Journal of
Hydroinformatics

Journal of Hydroinformatics Vol 25 No 3, 611 doi: 10.2166/hydro.2023.059

Effects of pumping flow rates on the estimation of hydrogeological parameters

Marios C. Kirlas [a,*] and Nikolaos Nagkoulis [b]

[a] Department of Agriculture, Aristotle University of Thessaloniki,Thessaloniki, Greece
[b] Department of Civil Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece
*Corresponding author. E-mail: kirlasmarios@agro.auth.gr

MCK, 0000-0003-2758-562X

ABSTRACT

In this paper, we examine the accuracy of estimating the hydrogeological parameters, transmissivity (T) and storativity (S), in a confined aquifer, when there are not enough available data for pumping flow rate values. While the most popular methods, used to estimate aquifer characteristics, assume that the pumping flow rate is constant during pumping, this is practically infeasible. Violation of this assumption results in errors, which are examined in this paper using field drawdown measurements. To find the aquifer characteristics, we use two methods, testing various pumping flow rates. Firstly, we employ the Cooper -Jacob equations to calculate (T) and (S) values. Afterwards, we use these values to create hypothetical drawdowns using Theis equation and finally we estimate the Root Mean Square Error (RMSE) between the actual and the hypothetical drawdowns. Then, we repeat the same process, replacing the Cooper -Jacob equations with Genetic Algorithms and Theis equation to find the aquifer characteristics by minimizing the RMSE between the actual and the hypothetical drawdowns. Although the process is applied only in three datasets, the results indicate that regardless of the method used, the obtained values of aquifer characteristics (T, S) are not considerably affected by inaccurate pumping flow rate estimations.

Key words: Cooper–Jacob, Genetic Algorithms, inverse groundwater problem, pumping flow rates, storativity, transmissivity

HIGHLIGHTS

- Actual data were used to find how the accuracy of the pumping flow rate's estimation affects the transmissivity and storativity results.
- Genetic Algorithms are applied to identify the aquifer characteristics, giving better results than Cooper-Jacob.
- The estimated aquifer characteristics' error varies linearly with the estimated pumping flow rate values.

INTRODUCTION

Groundwater, as a source of water supply, is of great importance for many rural and urban communities. It is a valuable resource that nearly half of the world's population uses for several activities such as irrigation, consumption and industrial use (Brindha & Elango 2015). Nonetheless, during the last decades, intensive agricultural activities contributed to the quantitative and qualitative degradation of groundwater, as traditional irrigation practices were widely used without consistent management of chemical fertilizers and pesticides (Ncibi *et al.* 2020). Moreover, groundwater quality and quantity are being threatened by increasing population and changing lifestyles, overconsumption, rapid urbanization, industrial wastewater and abusive farming practices (Gardner & Vogel 2005; Saidi *et al.* 2011; Hamed *et al.* 2022; Kirlas *et al.* 2022). Well-planned management of groundwater requires reasonable estimation of the hydrogeological parameters of an aquifer, such as transmissivity (T) and storativity (S). The calculation of T and S can be used for the modeling of groundwater flow, as well as for the prediction of contaminant transport as a step required for the planning and implementation of groundwater remediation activities (Sanchez-Vila & Fernàndez-Garcia 2016; Demir *et al.* 2017). Proper evaluation of these parameters, based on drawdown measurements, constitutes the inverse problem of groundwater hydraulics (Yeh 2015). The aim of the inverse modeling approach is to find the aquifer parameter values that minimize an objective function that calculates the differences between the observed and the simulated values of the state variables (Smaoui *et al.* 2018). Difficulties arising quite often in praxis are scarcity of accurate and sufficient

---

groundwater level data, economically unattainable measurements, as well as the inaccessibility of study areas (Kirlas 2017; Smaoui *et al.* 2018).

The Theis (1935) and the Cooper & Jacob (1946) methods are still predominant for the evaluation of hydrogeological parameters (Chapuis 1992; Osiensky *et al.* 2000; Avci *et al.* 2012; Anomohanran & Iserhien-Emekeme 2014; Kirlas 2021; Pfannkuch *et al.* 2021; Ali *et al.* 2022). The basic assumptions underlying the methods are the following: the aquifer is confined, homogenous, isotropic and pumped at a constant flow rate, the pumping well penetrates the total thickness of the aquifer, the piezometric surface is horizontal before pumping and the well diameter is small (Kruseman & de Ridder 2000). Moreover, Boulton (1954, 1963) proposed an analytical method for unconfined aquifers by exhibiting a delayed yield concept, whereas Prickett (1965) suggested a systematic graphical approach based on the type curve methods of Boulton. Moench (1995) proposed a combination of the Boulton and Newman methods for unconfined aquifers.

Notwithstanding, additional techniques have been proposed for the calculation of aquifer parameters, including Newman analytical solution (Neuman 1972; Naderi & Gupta 2020; Gunawardhana *et al.* 2021), numerical evaluation (Halford *et al.* 2006; Tumlinson *et al.* 2006; Lin *et al.* 2010; Chattopadhyay *et al.* 2015; Calvache *et al.* 2016), electrical resistivity tomography (González *et al.* 2021; Rao & Prasad 2021), hydraulic tomography based on geostatistical inversion (Yin & Illman 2009; Illman *et al.* 2015), direct push technologies (Dietrich *et al.* 2008; Bohling *et al.* 2012) and supervised committee machine with training algorithms (Tabari *et al.* 2021).

Genetic Algorithms (GAs) are increasingly used in groundwater hydraulics, because of their ability to solve multivariable complex problems, with a known objective function (Katsifarakis & Kontos 2020). They are probabilistic algorithms that mimic the functioning of natural phenomena, such as genetic inheritance and the Darwinian struggle for survival. They have been widely used in optimizing quantitative and qualitative aquifer management (McKinney & Lin 1994; Rauch & Harremoës 1999; Erickson *et al.* 2002; Kontos & Katsifarakis 2017; Seyedpour *et al.* 2021).

Moreover, GAs have been used to estimate the transmissivity of non-homogenous aquifers under a steady flow (Karpouzos *et al.* 2001). Applications include coastal aquifers, as well (Smaoui *et al.* 2018). Ha *et al.* (2020) used a GA combined with the Levenberg–Marquardt algorithm and with the Neuman and Witherspoon model and ratio method to accurately estimate aquifer parameters from pumping tests. Thomas *et al.* (2018) proposed a new simulation–optimization model for the estimation of aquifer parameters by coupling the radial point collocation meshfree method with cat swarm optimization. GAs are combined with Theis equation in this paper to solve the inverse groundwater problem. Their results are compared with the results obtained from Cooper–Jacob for a variety of pumping flow rates. The next section briefly presents the framework of this analysis.

STUDY AREA DESCRIPTION

The aquifer of Nea Moudania (Figure 1) is located in the south-western part of the Chalkidiki peninsula, Northern Greece. Its total area is approximately 127 km^2, and it administratively belongs to the municipalities of Nea Propontida and Polygyros. In general, the study area has a low altitude (≈210 m) and gentle slopes, and it is the prime agricultural land of Chalkidiki (Panteli & Theodossiou 2016). The average annual precipitation for the flat and hilly areas is about 420 and 510 mm, respectively, while the climate is described as semi-arid to humid (Siarkos & Latinopoulos 2016). In the Peonia geological zone, the Nea Moudania aquifer consists of rocky formations in the north (ophiolite, clay schists and gneiss) and Neogene sediments and alluvial deposits in the south (sandstones, red clay, gravels, silts, sand and conglomerates) (Syridis 1990; Svigkas *et al.* 2020). Since rocky formations in the area are typically thought to be impermeable, recent deposits with significant sediment thickness and important water storage capacity are of significant hydrogeological interest, composing the main aquifer system (Kirlas 2017; Kirlas & Katsifarakis 2020). The aquifer system consists of an alternation of permeable and impermeable beds without standard geometric development and exhibits severe heterogeneity and complexity (Siarkos & Latinopoulos 2016).

Furthermore, in the study area, there is a high demand for water for domestic and agricultural irrigation, particularly during summer. However, there is an intense lack of surface water and low annual precipitation, making groundwater the only source of water that is viable. For this reason, a basic system of municipal and private wells can partially meet the total water demands (Latinopoulos *et al.* 2003). Figure 2 shows the representative lithological profile of the investigated wells, which is similar for all three wells. Moreover, it shows that they penetrate successfully different beds of clay, clay with gravels and gravels.

