Article

Geospatial Analysis of Transmissivity and Uncertainty in a Semi-Arid Karst Region

Thiago dos Santos Gonçalves [ID], Harald Klammler *[ID] and Luíz Rogério Bastos Leal

Institute of Geosciences, Federal University of Bahia, Salvador 40170, Brazil; t.gon@outlook.com (T.d.S.G.); lrogerio@ufba.br (L.R.B.L.)
* Correspondence: hklammler@ufba.br

Abstract: Aquifer properties, such as hydraulic transmissivity T and its spatial variability, are fundamental for sustainable groundwater exploitation in arid regions. Especially in karst aquifers, spatial variability can be considerable, and the application of geostatistical methods allows for spatial interpolation and mapping based on observations combined with the quantification of uncertainties. Moreover, direct measurements of T are typically scarce, while those of specific capacity S_c are more frequent. In this study, we establish the linear regression relationship between the logarithms of T and S_c measured in 51 wells in a semi-arid karst region in Northeastern Brazil. This relationship is used to estimate empirical values $\log T_{emp}$ based on measurements of $\log S_c$ at 269 wells. $\log T_{emp}$ values are found to be normally distributed with an isotropic variogram of a significant nugget effect (attributed to local-scale karst features) and approximately 10 km range (attributed to larger-scale gradual changes in karst feature density). Ordinary kriging cross-validation indicates an optimum number of 25 neighboring wells for interpolation, which is used in a conditional sequential Gaussian simulation (SGSIM) to generate 500 realizations of $\log T_{emp}$ with respective maps of standard deviations and probabilities of (not) exceeding threshold values. High-transmissivity areas mostly coincide with karstified river valleys, while low-transmissivity areas occur toward the edges where aquifer thickness decreases. The resulting transmissivity maps are relevant for optimizing regional water management strategies, which includes stochastic approaches where transmissivity realizations can be used to parameterize multiple runs of numerical groundwater models.

Keywords: groundwater; aquifer; specific capacity; variogram; kriging; stochastic simulation

1. Introduction

Hydraulic transmissivity T expresses an aquifer's capacity to transmit water and is frequently used for describing the groundwater exploitation potential. It is considered an important parameter for designing groundwater management models, modeling groundwater flow and contaminant transport, as well as selecting areas for well drilling and artificial recharge projects [1–3]. Values of T are generally obtained with long-term pumping tests that last a few hours or days. In contrast, tests to measure the specific capacity S_c of a well require a shorter time to be carried out, where S_c represents the pumping rate per unit of water table drawdown in a well at the end of a pumping period.

Karst aquifers typically have complex groundwater flow patterns as a result of depositional heterogeneities, fracturing, and post-lithification karstification [4,5]. Therefore, it is not practical to perform a sufficient number of pump tests for characterizing the spatial variability of T, and simpler empirical models are generally resorted to. Most observations under heterogeneous conditions indicate that T values calculated from S_c data are consistent with those measured throughout pumping tests. Hence, empirical (log–log) regression relationships for determining T from S_c in fractured karst aquifers have been developed and applied to estimate T at pumped wells without requiring more cumbersome pump tests [6–8].

Geostatistical kriging methods have been applied for spatial interpolation of T between available data points of S_c at such wells [1,7–10]. The typical premise in kriging is that close-by samples are more similar to each other than those further apart (i.e., spatially decreasing correlation as expressed by a variogram). One advantage of kriging is the calculation of expected mean square errors for interpolated data (kriging variance). This interpolation error is small near the observation points and increases in sparsely sampled regions. However, the smoothing effect of kriging and the data-independence of the kriging variance, which only relies on the variogram, are known limiting aspects. Therefore, the kriging variance (or standard deviation) is not a reliable measure of estimation uncertainty, although it may be useful to define the location of future observation points (e.g., monitoring network optimization).

Stochastic simulation is a more sophisticated geostatistical method and commonly used to produce realistic images (i.e., actual realizations and not mean estimates as kriging) of a process from a dataset based on known data points, their univariate distribution (histogram), and a variogram. A sufficient number of realizations allow for the calculation of the best estimates at each location (e.g., simulation-based expectations) along with their respective uncertainty measurements (e.g., simulation-based and, hence, data-dependent standard deviations, confidence intervals, exceedance probabilities, or entire probability distributions). A stochastic simulation has to consist of a sufficient number of realizations to explore the complete uncertainty space of a parameter, and those realizations may be used to propagate such uncertainty into subsequent results (e.g., numerically simulated groundwater levels) by a Monte Carlo approach. Despite these advantages, the application of such stochastic simulation approaches has so far been very limited in the context of spatially variable hydraulic transmissivity [11] or other hydrogeological parameters [1,12] in karst aquifers.

