the necessity of fitting a model to a complex cross-covariance that is anisotropic, antisymmetric and difficult to visualize from the experimental cross-covariances. However, a verification of the analytical solution using the experimental data is desirable. A comparison of the different estimated log-transmissivity values can be used to improve knowledge of the transmissivity field and could be used to optimize the locations for acquiring additional experimental data—pumping test data, borehole data and geophysical data. In the interests of open science, the experimental data used in the work have been provided. This will allow others to verify the work or assess alternative methodologies to obtain more reliable log-transmissivity fields.

Declarations

References

Ahmed S, de Marsily G (1987) Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resour Res 23:1717–1737

Ahmed S, de Marsily G (1993) Co-kriged estimation of aquifer transmissivity as an indirect solution of the inverse problem: a practical approach. Water Resour Res 29:521–530

Anderson MP, Woessner WW, Hunt RJ (2015) Applied groundwater modeling. Simulation of flow and advective transport. Academic Press, London, Second Edition, p. 564

Boucher M, Favreau G, Vouillamoz JM, Nazoumou Y, Legchenko A (2009) Estimating specific yield and transmissivity with magnetic resonance sounding in an unconfined sandstone aquifer (Niger). Hydrogeol J 17:1805–1815

Carrera J, Alcolea A, Medina A, Hidalgo J, Slooten LJ (2005) Inverse problem in hydrogeology. Hydrogeol J 13:206–222

Castillo A (1986) Estudio Hidroquímico del acuífero de la Vega de Granada [Hydrochemical study of the Aquifer de la Vega, Granada]. PhD Thesis, University of Granada, Spain, 658 pp

Chilès J-P, Delfiner P (1999) Geostatistics: modelling spatial uncertainty. Wiley, New York

Christensen S (1996) On the strategy of estimating regional-scale transmissivity fields. Ground Water 35:131–139

Cordano E, Rigon R (2013) A mass-conservative method for the integration of the two-dimensional groundwater (Boussinesq) equation. Water Resour Res 49:1058–1078

Dagan G (1985) Stochastic modelling of groundwater flow by unconditional and conditional probabilities: the inverse problem. Water Resour Res 21:65–72

Demir MT, Copty NK, Trinchero P, Sanchez-Vila X (2017) Bayesian estimation of the transmissivity spatial structure from pumping test data. Adv Water Resour 104:174–182

Dowd PA, Pardo-Igúzquiza (2006) Core-log integration: optimal geostatistical reconstruction from secondary information. Appl Earth Sci (Trans Inst Min Metall B) 115:59–78

Dowd PA, Pardo-Igúzquiza E (2012) Integration of spatial geophysical data by geostatistical simulation. Zeitschr Geol Wissenschaft 40:267–280

Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs, NJ

Freixas G, Fernández-Garcia F, Sanchez-Vila X (2017) Stochastic estimation of hydraulic transmissivity fields using flow connectivity indicator data. Water Resour Res 53:602–618

García-Soldado MJ (2009) Metodología basada en SIG para el desarrollo de un Sistema Soporte de Decisión en la gestión de la calidad de los recursos hídricos subterráneos de la Vega de Granada [Methodology based on SIG for the development of a decision support system in the management of the quality of subterranean water resources of the Vega of Granada]. PhD Thesis, University of Granada, Spain, 352 pp

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

Hoeksema RJ, Kitanidis PK (1984) An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modelling. Water Resour Res 20:1003–1020

Isaaks EH, Srivastava RM (1989) Applied geostatistics. Oxford University Press, New York

Journel A, Huijbregts C (1978) Mining geostatistics. Academic, New York, 600 pp

Kitanidis PK (1983) Statistical estimation of polynomial generalized covariance functions and hydrologic applications. Water Resour Res 19:909–921

Kitanidis PK (1996) On the geostatistical approach to the inverse problem. Adv Water Resour 19(6):333–342

Kitanidis PK (1997) Introduction to geostatistics: applications in hydrogeology. Cambridge University Press, Cambridge, UK

Knoben WJM, Freer JE, Woods RA (2019) Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrol Earth System Sci 23:4323–4331

Kuhlman K, Pardo-Igúzquiza E (2010) Universal co-kriging of hydraulic heads accounting for boundary conditions. J Hydrol 384:14–25

Kupfersberger H, Blöschol G (1995) Estimating aquifer transmissivities: on the value of auxiliary data. J Hydrol 165:85–99