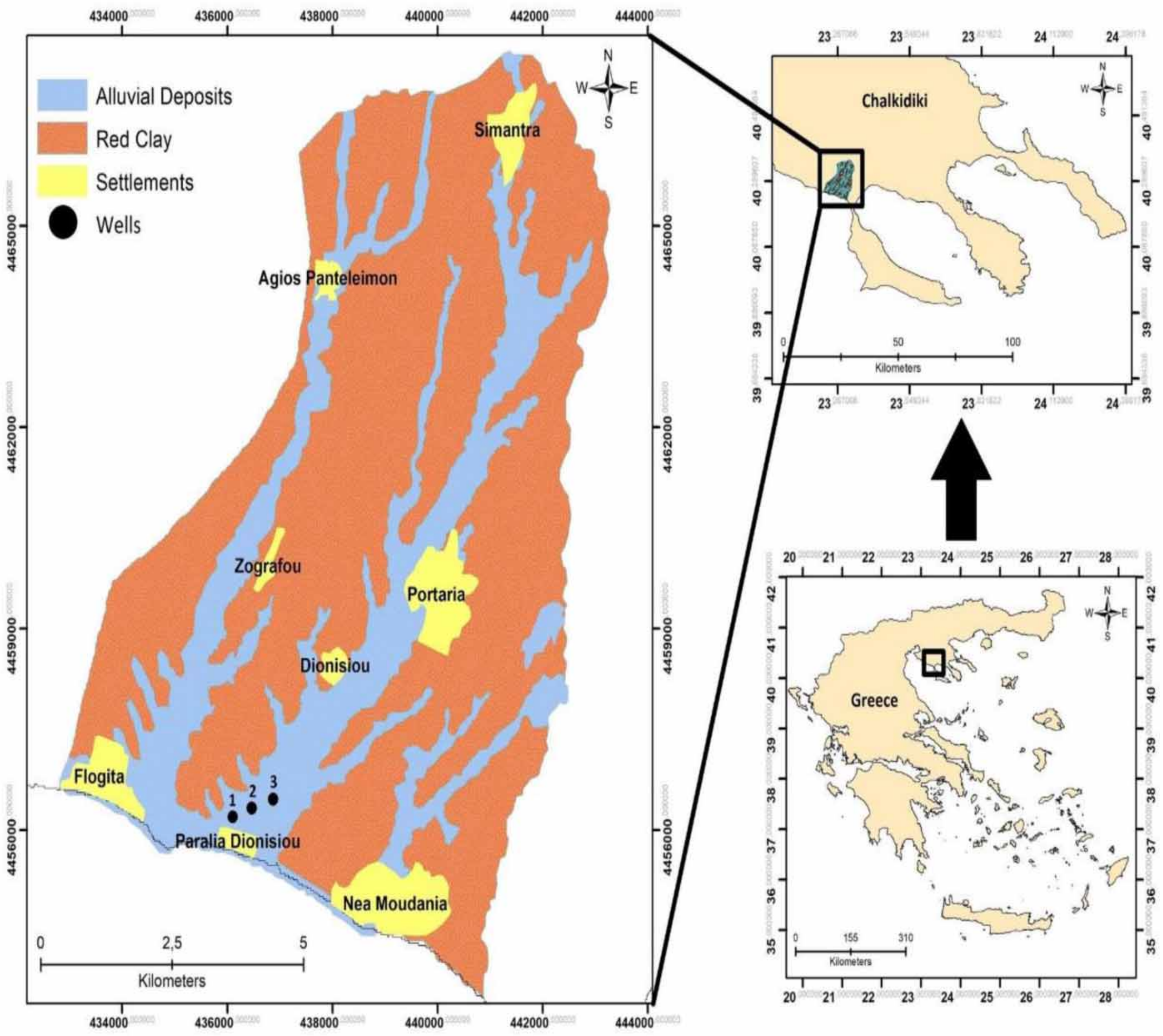

Figure 1 | The study area (Nea Moudania aquifer).

METHODOLOGY

Methodological framework

In this paper, we use three sets of hydraulic head drawdown measurements to test the effect of inaccuracy in pumping flow rate estimation on transmissivity (T) and storativity (S) evaluation. Two methods are employed using these datasets, for a number of possible pumping flow rate values. Figure 3 can be used to summarize the calculation of T and S for every pumping flow rate. During the first approach:

1. Cooper–Jacob equations are used in order to estimate T and S values.
2. The estimated values are used to create hypothetical drawdown curves, using Theis equation.
3. Root Mean Square Error (*RMSE*) is used to estimate the difference between the actual and the hypothetical drawdowns.
4. The process is repeated for a new pumping flow rate hypothesis.

During the second approach:

1. GAs are used to create populations and generations of possible T, S solutions.

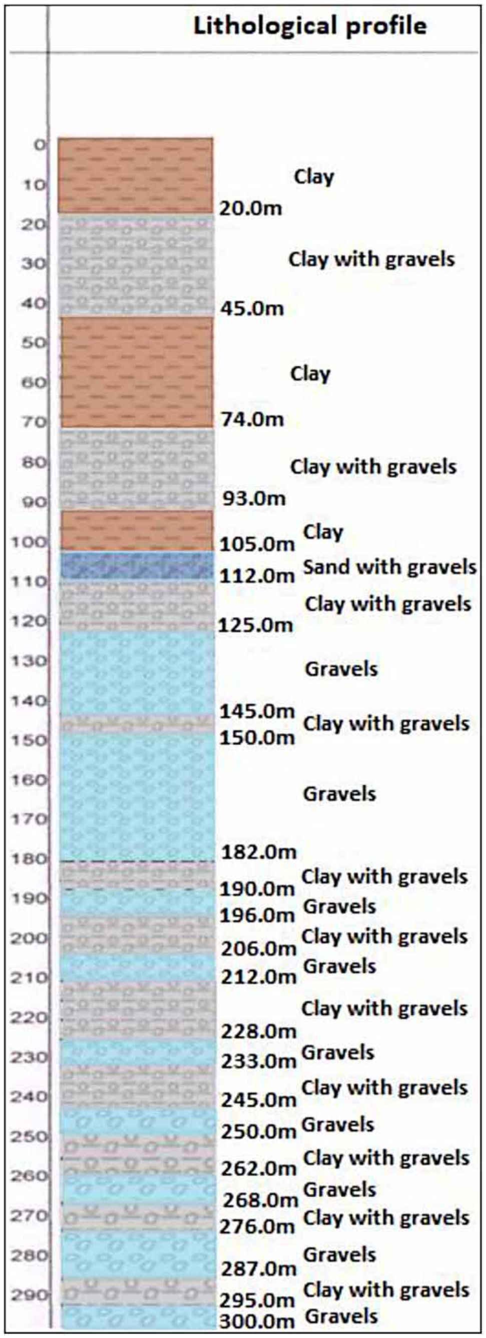

Figure 2 | Lithological profile of the tested wells.

2. Hypothetical drawdown curves are created using these T, S values and Theis equation.
3. *RMSE* is used to estimate the difference between the actual and the hypothetical drawdowns.
4. The set of T, S that minimizes the *RMSE* between the actual and the hypothetical drawdowns is considered the solution to the inverse problem. In other words, *RMSE* is the fitness value of the genetic algorithm. This set is the solution of the optimization.
5. The process is repeated for a new pumping flow rate hypothesis.

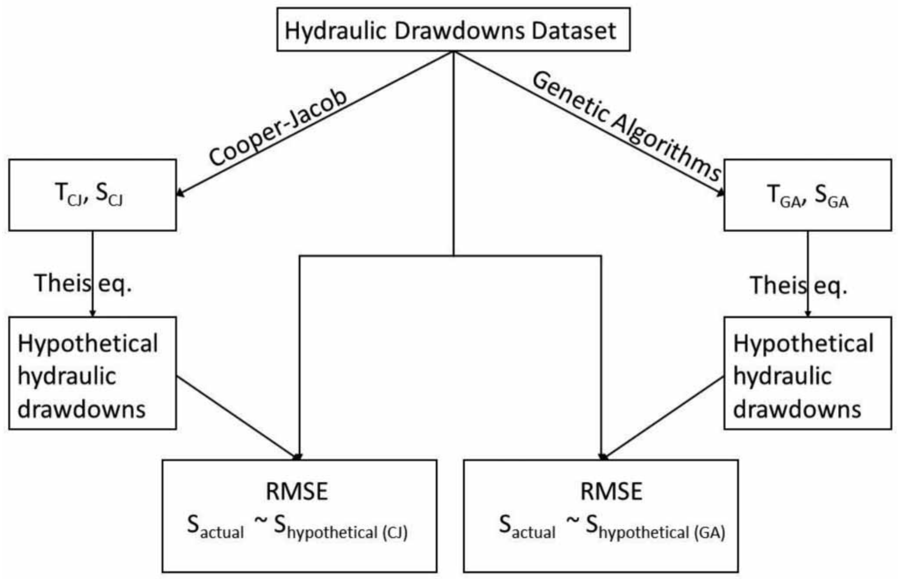

Figure 3 | Flow chart of the calculation of *T*, *S* and *RMSE* using actual hydraulic drawdowns. The calculation is repeated for a set of possible pumping flow rate values, for three hydraulic drawdown datasets.

Basic equations

According to Theis (1935), transient groundwater head level drawdown s_i at point i can be accurately calculated by Equation (1), as long as the assumptions mentioned in the previous section hold. The term $W(u)$ and u, appearing in Equation (1), are given by Equations (2) and (3).

$$s_{i,\Delta t} = \frac{1}{4\pi T} \sum_{j=1}^{j=N} Q_j \, W(u_{i,j,\Delta t}) \tag{1}$$

$$W(u_{i,j,\Delta t}) = -\gamma - \ln(u_{i,j,\Delta t}) - \sum_{k=1}^{\infty} \frac{(-1)^k (\, u_{i,j,\Delta t})^k}{k \, k!} \tag{2}$$

$$u_{i,j,\Delta t} = \frac{S \, r_{i,j}^2}{4T\Delta t} \tag{3}$$

In the above formulas, T represents aquifer's transmissivity, Q_j is the pumping flow rate of well j, γ is the Euler's constant, S is the aquifer's storativity, $r_{i,j}$ is the distance between point i and well j and Δt represents the duration of pumping.

For small u values, namely r and/or Δt, the third term of the right-hand side of Equation (2) can be neglected. Moreover, taking into account that 0.5772–ln(4/2.25) and substituting Napierian by decimal logarithm, Equation (1) can be transformed to Equation (4), which is known as Cooper–Jacob equation:

$$s_i = \frac{2.3Q}{4\pi \mathrm{T}} \log\left(\frac{2.25Tt}{r^2 S}\right) \tag{4}$$

For their method to be applicable, Cooper & Jacob (1946) recommended that u values should not exceed 10^{-2}. Many authors, such as Freeze & Cherry (1979), Schwartz & Zhang (2003), Todd & Mays (2005) follow their suggestion, while Fetter (2001) and Sterrett (2007) affirm that a maximum value of $u_{max} = 0.05$ is satisfactory. Nevertheless, Alexander & Saar (2011) propose a significantly higher value of $u_{max} = 0.2$.

Gomo (2019) demonstrated that in Cooper–Jacob method there might be some inaccuracies in transmissivity and storativity values, irrespective of the u value. For this reason, instead of using u, another objective criterion was proposed, namely the

Infinite Acting Radial Flow (IARF) condition, in order to determine the applicability of Cooper–Jacob method (Spane 1993; Renard *et al.* 2009; Gomo 2020).