Our study area is the Salitre Karst Aquifer (SKA) located in the central region of Bahia, Brazil. It has been the subject of different hydrogeological, hydrochemical, and contamination vulnerability studies over the last 30 years [13–15], including preliminary studies correlating hydraulic parameters, such as T and S_c [4,9,16]. The SKA is located in a region with stunning landscapes and cultural diversity called Chapada Diamantina, where nature-based tourism is an important economic activity. The network of galleries and caves in the region make up one of the most relevant speleological sites in Brazil. This area is possibly the place with the highest density of underground galleries per unit area in Brazil with an enormous potential for speleological tourism [17,18], in addition to hosting important archaeological, paleontological, and underground biodiversity in South America [19–23]. As such, it is legally protected, although it lacks management plans and environmental water strategies as well as a first fundamental characterization of its hydraulic transmissivity that includes spatial variability and uncertainty.

The present study aims at performing stochastic simulations to map the transmissivity T of the SKA and its uncertainty. Values of specific capacity S_c are used to estimate T from an empirical relationship, while ordinary kriging and a conditional stochastic simulation are used to produce best linear unbiased estimates ("BLUE") and multiple realizations of transmissivity maps, respectively. The mapping of the spatial variability and uncertainty of T in the SKA will contribute to a better understanding of groundwater flow patterns and location of areas with the greatest potential for well drilling and groundwater exploitation. It will also help in artificial recharge projects and numerical modeling and for the sustainable management of speleological tourism and water resources in the region.

2. Materials and Methods

2.1. Study Area and Hydrogeological Conditions

The SKA is located in the central region of the Chapada Diamantina, Bahia, Brazil (Figure 1a). It is an unconfined aquifer with fracture and karst porosity, which is hosted by an extensive plateau of Neoproterozoic carbonate rocks that can reach up to 900 m in thickness [24,25]. The region has a tropical rainy and hot semi-arid climate, with an

average annual temperature of 27.8 °C (varying between 18 °C and 38 °C) and an average rainfall of 846 mm per year. Chapada Diamantina is a "hotspot" of the world's tropical biodiversity and extreme heat during the summer, which greatly increases the regional fire hazard [26,27].
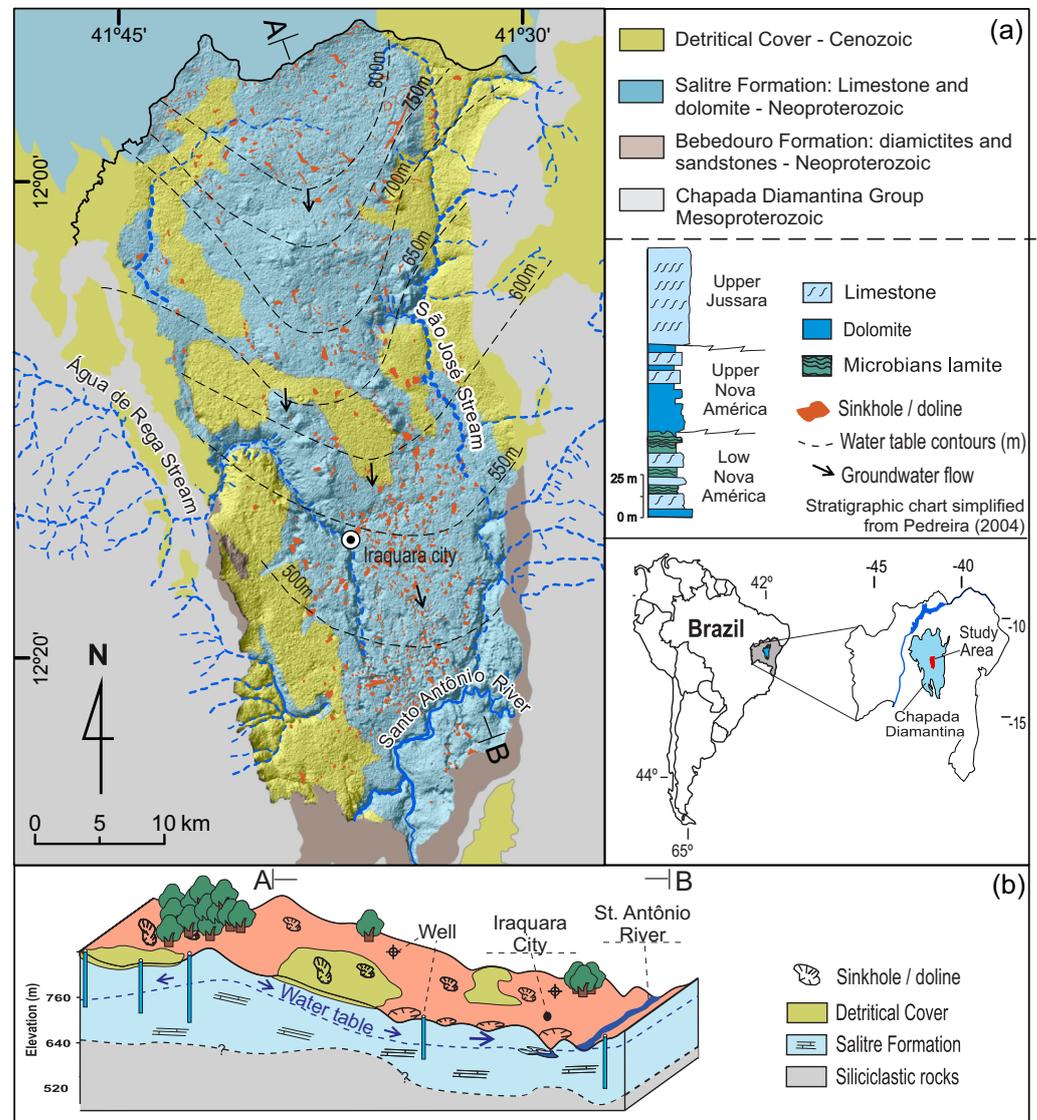
Figure 1. (**a**) Location, schematic geological map of the central region of Chapada Diamantina and detailed stratigraphy for the Salitre Formation. (**b**) Detailed hydrogeological conditions for the Salitre Karst Aquifer (SKA). Modified from [15,28].