Luque-Espinar JA (2001) Analisis geoestadistico espaciotemporal de la variabilidad piezometrica: aplicacion a la Vega de Granada [Spatiotemporal geostatistical analysis of piezometric variability: application to Vega de Granada]. PhD Thesis, University of Granada, Spain, 313 pp

Mardia KV, Marshall RJ (1984) Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika 71:135–146

Nash JE, Sutcliffe (1970) River flow forecasting through conceptual models. J Hydrol 10:282–290

Pardo-Igúzquiza E (1997) MLREML: A computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood. Comput Geosci 23:153–162

Pardo-Igúzquiza E (1998) Maximum likelihood estimation of spatial covariance parameters. Math Geol 30:95–108

Pardo-Igúzquiza E, Chica-Olmo M (2004) Estimation of gradients from sparse data by universal kriging. Water Resour Res 40(W12418):1–17

Pardo-Igúzquiza E, Chica-Olmo M (2007) KRIGRADI: a co-kriging program for estimating the gradient of spatial variables from sparse data. Comput Geosci 33:497–512

Pardo-Igúzquiza E, Chica-Olmo M, Luque-Espinar JA, García-Soldado MJ (2009) Using semi-variogram parameter uncertainty in hydrogeological applications. Groundwater 47:25–34

Pardo-Igúzquiza E, Chica-Olmo M, Rigol-Sánchez JP, Luque-Espinar JA, Rodríguez-Galiano V (2011) Una revisión de las nuevas aplicaciones metodológicas del cokrigeaje en Ciencias de la Tierra [A review of the new methodological applications of cokriging in Earth Sciences]. Bol Geol Min 122:497–516

Pulido-Velazquez D, Sahuquillo A, Andreu J, Pulido-Velazquez M (2007) A general methodology to simulate groundwater flow of unconfined aquifers with a reduced computational cost. J Hydrol 338:42–56

Razack M, Huntley D (1991) Assessing transmissivity from specific capacity in a large and heterogeneous alluvial aquifer. Groundwater 29:856–861

Renard P (2005) The future of hydraulic tests. Hydrogeol J 13:259–262

Renard F, Jeannée N (2008) Estimating transmissivity fields and their influence on flow and transport: the case of Champagne mounts. Water Resour Res 44:W11414

Richard SK, Chesnaux R, Rouleau A, Coupe RH (2016) Estimating the reliability of aquifer transmissivity values obtained from specific capacity tests: examples from the Saguenay-Lac-Saint-Jean aquifers, Canada. Hydrol Sci J 61:173–185

Rubin Y (2003) Applied stochastic hydrogeology. Oxford University Press, Oxford, UK, 391 pp

Soupios PM, Kouli M, Vallianatos F, Vafidis A, Stavroulakis G (2007) Estimation of aquifer hydraulic parameters from surficial geophysical methods: a case study of Keritis Basin in Chania (Crete – Greece). J Hydrol 338:122–131

Teramoto EH, Montanheiro F, Chang HK (2021) An alternative approach to designing conceptual models in cases of scarce field data. Groundwater Sustain Dev 15:100695

Wackernagel H (2003) Multivariate geostatistics: an introduction with applications, 3rd edn. Springer, Heidelberg, Germany

Xue P, Wen Z, Park E, Jakada H, Zhao D, Liang X (2022) Geostatistical analysis and hydrofacies simulation for estimating the spatial variability of hydraulic conductivity in the Jianghan Plain, central China. Hydrogeol J 30:1135–1155

Yeh WW-G (1986) Review of parameter identification procedures in groundwater hydrology: the inverse problem. Water Resour Res 22:95–108

Zhu L, Gong H, Chen Y, Li X, Chang X, Cui Y (2016) Improved estimation of hydraulic conductivity by combining stochastically simulated hydrofacies with geophysical data. Sci Rep 6:22224

Zimmerman DA, de Marsily G, Gotway CA, Marietta MG, Axness CL, Beauheim RL, Bras RL, Carrera J, Dagan G, Davies PB, Gallegos DP, Galli A, Gómez-Hernández J, Grindrod P, Gutjahr AL, Kitanidis PK, Lavenue AM, McLaughlin D, Neuman SP et al (1998) A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modelling advective transport by groundwater flow. Water Resour Res 34:1373–1413