Furthermore, Kirlas & Katsifarakis (2020) investigated the accuracy of T and S values and showed that the precise recording of pumping initiation and shutdown time is of crucial importance, because when pumping is estimated to start earlier than it actually does, transmissivity is underestimated, while storativity is overestimated. Additionally, they concluded that when the residual drawdown is substantial, the transmissivity value might be overestimated.

Despite the aforementioned restrictions, the Cooper–Jacob equation has been widely used to solve the inverse problem of groundwater, due to its simplicity. The simplicity stems from the linear relationship between s_i and the logarithm appearing in Equation (4). This allows easy graphical application of the Cooper–Jacob method, marking field measurement data on semi-logarithmic paper and plotting the straight line that best fits them. The transmissivity T is calculated first, based on the straight-line slope (Equation (5)); then, the point of its intersection with the logarithmic axis is used to calculate storativity S (Equation (6)). The respective formulas are:

$$T = \frac{2.3Q}{4\pi\Delta s/\Delta logt} \tag{5}$$

$$S = \frac{2.25Tt_0}{r^2} \tag{6}$$

Checking the validity of the results, using *RMSE*

The question that arises after the calculation of the aquifer characteristics is how to estimate their accuracy. In this paper, we propose *RMSE* as accuracy criterion. When solving an inverse problem, the calculated T and S values can be considered as accurate, if they can be used to 'reconstruct' the physical phenomenon. From this perspective, what Equations (5) and (6) do is finding the parameters T and S that can be used to 'reconstruct' the hydraulic drawdown curve. If this 'reconstruction' is accurate, then the differences between calculated and measured s_i values should be small. Supposing that we have calculated T and S values, we use them to reconstruct a hydraulic head drawdown curve, by means of the Theis equation. Then, the difference between this 'hypothetical' curve and the initial (real) curve can be calculated using *RMSE* (Equation (7)). When the values of *RMSE* are low, the hypothetical drawdown curve is very similar to the actual one.

$$RMSE = \sqrt{\sum_{i=1}^{i=m} \frac{\left(s_{i\,(\mathrm{hypothetical})} - s_{i\,(observed)}\right)^2}{m}} \tag{7}$$

In relationship (7), m indicates the number of actual (or observed) drawdown measurements. The density of the hypothetical measurements is higher than the density of the actual measurements. The hypothetical measurements can be easily obtained using a set of T and S, for any possible time discretization. Therefore, in order to make it possible to compare the actual values with the hypothetical values, we choose only these hypothetical values that take place at the same time as the observed values.

It is important to mention that *RMSE* indicates that the results are accurate in case the input data are accurate. In this paper, we use inaccurate data (testing a number of possible pumping flow rates), therefore this inaccuracy is inserted in the outputs of the algorithm. In other words, low *RMSE* values indicate that the inverse problem algorithm-method operates correctly. However, this does not guarantee that the outcomes should be trusted. This point is better explained in the following sections.

Use of GAs to calculate *T* and *S*

As mentioned in the previous sections, GAs have been used to estimate T and S values, based on field measurements. In this case, the decision variables, which are included in the chromosomes, are T and S. *RMSE* as defined above, can serve as an evaluation function. *RMSE* has been already efficiently combined with GAs (Bastani *et al.* 2010; Amaranto *et al.* 2018). *RMSE* is a useful tool, when combined with GAs, because it can be used to calculate the fitness values of the chromosomes. Bastani *et al.* (2010) use *RMSE* in groundwater flow modeling to compare how close observed and simulated values are and Amaranto *et al.* (2018) use *RMSE* to test the accuracy of forecasting. One recent study uses *RMSE* to find the characteristics of aquifers (*T*, *S* values) and the time schedules of pumping, when more than one wells pump simultaneously (Nagkoulis 2021). The main idea behind that is that the aquifer behaves like a conduit and the well-piezometer as a receiver. For each aquifer, there is a unique hydraulic drawdown curve 'received' by the well-piezometer. When the hypothetical hydraulic drawdown, created for a hypothetical set of *T* and *S* values, matches to the actual hydraulic drawdown measured at the well-piezometer, the hypothetical *T* and *S* values correspond to the actual aquifer characteristics. Consequently, the objective of GAs is to find the aquifer characteristics (T_0, S_0) that minimize $RMSE_{(T,S)}$. In Figure 4, we can see that after a number of generations, *RMSE* stabilizes, reaching its minimum values. The script of the aforementioned paper is applied in this one too, using actual drawdown data. The algorithm is written in R studio, using the GA package (Scrucca 2013).

In the aforementioned paper, it has been noticed that even though 1,000 generations were used, the solution was usually found before generation 300. The variable inputs are inserted in the algorithm in binary form using 20 digits for *T* and 17 digits for *S*. These digits are chosen so that the algorithm can search in a wide space from 10^{-1} to 10^{-8} approximately for possible *T* and *S* solutions. A typical 'rank selection' is used, following the package's initial settings. The computational time was approximately 10 h using a typical i7 CPU. It should be mentioned however that this time can even be reduced in 1 h in case that less generations are used. We have used the following parameters in our code: Number of generations: 1,000; population size: 100, crossover probability: 0.85, mutation probability: 0.45. Running the same code for 30 min (using 100 generations) for some specific cases, resulted in very similar results. The mutation probability was chosen after a number of tests. For low probabilities, the algorithm was often trapped in local minima. The selection process included elitism, preserving the three best chromosomes of the current generation for the new one.

Estimation of pumping flow rates

In most cases, the pumping flow rate is unsteady in practice. There can be many technical reasons for unintentional variation in pumping flow rates, such as pumping well's diameter (attachment of a smaller than suggested hose), intake line obstruction (a common problem is debris blockage and slurry flow) and improperly connected motor (due to incorrectly electrical connections to the electric motor) (Farokhzad *et al.* 2012; Derakhshan & Bashiri 2018). In the aforementioned cases, the error increases in time. The more time that a well operates, the higher the drawdowns get, the more difficult it is for a pumping system that operates insufficiently to pump water. Nonetheless, there is one more situation that has critical differences from the previous ones and therefore it should be considered a separate case, outside of the scopes of this paper. There can also be variations in the pumping flow rate, due to nearby pumping wells, which might start pumping during the pumping phase of the examined well. Whereas, in the previous cases the energy loss increases with time, in this case, there can be radical increases in drawdowns in the examined well, which cannot be modeled (without additional information) from the tools used.

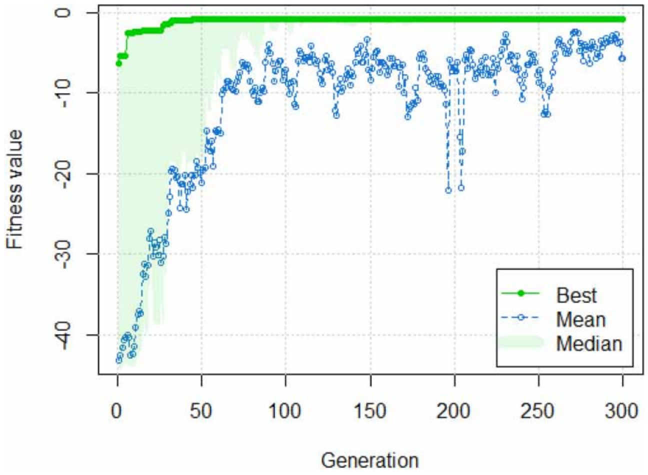

Figure 4 | Flow chart representing the fitness values (*RMSE*) of each generation for $Q = 30$ m^3/h (dataset 3) and for 300 generations. Although we used 1,000 generation in the main tests, we found out that in most cases the result had been found in the first 100–300 generations.

Even if we suppose that there is a perfect pump that supplies the system with energy constantly, the pumping flow rate will gradually reduce, since the flow rate and pump power are related through Equation (8):

$$HP_{pump} = \gamma \times Q \times s \tag{8}$$

In Equation (8), HP_{pump}(W) represents the power that the pump provides to the system, γ (N/m^3) is the specific weight of water, Q (m^3/s) is the pumping flow rate and s (m) is the hydraulic drawdown. For $HP_{pump}(t) = constant$ we get $Q(t)s(t) = constant$. This way, as the hydraulic drawdown increases, the pumping flow rate necessarily decreases. The practical solution to this problem is using an inverter to increase the energy offered by the pump to the system so that $HP_{pump}(t_1) < HP_{pump}(t_2)$.

Nonetheless, the main equations used in praxis (Theis and Cooper–Jacob) are derived under the assumption that pumping flow rates are constant. In order to deal with the fact that the flow rate in many cases decreases with time as the water level drops, Kruseman & de Ridder (2000) suggested checking and, if necessary, even adjusting the well flow rate on an hourly basis.

In this paper, on the one hand, we propose *RMSE* as a parameter that can be used to find out if Theis equation is still valid (under unsteady pumping flow rates) and on the other hand, we prove that errors in pumping flow rates estimation can result in minor errors in *T* and *S* evaluation.

RESULTS

Evaluation of *T* and *S* and calculation of *RMSE*

Firstly, we applied the Cooper–Jacob method to three groundwater level datasets (drawdown data), in order to determine the hydrogeological parameters, such as transmissivity *T* and storativity *S*. To create the diagrams and calculate the transmissivity and storativity values, we used MS Excel, considering that the deviations between MS Excel calculations and other computational tools, regarding the *T* and *S* values, are negligible (Kirlas 2017; Kirlas & Katsifarakis 2020).

In order to apply Cooper–Jacob method we used the following drawdown datasets. The first dataset (hourly data) was from one pumping cycle on April 11, 2018 (Figure 5). The duration of pumping was 1,380 min and the u value for the first value of drawdown was $5.14 \times 10^{-3} < 0.01$. The second dataset (5-min data) was from one pumping cycle on April 14, 2018. The duration of pumping was 975 min and u for the first drawdown value was $9.75 \times 10^{-3} < 0.01$ (Figure 6). The third dataset (5-min data) was from one pumping cycle on April 15, 2018. The duration of pumping was 810 min and the u value for the first drawdown equals to $9.75 \times 10^{-3} < 0.01$ (Figure 7). The pumping flow rates were estimated to be $Q = 30$ m^3/h. The results of both transmissivity and storativity values are shown in Table 1.