The SKA is part of the Rio Santo Antonio watershed at approximately 800 m above sea level. It is surrounded by a mountainous region known as the Chapada Diamantina Group, which, in turn, is at an average altitude of approximately 1500 m [28]. Geologically, the study area consists of a sequence of Precambrian (Neoproterozoic) limestone and dolostone layers of the Salitre Formation, which is further divided into Nova América and Jussara units (Figure 1a). This carbonate sequence overlies the Precambrian terrigenous clastic rocks of the Bebedouro Formation and the Chapada Diamantina Group [29]. The sequence of carbonate rocks is partially covered by a thin layer of tertiary–quaternary detrital sediments.

The Salitre Formation has a complex history with successive low-grade deformation and metamorphism events that gave rise to a system of faults, fractures, and folds. Those are related to two phases of deformation linked to Neoproterozoic compressive

tectonics that led to the formation of mobile belts along the margins of the São Francisco Craton [29–32]. In the study area, the carbonate rocks exhibit a pattern of sub-horizontal layers and E–W- and N–S-trending gentle folds, as well as N–S- and NNE–SSW-striking fracture and vertical joint systems [32–34]. Anticlines and fractures favored the formation of cave systems and paleokarst tunnels exposed above the water table that are widely used by the population for speleological visits [35,36]. The intense deformation and metamorphism undergone by rocks of the Salitre Formation highly increased the aquifer heterogeneity affecting the dissolution of carbonate rocks. This led to the development of a network with multiple conduits and channels for groundwater storage and flow, which makes the process of mapping zones with great potential for well drilling and groundwater capture even more complex [16,24].

The surface geomorphological features include dolines (Figure 1b) and karst depressions known as collapse and suffosion sinkholes [16,37]. The annual recharge of the Salitre aquifer, as calculated on the basis of a hydrometeorological water balance, varies between 55 and 65 mm/year, which is approximately 7% of precipitation. Infiltration events are concentrated in the rainy season (November–March) [38,39]. Aquifer recharge is divided into autogenic or allogenic. Autogenic recharge comes solely from rainfall that falls directly on karst outcrops [15]. Allogenic recharge is sourced from underground flow through underlying rocks or from surface water derived from the mountainous regions of the Chapada Diamantina Group surrounding the karst plateau.

2.2. Data and Well Characteristics

The wells used in this work are generally for public and private supply, with residential and irrigation use. Two different databases were explored, the first consisting of 51 "long-term" (typically around 12 h) single-well pumping test data provided by the Companhia de Engenharia Hídrica e de Saneamento da Bahia (Water and Sanitation Engineering Company of Bahia State, CERB, Salvador, Brazil; Figure 2a). These tests were used for the estimation of T with the Theis recovery method [40] and of S_c as the ratio of pumping rate and final drawdown. To the best of our knowledge, multi-well pumping test data using additional observation wells for an improved estimation of T are not available for the SKA. The second group of data included specific capacity S_c values measured during 218 short-term pumping tests (often shorter than 10 h). This database is available on the website of the Sistema de Informações de Águas Subterrâneas do Serviço Geológico do Brasil [41] and the Sistema Estadual de Informações Ambientais e Recursos Hídricos [42]. This resulted in a total of 269 measurements of S_c (Figure 2b).