Secondly, after the Cooper–Jacob method application, we calculated *RMSE* to evaluate the accuracy of the values of the aforementioned parameters (T, S, Q). Specifically, we calculated the *RMSE* between the actual (red line) and the hypothetical

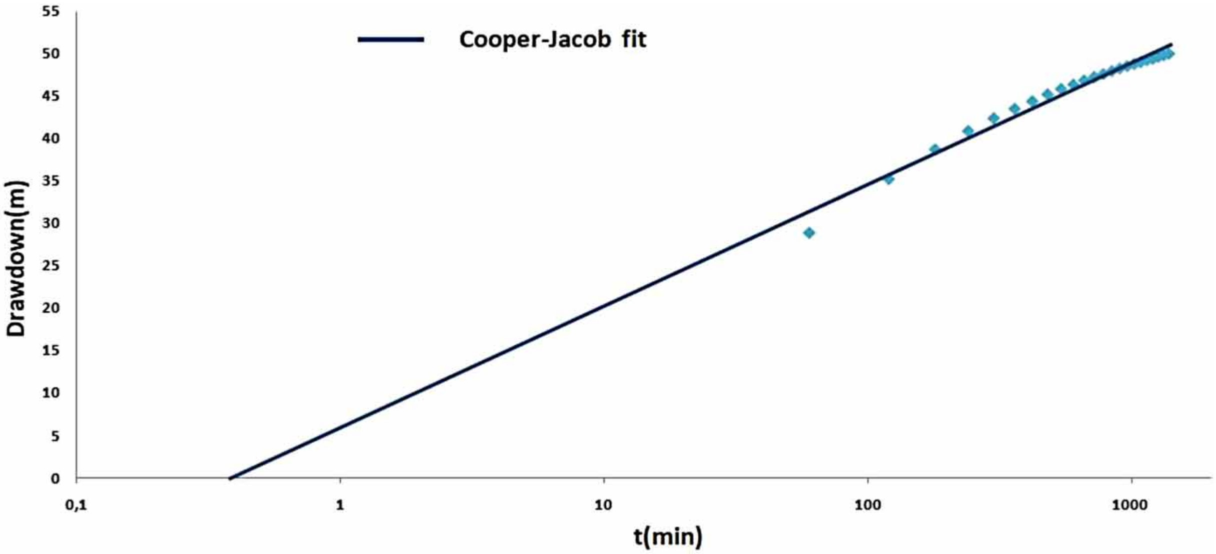

Figure 5 | Scatter plot showing the Cooper–Jacob model fit on drawdown data on April 11, 2018.

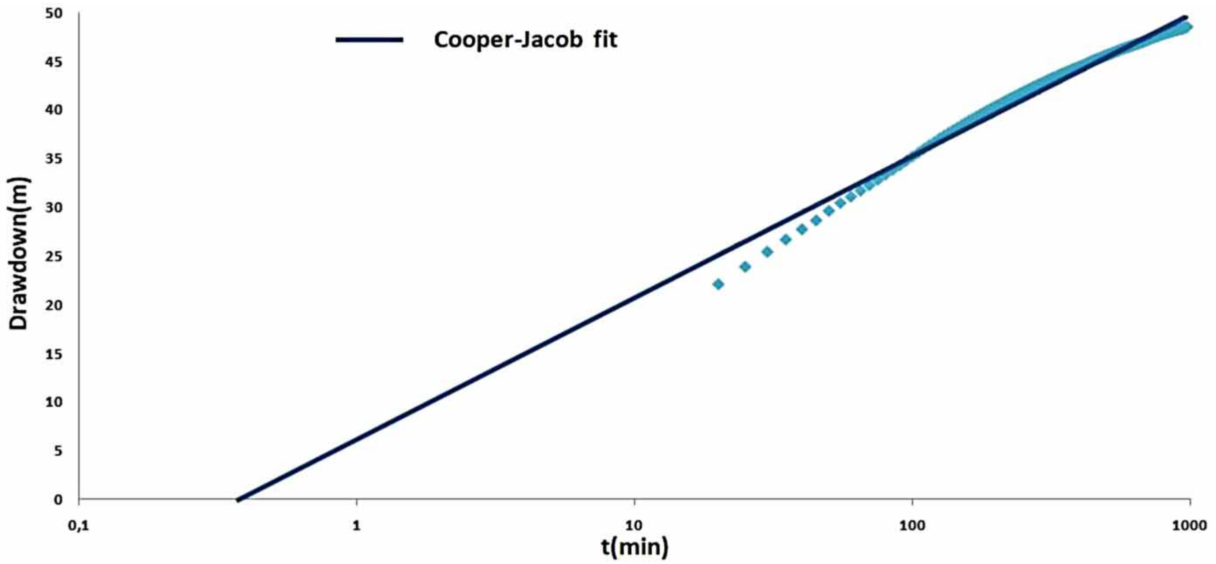

Figure 6 | Scatter plot showing the Cooper–Jacob model fit on drawdown data on April 14, 2018.

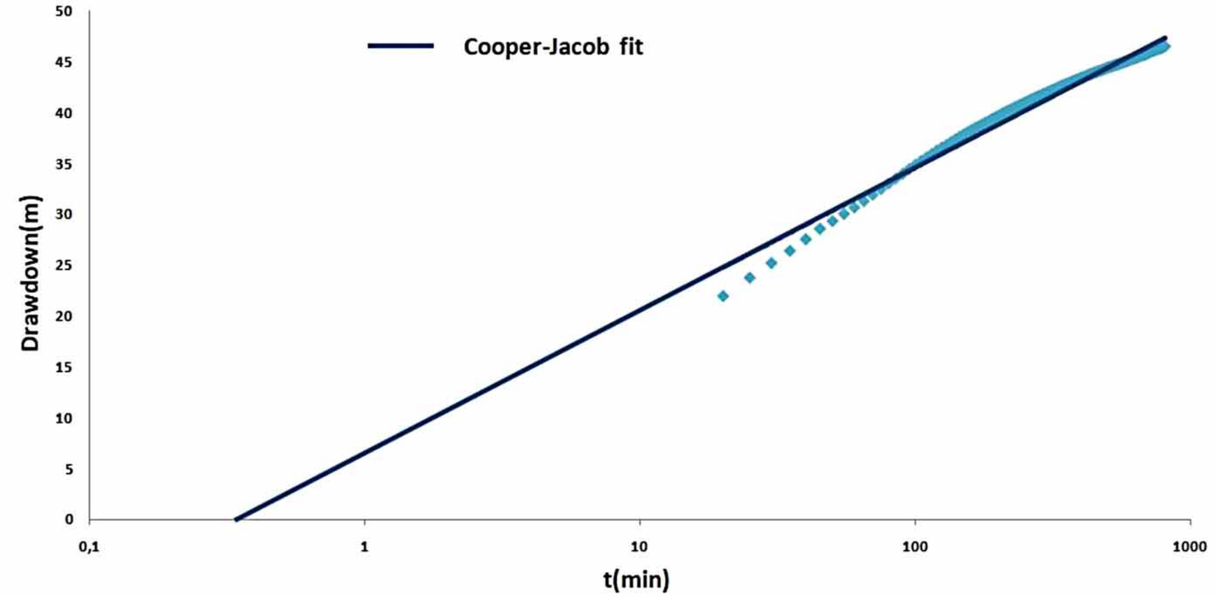

Figure 7 | Scatter plot showing the Cooper–Jacob model fit on drawdown data on April 15, 2018.

Table 1 | Results of *T* and *S* for different datasets

Cooper–Jacob analysis

Datasets	*Q* (m^3/h)	*T* (m^2/s)	*S*	Pumping duration (min)
First	30	1.028×10^{-4}	8.406×10^{-2}	1,380
Second	30	1.059×10^{-4}	5.766×10^{-2}	975
Third	30	1.085×10^{-4}	5.646×10^{-2}	810

(black line) drawdowns. The scatter plots of the actual and hypothetical drawdown for the three datasets can be seen in Figure 8. The resulted *RMSE* values are shown in Table 2.

We can see that the *RMSE* takes low values and varies from 0.6916 to 0.8980. Finally, it can be seen that the first dataset which contains fewer drawdown measurements results in higher *RMSE* than the next two datasets. This is because Cooper–Jacob method can solve more accurately the inverse problem when there are more data available.

Reevaluation of *T* and *S* and calculation of *RMSE* for *Q* variations

However, if the pumping flow rate values obtained are not accurate, the results obtained will differ. To test that, we used 16 pumping flow rate values from 17 to 32 m^3/h (the field measurements indicate $Q = 30\ m^3/h$) and we reevaluated the *T* and *S* values by using the graphical Cooper–Jacob method. We came up with the results of *T* and *S* appearing in Figures 9 and 10, respectively. It can be seen that reduction of *Q* results in a proportional decrease of both *T* and *S* values in all datasets. For instance, a 46.8% reduction of *Q* (from 17 to 32 m^3/h) leads to an equal 46.8% reduction of *T* and *S* values, implying a linear relationship between the values.

Then, we calculated the *RMSE* for different *T*, *S* and *Q* values. *RMSE* appear constant, regardless of the pumping flow rate values (Figure 11).

Variation of flow rate and calculation of *T*, *S* and *RMSE* using the GA

Finally, the *RMSE* is used to find the solution of *T* and *S* values for a possible range of pumping flow rates using the GA (Figures 12 and 13). It can be seen that the GA resulted in slightly lower *RMSE* than the graphical Cooper–Jacob method. This signifies the accurate and successful reproduction of the actual hydraulic drawdown as well as the accurate evaluation of the hydrogeological parameters *T* and *S*. Again, the *RMSE* is not affected by using different pumping flow rates (Figure 14).

DISCUSSION

From Figures 11 and 14, we can see that *RMSE* is not affected bythe pumping flow rate variations. The nearly-constant *RMSE* line means the methods solving the inverse problem operate ‘normally’ (Cooper–Jacob and GA–Theis). This means that there exists a logarithmic line or curve that can be used to ‘reconstruct’ the hydraulic drawdowns using these characteristics.

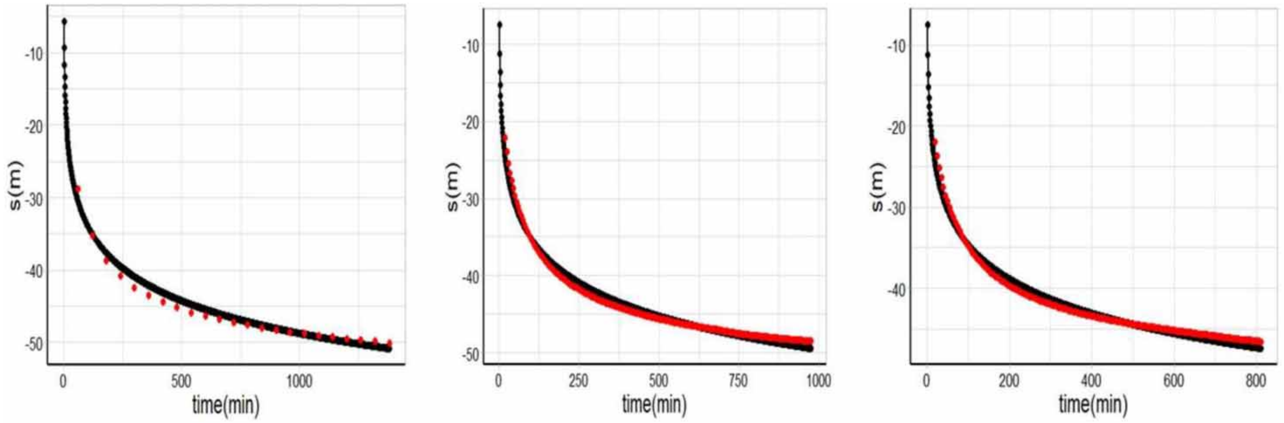

Figure 8 | Scatter plot for the first (left), second (middle) and third (right) dataset. Red line shows the real drawdown and black line shows the hypothetical drawdown. Please refer to the online version of this paper to see this figure in color: http://dx.doi.org/10.2166/hydro.2023.059.

Table 2 | Results of the *RMSE* between the real and the hypothetical drawdown

Datasets	*Q* (m^3/h)	Method	*RMSE*
First	30	Copper–Jacob	0.8980
Second	30	Copper–Jacob	0.7082
Third	30	Copper–Jacob	0.6916

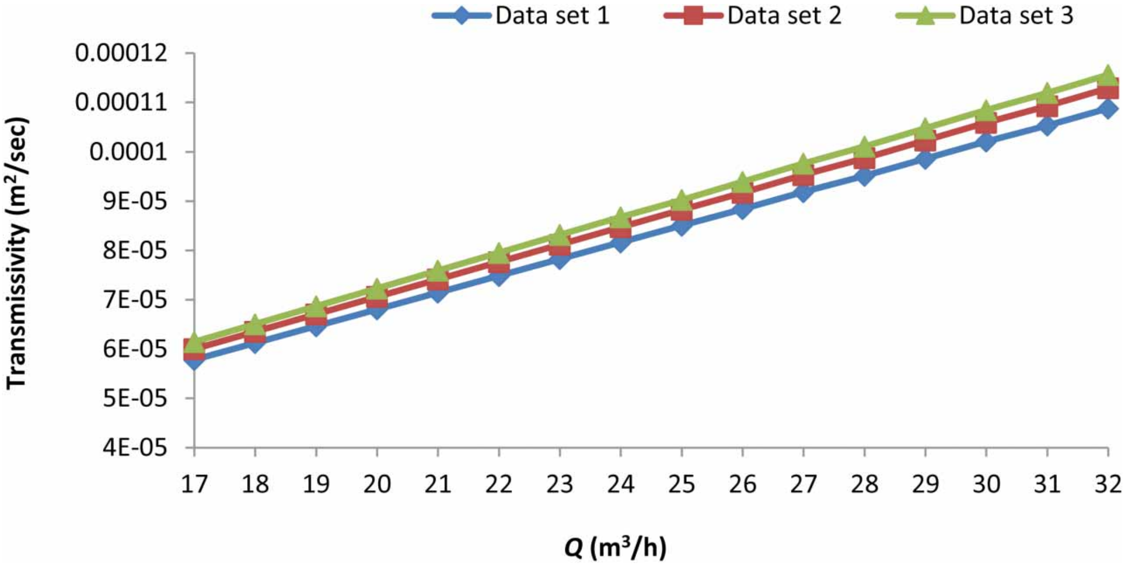

Figure 9 | Variation of flow rates and transmissivity for all datasets.

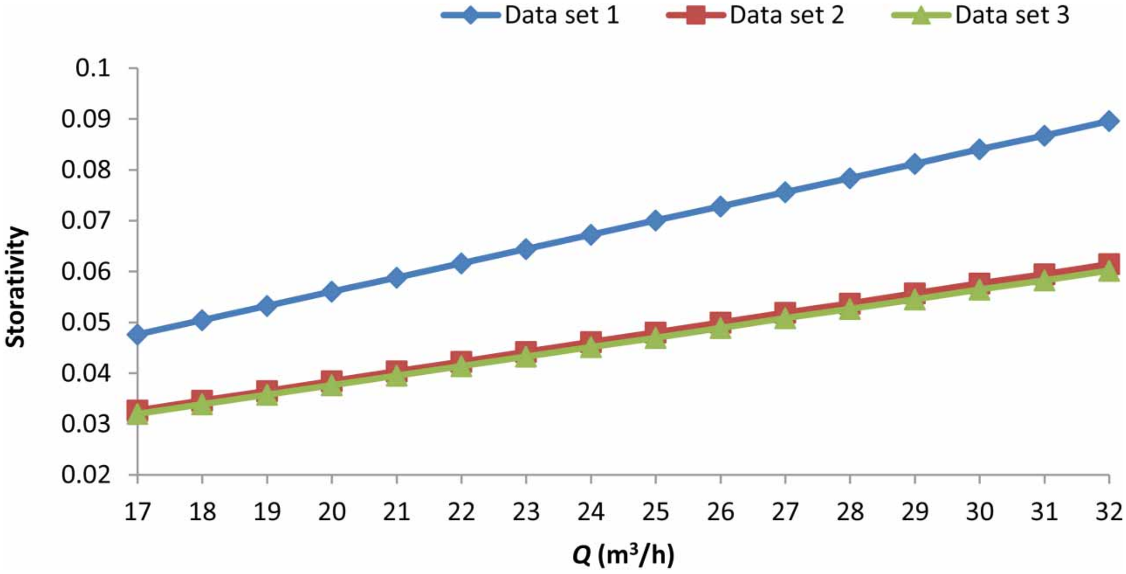

Figure 10 | Variation of flow rates and storativity for all datasets.

From the *RMSE* charts, we can see that there is no unique set of (Q, T, S) that minimizes that *RMSE*. If the pumping flow rate is not accurately obtained from the field survey studies, an error will be inserted in finding the T, S parameters. The variability of the pumping flow rate (which is assumed to be constant in the aforementioned equations), results in errors. This variability might be due to intense heterogeneity of aquifer's geological formations as well as technical issues of the water pumping system (e.g. mechanical jamming in pump, unsuitable pump selection, pump efficiency, power failure and short blackouts with zero pumping flow and intake line obstruction) (Trabucchi *et al.* 2018). Having tested both Cooper–Jacob and GA–Theis methods we can see that the GA–Theis approach results in lower *RMSE* values. However, the differences between the *RMSE* values are not strong. What is the most important is that in both cases the *RMSE* values present the same characteristics in terms of pumping flow rate variations

One of the most important results of this paper is the linear form of Q–T and Q–S graphs. In the inverse groundwater problem, researchers are usually interested in the order of magnitude of T and S. Thence, the linear relationship between T, S and

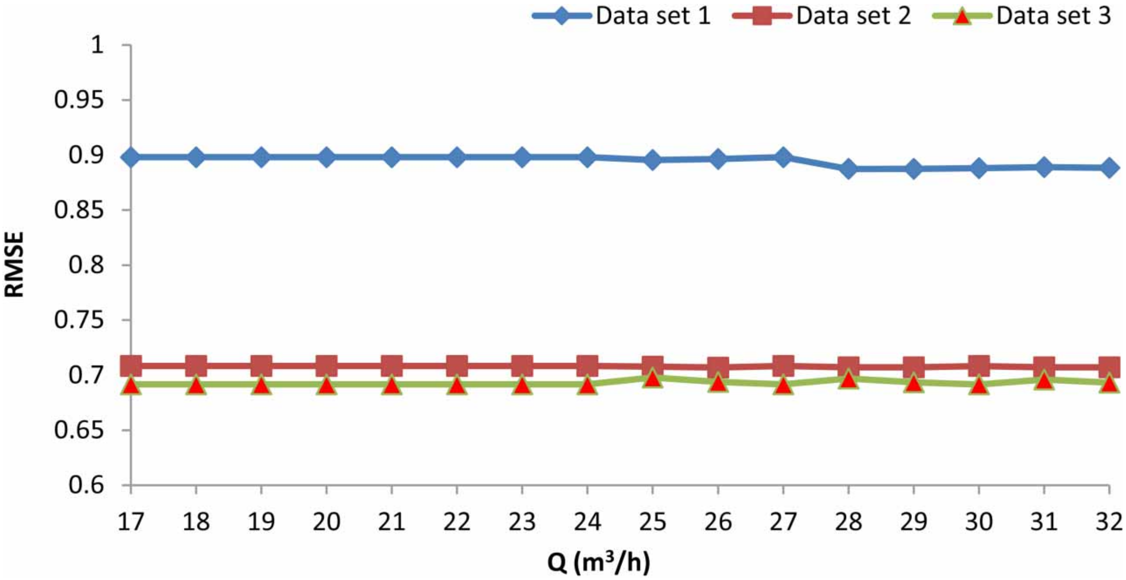

Figure 11 | Variation of flow rates and *RMSE* results for all datasets.

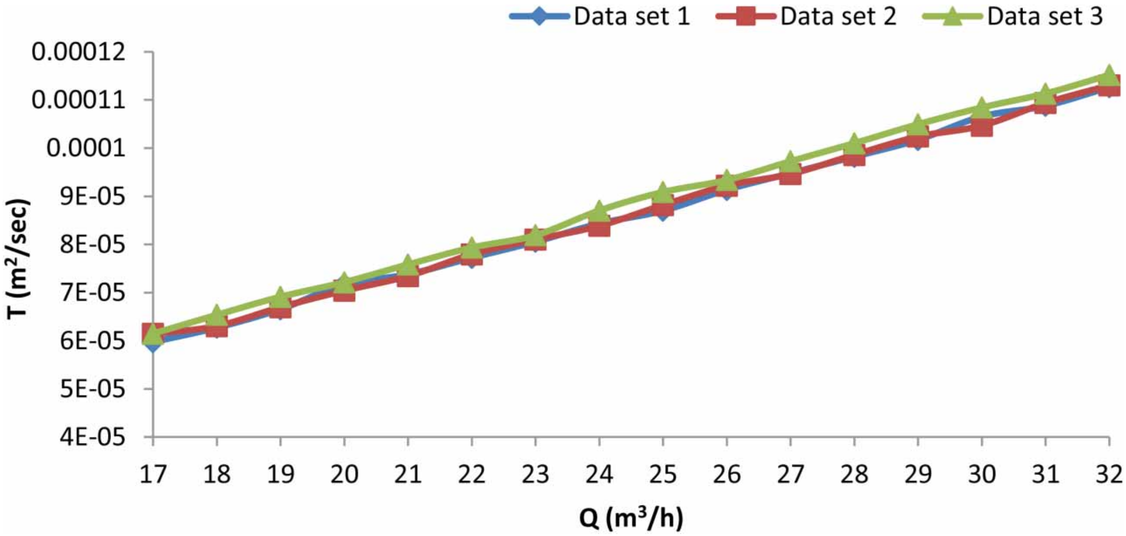

Figure 12 | Variation of flow rates and transmissivity for all datasets.

Q indicates that errors in Q do not strongly affect the results. For instance, using data from the third dataset, for $Q = 30\ \mathrm{m}^3/\mathrm{h}$ we get $T \simeq 1.1 \times 10^{-5}\ \mathrm{m}^2/\mathrm{s}$ and for $Q = 25$ we get $T \simeq 0.9 \times 10^{-5}\ \mathrm{m}^2/\mathrm{s}$. An error of 16% in Q will result in an error of 18% in T. This variation is usually inconsiderable in practice. The main reason that makes hydrogeologists interested in T and S values is that they can be used to find hydraulic drawdowns for different pumping flow rates. The logarithmic form of the relationships of groundwater has led the scientific community to target to the order of magnitude of the T and S values. On the other hand, 16% flow variation (5/30) is a large number and should not be considered common. An error of 18% for T should be considered mirror, whereas a 16% miscalculation of Q should be considered major. Schematically, a major miscalculation of Q results in minor miscalculations of T and S, because of the linear relationships between the errors and the non-linear requirements for T and S.

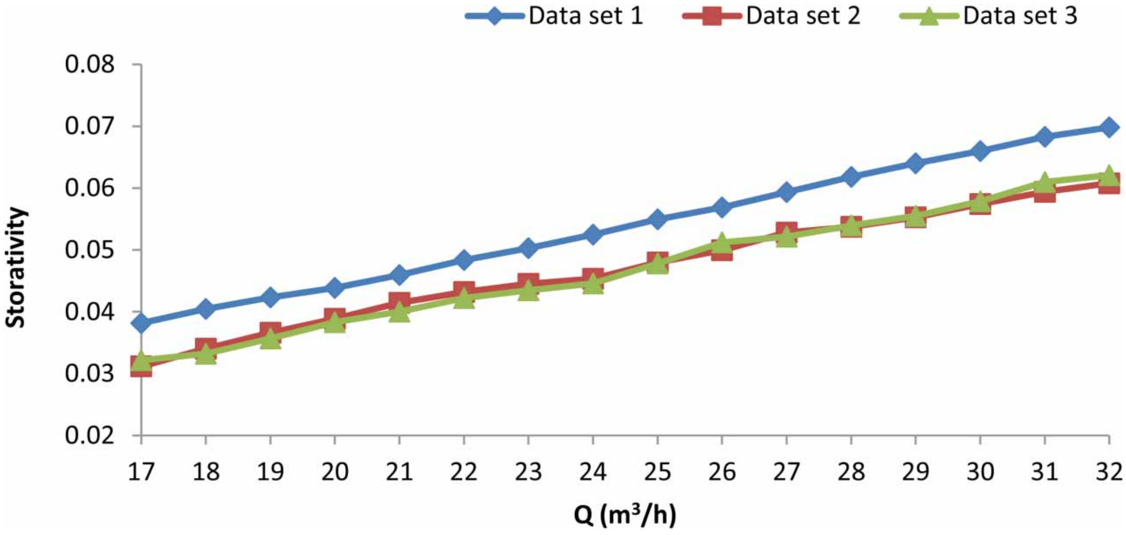

Figure 13 | Variation of flow rates and storativity for all datasets.

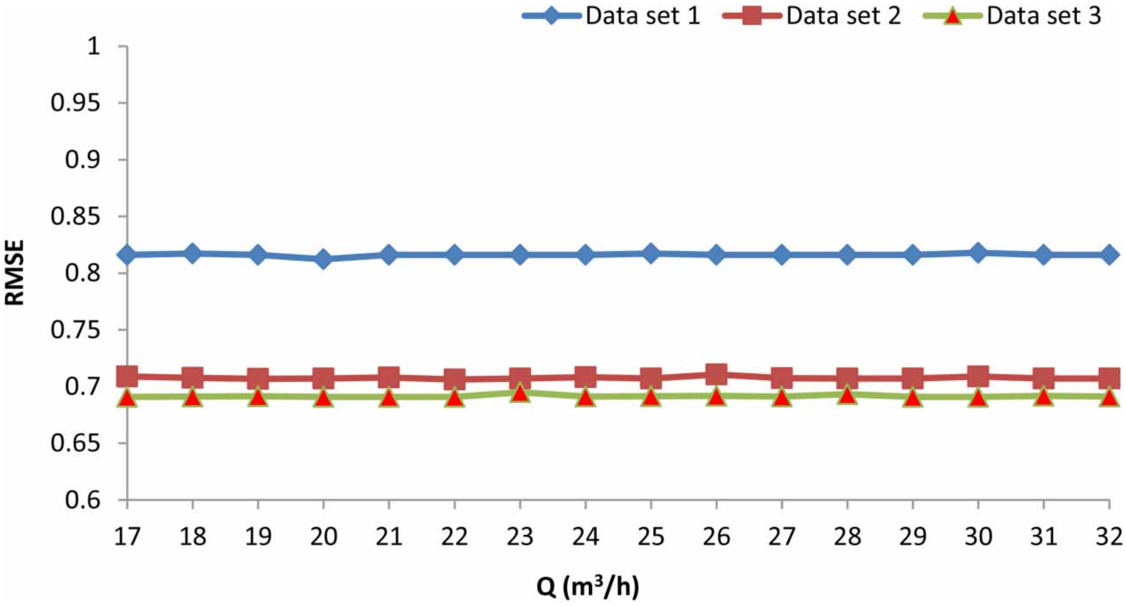

Figure 14 | Variation of flow rates and *RMSE* results for all datasets.

CONCLUSION

In this paper, we investigate the validity of hydrogeological parameters calculation using the Cooper–Jacob method, the *RMSE* and the GA. In particular, we examine how pumping flow rate errors affect the hydrogeological parameters estimation, considering a confined aquifer and using actual hydraulic drawdown measurements. Low values of *RMSE* mean that the calculated values of transmissivity (T) and storativity (S) can accurately reproduce the actual hydraulic drawdown curve. As indicated in the literature review, the GA combined with Theis resulted in lower *RMSE* that the graphical Cooper–Jacob method. This means that GA can be used to accurately reproduce the actual hydraulic drawdown and therefore the results of T and S obtained by GA should be considered more accurate than those obtained from Cooper–Jacob. We have found out that:

- The *RMSE* is not affected by inaccurate pumping flow rate estimations. This means that both Cooper–Jacob and GAs can be successfully used to reproduce the actual hydraulic drawdown, regardless the accuracy of the pumping flow rate value.
- The transmissivity and storativity values decrease as the pumping flow rate decreases, in a linear way. Hence, errors in pumping flow rate estimation should be considered minor by researchers who are interested in finding the order of magnitude of an aquifer's T and S values.

Overall, we have conducted 48 GAs and Cooper–Jacob tests using three drawdown datasets. The main challenge in terms of future research is to apply this method using more datasets. This way it will be possible to find the characteristics of the linear relationship between Q and T, S errors (slope and constant values). At the same time, it is interesting to use GA to find out how accurately they operate (in terms of solving the inverse problem) when external parameters are included (e.g. simultaneous pumping from a system of wells, geological faults, recharging act). Applying the proposed methodology to a range of aquifers will help in generalizing the results about the pumping flow rates' effect in transmissivity and storativity errors' generation. It is also interesting to test how *RMSE* varies according to the range of values chosen as acceptable (e.g. removing the last drawdowns). The results can be used as a rule of thumb to practitioners, indicating that they should not reject measurements obtained from pumping tests when there are reasonable uncertainties about flow rate estimations.

ACKNOWLEDGEMENTS

The authors are indebted to Prof. N. Theodossiou, professor at the Aristotle University of Thessaloniki, Greece, for providing the field data and to Prof. K.L. Katsifarakis for his insightful comments. The genetic algorithm script used is available upon request.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Alexander, S. C. & Saar, M. O. 2011 Improved characterization of small 'u' for Jacob pumping test analysis methods. *Ground Water* **50** (2), 256–265. https://doi.org/10.1111/j.1745-6584.2011.00839.x.

Ali Md., H., Zaman Md., H., Biswas, P., Islam Md., A. & Karim, N. N. 2022 Estimating hydraulic conductivity, transmissibility and specific yield of aquifer in Barind Area, Bangladesh using pumping test. *European Journal of Environment and Earth Sciences* **3** (4), 90–96. European Open Science Publishing. https://doi.org/10.24018/ejgeo.2022.3.4.308.

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P. & Meyer, G. 2018 Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland. *Journal of Hydroinformatics* **20**, 1227–1246. https://doi.org/10.2166/hydro.2018.002.

Anomohanran, O. & Iserhien-Emekeme, R. E. 2014 Estimation of aquifer parameters in Erho, Nigeria using the Cooper–Jacob evaluation method. *American Journal of Environmental Sciences* **10** (5), 500–508. https://doi.org/10.3844/ajessp.2014.500.508.

Avci, C. B., Şahin, A. U. & Çiftçi, E. 2012 A new method for aquifer system identification and parameter estimation. *Hydrological Processes* **27** (17), 2485–2497. Wiley. https://doi.org/10.1002/hyp.9352.

Bastani, M., Kholghi, M. & Rakhshandehroo, G. R. 2010 Inverse modeling of variable-density groundwater flow in a semi-arid area in Iran using a genetic algorithm. *Hydrogeology Journal* **18**, 1191–1203. https://doi.org/10.1007/s10040-010-0599-8.

Bohling, G. C., Liu, G., Knobbe, S. J., Reboulet, E. C., Hyndman, D. W., Dietrich, P. & Butler Jr., J. J. 2012 Geostatistical analysis of centimeter-scale hydraulic conductivity variations at the MADE site. *Water Resources Research* **48** (2). American Geophysical Union (AGU). https://doi.org/10.1029/2011wr010791.

Boulton, N. S. 1954 Unsteady radial flow to a pumped well allowing for delayed yield from storage. *Int Assoc Sci HydrolPubl* **2**, 472–477.

Boulton, N. S. 1963 Analysis of data from non-equilibrium pumpingtests allowing for delayed yield from storage. *Proceedings of the Institution of Civil Engineers* **26** (3), 469–482. https://doi.org/10.1680/iicep. 1963.10409.

Brindha, K. & Elango, L. 2015 Cross comparison of five popular groundwater pollution vulnerability index approaches. *Journal of Hydrology* **524**, 597–613. https://doi.org/10.1016/j.jhydrol.2015.03.003.

Calvache, M. L., Sánchez-Úbeda, J. P., Duque, C., López-Chicano, M. & De la Torre, B. 2016 Evaluation of analytical methods to study aquifer properties with pumping tests in coastal aquifers with numerical modelling (Motril-Salobreña aquifer). *Water Resources Management* **30** (2), 559–575. https://doi.org/10.1007/s11269-015-1177-6.

Chapuis, R. P. 1992 Using Cooper-Jacob approximation to take account of pumping well pipe storage effects in early drawdown data of a confined aquifer. *Groundwater* **30**, 331–337.

Chattopadhyay, P. B., Vedanti, N. & Singh, V. S. 2015 A conceptual numerical model to simulate aquifer parameters. *Water Resourses Management* **29**, 771–784. https://doi.org/10.1007/s11269-014-0841-6.

Cooper Jr., H. H. & Jacob, C. E. 1946 A generalized graphical method for evaluating formation constants and summarizing well-field history. *Transactions, American Geophysical Union* **27** (4), 526–534. https://doi.org/10.1029/TR027i004p00526.

Demir, M. T., Copty, N. K., Trinchero, P. & Sanchez-Vila, X. 2017 Bayesian estimation of the transmissivity spatial structure from pumping test data. *Advances in Water Resources* **104**, 174–182. Elsevier BV. https://doi.org/10.1016/j.advwatres.2017.03.021.

Derakhshan, S. & Bashiri, M. 2018 Investigation of an efficient shape optimization procedure for centrifugal pump impeller using eagle strategy algorithm and ANN (case study: slurry flow). *Struct Multidisc Optim* **58**, 459–473. https://doi.org/10.1007/s00158-018-1897-3.

Dietrich, P., Butler, J. J. & Faiß, K. 2008 A rapid method for hydraulic profiling in unconsolidated formations. *Ground Water* **46** (2), 323–328. Wiley. https://doi.org/10.1111/j.1745-6584.2007.00377.x.

Erickson, M., Mayer, A. & Horn, J. 2002 Multi-objective optimal design of groundwater remediation systems: application of the niched Pareto genetic algorithm (NPGA). *Advances in Water Resources* **25**, 51–65. https://doi.org/10.1016/S0309-1708(01)00020-3.

Farokhzad, S., Ahmadi, H., Jaefari, A., Asadi, A. A. M. R. & Ranjbar, K. M. 2012 Artificial neural network based classification of faults in centrifugal water pump. *Vibro Engineering* **14** (4), 1734–1744.

Fetter, C. W. 2001 *Applied Hydrogeology*, 4th edn. Prentice-Hall, Upper Saddle River, New Jersey.

Freeze, R. A. & Cherry, J. A. 1979 *Groundwater*. Prentice-Hall, Upper Saddle River, New Jersey.

Gardner, K. K. & Vogel, R. M. 2005 Predicting ground water nitrate concentration from land use. *Ground Water* **43** (3), 343–352. https://doi.org/10.1111/j.1745-6584.2005.0031.x.

Gomo, M. 2019 On the interpretation of multi-well aquifer-pumping tests in confined porous aquifers using the Cooper and Jacob (1946) method. *Sustainable Water Resources Management* **5**, 935–946. https://doi.org/10.1007/s40899-018-0259-z.

Gomo, M. 2020 On the practical application of the Cooper and Jacob distance-drawdown method to analyse aquifer-pumping test data. *Groundwater for Sustainable Development* **11**, 100478. https://doi.org/10.1016/j.gsd.2020.100478.

González, J. A. M., Comte, J.-C., Legchenko, A., Ofterdinger, U. & Healy, D. 2021 Quantification of groundwater storage heterogeneity in weathered/fractured basement rock aquifers using electrical resistivity tomography: sensitivity and uncertainty associated with petrophysical modelling. *Journal of Hydrology* **593**. https://doi.org/10.1016/j.jhydrol.2020.125637.

Gunawardhana, L. N., Al-Harthi, F., Sana, A. & Baawain, M. S. 2021 Analytical and numerical analysis of constant-rate pumping test data considering aquifer boundary effect. *Environmental Earth Sciences* **80** (17). Springer Science and Business Media LLC. https://doi.org/10.1007/s12665-021-09833-x.

Ha, D., Zheng, G., Zhou, H., Zeng, C. & Zhang, H. 2020 Estimation of hydraulic parameters from pumping tests in a multiaquifer system. *Underground Space* **5** (3), 210–222. Elsevier BV. https://doi.org/10.1016/j.undsp.2019.03.006.

Halford, K. J., Weight, W. D. & Schreider, R. P. 2006 Interpretation of transmissivity estimates from single-well pumping aquifer tests. *Ground Water* **3**, 467–471. https://doi.org/10.1111/j.1745-6584.2005.00151.x.

Hamed, M. H., Dara, R. N. & Kirlas, M. C. 2022 Groundwater vulnerability assessment using a GIS-based DRASTIC method in Erbil Dumpsite area (Kani Qirzhala), Central Erbil Basin, North Iraq. *Research Square Platform LLC*. https://doi.org/10.21203/rs.3.rs-2074088/v1.

Illman, W. A., Berg, S. J. & Zhao, Z. 2015 Should hydraulic tomography data be interpreted using geostatistical inverse modeling? A laboratory sandbox investigation. *Water Resources Research* **51** (5), 3219–3237. American Geophysical Union (AGU). https://doi.org/10.1002/2014wr016552.

Karpouzos, D. K., Delay, F., Katsifarakis, K. L. & de Marsily, G. 2001 A multi-population genetic algorithm to solve the inverse problem in hydrogeology. *Water Resources Research* **37** (9), 2291–2302. https://doi.org/10.1029/2000WR900411.

Katsifarakis, K. L., Kontos, Y. N., 2020 Genetic algorithms: a mature bio-inspired optimization technique for difficult problems. In: *Nature-Inspired Methods for Metaheuristics Optimization: Algorithms and Applications in Science and Engineering, Modeling and Optimization in Science and Technologies* (Bennis, F. & Bhattacharjya, R. K., eds). Springer International Publishing, Cham, pp. 3–25. https://doi.org/10.1007/978-3-030-26458-1_1.

Kirlas, M. C. 2017 *Hydrogeological Parameters Determination and Investigation of Their Variability in Nea Moudania Aquifer, Greece*. MSc Thesis, School of Civil Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Greece. https://doi.org/10.13140/RG.2.2.20396.16009.

Kirlas, M. C. 2021 Assessment of porous aquifer hydrogeological parameters using automated groundwater level measurements in Greece. *Journal of Groundwater Science and Engineering* **9** (4), 269–278. https://doi.org/10.19637/j.cnki.2305-7068.2021.04.001.

Kirlas, M. C. & Katsifarakis, K. L. 2020 Evaluation of automated groundwater level measurements for transmissivity and storativity calculation. *Journal of Water Supply: Research and Technology – AQUA* **69** (4), 332–344. https://doi.org/10.2166/aqua.2020.100.

Kirlas, M. C., Karpouzos, D. K., Georgiou, P. E. & Katsifarakis, K. L. 2022 A comparative study of groundwater vulnerability methods in a porous aquifer in Greece. *Applied Water Science* **12** (6). Springer Science and Business Media LLC. https://doi.org/10.1007/s13201-022-01651-1.

Kontos, Y. N. & Katsifarakis, K. L. 2017 Optimal management of a theoretical coastal aquifer with combined pollution and salinization problems, using genetic algorithms. *Renew. Energy Energy Storage Syst.* **136**, 32–44. https://doi.org/10.1016/j.energy.2016.10.035.

Kruseman, G. P. & de Ridder, N. A. 2000 Analysis and Evaluation of Pumping Test Data, *second ed. International Institute for Land Reclamation and Improvement*, 2nd ed. Wageningen.

Latinopoulos, P., Theodosiou, N., Papageorgiou, A., Xefteris, A., Fotopoulou, E. & Mallios, Z. 2003 Investigation of water resources in the basin of Moudania, Chalkidiki. In: *Proc. 9th Conference of Hellenic Hydrotechnical Association*, Thessaloniki, pp. 401–408. (in Greek).

Lin, H. T., Tan, Y. C., Chen, C. H., Yu, H. L., Wu, S. C. & Ke, K. Y. 2010 Estimation of effective hydrogeological parameters in heterogeneous and anisotropic aquifers. *Journal of Hydrology* **389**, 57–68. https://doi.org/10.1016/j.jhydrol.2010.05.021.

McKinney, D. C. & Lin, M. D. 1994 Genetic algorithm solution of groundwater management models. *Water Resources Research* **30**, 1897–1906. https://doi.org/10.1029/94WR00554.

Moench, A. F. 1995 Combining the Neuman and Boulton models for flow to a well in an unconfined aquifer. *Groundwater* **33** (3), 378–384. https://doi.org/10.1111/j.1745- 6584.1995.tb00293.x.

Naderi, M. & Gupta, H. V. 2020 On the reliability of variable-rate pumping test results: sensitivity to information content of the recorded data. *Water Resources Research* **56** (5). American Geophysical Union (AGU). https://doi.org/10.1029/2019wr026961.

Nagkoulis, N. 2021 A solution to the groundwater inrverse problem, considering a system of wells. *Indian Water Resources Society Journal* **41** (2). Available from: http://iwrs.org.in/journal/apr2021/1apr.pdf

Ncibi, K., Chaar, H., Hadji, R., Baccari, N., Sebei, A., Khelifi, F. & Hamed, Y. 2020 A GIS-based statistical model for assessing groundwater susceptibility index in shallow aquifer in Central Tunisia (Sidi Bouzid basin). *Arabian Journal of Geosciences* **13** (2). https://doi.org/10.1007/s12517-020-5112-7.

Neuman, S. P. 1972 Theory of flow in unconfined aquifers considering delayed response of the water table. *Water Resources Research* **8** (4), 1031–1045.

Osiensky, J. L., Williams, R. E., Williams, B. & Johnson, G. 2000 Evaluation of drawdown curves derived from multiple well aquifer tests in heterogeneous environments. *Mine Water and the Environment* **19**, 30–55. https://doi.org/10.1007/BF02687263.

Panteli, N. M. & Theodossiou, N. 2016 Analysis of groundwater level measurements – application in the Moudania aquifer in Greece. *European Water* **55**, 79–89.

Pfannkuch, H. O., Mooers, H. D., Siegel, D. I., Quinn, J. J., Rosenberry, D. O. & Alexander, S. C. 2021 Review: 'Jacob's Zoo'–how using Jacob's method for aquifer testing leads to more intuitive understanding of aquifer characteristics. *Hydrogeology Journal* **29**, 2001–2015. https://doi.org/10.1007/s10040-021-02363-7.

Prickett, T. A. 1965 Type-curve solution to aquifer tests under water table conditions. *Groundwater* **3** (3), 5–14. https:// doi. org/ 10.1111/ j.1745- 6584. 1965. tb012 14.x.

Rao, B. V. & Prasad, Y. S. 2021 Evaluation of aquifer characteristics using well logging tools and pumping tests in a typical khondalitic aquifer in the Northern parts of Eastern Ghats of India. *Environmental Earth Sciences* **80**, 38. https://doi.org/10.1007/s12665-020-09286-8.

Rauch, W. & Harremoës, P. 1999 Genetic algorithms in real time control applied to minimize transient pollution from urban wastewater systems. *Water Research* **33**, 1265–1277. https://doi.org/10.1016/S0043-1354(98)00304-2.

Renard, P., Glenz, D. & Mejias, M. 2009 Understanding diagnostic plots for well-test interpretation. *Hydrogeology Journal* **17** (3), 589–600. https://doi.org/10.1007/s10040-008-0392-0.

Saidi, S., Bouri, S., Ben Dhia, H. & Anselme, B. 2011 Assessment of groundwater risk using intrinsic vulnerability and hazard mapping: application to Souassi aquifer, Tunisian Sahel. *Agricultural Water Management* **98** (10), 1671–1682. https://doi.org/10.1016/j.agwat.2011.06.005.

Sanchez-Vila, X. & Fernàndez-Garcia, D. 2016 Debates-Stochastic subsurface hydrology from theory to practice: why stochastic modeling has not yet permeated into practitioners? *Water Resources Research* **52** (12), 9246–9258. American Geophysical Union (AGU). https://doi.org/10.1002/2016wr019302.

Schwartz, F. W. & Zhang, H. 2003 *Fundamentals of Ground Water*. John Wiley & Sons, New York.

Scrucca, L. 2013 GA: A Package for Genetic Algorithms in R. J. Stat. Softw.

Seyedpour, S. M., Valizadeh, I., Kirmizakis, P., Doherty, R. & Ricken, T. 2021 Optimization of the groundwater remediation process using a coupled genetic algorithm-finite difference method. *Water* **13**. https://doi.org/10.3390/w13030383.

Siarkos, I. & Latinopoulos, P. 2016 Modeling seawater intrusion in overexploited aquifers in the absence of sufficient data: application to the aquifer of Nea Moudania, Northern Greece. *Hydrogeology Journal* **24** (8), 2123–2141. doi: 10.1007/s10040-016-1455-2.

Smaoui, H., Zouhri, L., Kaidi, S. & Carlier, E. 2018 Combination of FEM and CMA-ES algorithm for transmissivity identification in aquifer systems. *Hydrological Processes* **32** (2), 264–277. Wiley. https://doi.org/10.1002/hyp.11412.

Spane Jr., F. A. 1993 *Selected Hydraulic Test Analysis Techniques for Constant-Rate Discharge Tests. PNL-8539*. Pacific Northwest Laboratory, Washington.

Sterrett, R. J. 2007 *Groundwater and Wells: A Comprehensive Guide for the Design, Installation and Maintenance of Water Well. Johnson Screens, 3rd ed.* Johnson Screens/A Weatherford Company, New Brighton.

Svigkas, N., Loupasakis, C., Papoutsis, I., Kontoes, C. (Haris), Alatza, S., Tzampoglou, P., Tolomei, C. & Spachos, T. 2020 InSAR campaign reveals ongoing displacement trends at high impact sites of Thessaloniki and Chalkidiki, Greece. *Remote Sensing* **12** (15), 2396. MDPI AG. https://doi.org/10.3390/rs12152396.

Syridis, G. 1990 *Lithostromatographical, Biostromatographical and Paleostromatographical Study of Neogene-Quaternary Formation of Chalkidiki Peninsula (in Greek)*. PhD Thesis, School of Geology, Aristotle University of Thessaloniki, Thessaloniki, Greece.

Tabari, M. M. R., Azari, T. & Dehghan, V. 2021 A supervised committee neural network for the determination of aquifer parameters: a case study of Katasbes aquifer in Shiraz plain, Iran. *Soft Computing* **25** (6), 4785–4798. Springer Science and Business Media LLC. https://doi.org/10.1007/s00500-020-05487-2.

Theis, C. V. 1935 The relation between lowering of the piezometric surface and the rate and duration of discharge of a well using ground water storage. *Transactions, American Geophysical Union* **16**, 519–524.

Thomas, A., Majumdar, P., Eldho, T. I. & Rastogi, A. K. 2018 Simulation optimization model for aquifer parameter estimation using coupled meshfree point collocation method and cat swarm optimization. *Engineering Analysis with Boundary Elements* **91**, 60–72. Elsevier BV. https://doi.org/10.1016/j.enganabound.2018.03.004.

Todd, D. K. & Mays, L. W. 2005 *Groundwater Hydrology*, 3rd edn. John Wiley and Sons, New York, pp. 163–164.

Trabucchi, M., Carrera, J. & Fernàndez-Garcia, D. 2018 Generalizing Agarwal's method for the interpretation of recovery tests under non-ideal conditions. *Water Resources Research* **54** (9), 6393–6407. American Geophysical Union (AGU). https://doi.org/10.1029/2018wr022684.

Tumlinson, L. G., Osiensky, J. L. & Fairley, J. P. 2006 Numerical evaluation of pumping well transmissivity estimates in laterally heterogeneous formations. *Hydrogeology Journal* **14** (1), 21–30. https://doi.org/10.1007/s10040-004-0386-5.

Yeh, W. W.-G. 2015 Review: optimization methods for groundwater modeling and management. *Hydrogeology Journal* **23** (6), 1051–1065. Springer Science and Business Media LLC. https://doi.org/10.1007/s10040-015-1260-3.

Yin, D. & Illman, W. A. 2009 Hydraulic tomography using temporal moments of drawdown recovery data: a laboratory sandbox study. *Water Resources Research* **45** (1). American Geophysical Union (AGU). https://doi.org/10.1029/2007wr006623.

First received 8 May 2022; accepted in revised form 16 March 2023. Available online 28 March 2023