These wells have an average and maximum depth of 123 m and 200 m, respectively. The average pumping rate is approximately 300 m^3/d but can exceed 2100 m^3/d (Table 1). The water levels measured at the wells vary widely in depth from near surface levels in low-lying regions to depths of up to 147 m in higher elevated areas, and correlation with logS_c is insignificant ($R^2 = 2 \times 10^{-4}$ with $p = 0.92$). Since this is an aquifer with dual porosity given the fracture systems and interconnected karst conduits, water entrances (i.e., well intersections with water bearing fractures or conduits) occur at depths greater than 6.5 m and sometimes close to the total depth of a well (Table 1). Individual wells in the dataset possessed between one and six water entrances, but correlation of the number of entrances to logS_c was also insignificant ($R^2 = 3 \times 10^{-3}$ with $p = 0.73$). Generally, wells do not possess a filter at all and exist as open boreholes in the karst rock below a top casing.
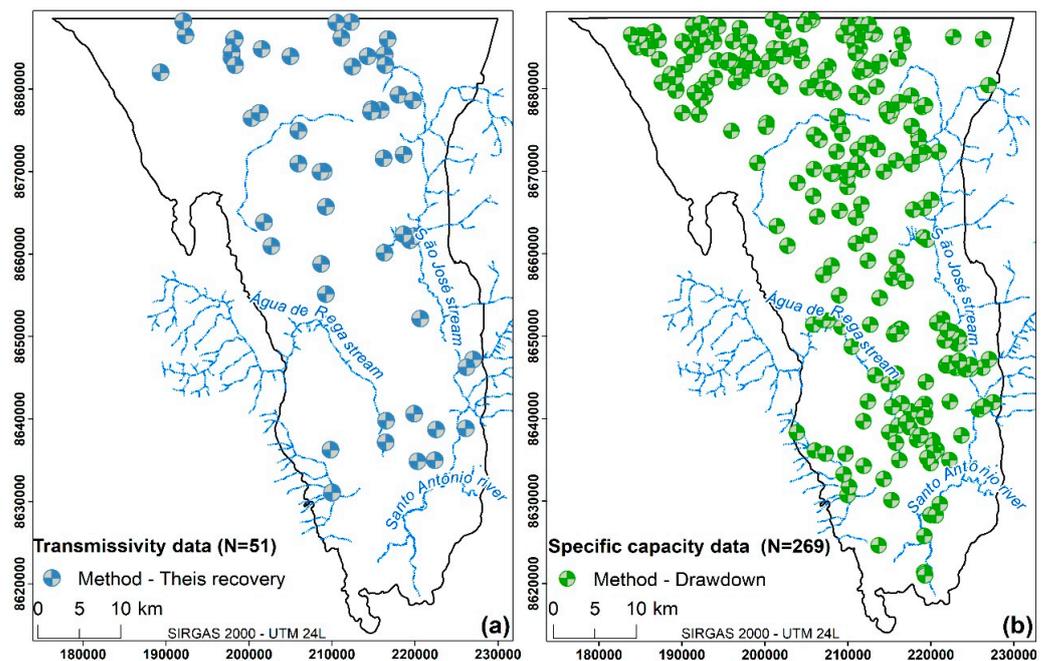
Figure 2. Location of the Salitre aquifer pumping wells used in this study. (**a**) Well data from the CERB database used for the calculation of transmissivity *T* by the recovery method [40] and of specific capacity S_c. (**b**) Specific capacity well data available on all databases.

Table 1. Main characteristics of the wells located in the SKA within the study area. The locations of the wells are shown in Figure 2.

Parameters	Min [1]	Max [2]	Mean	Median	SD [3]	CV [4]
Pumping rate (m^3/d)	5	2112	306	262	247	0.8
Water table depth (m)	0	147	46	40	32	0.7
Water entrance depths (m)	6.5	190	77	73	41	1
Well depth (m)	21	200	123	120	39	0.7

[1] Minimum; [2] Maximum; [3] Standard deviation; [4] Coefficient of variation.

2.3. Linear Regression

Values of transmissivity *T* were calculated using pumping test data according to the Theis recovery method [40]. To establish the linear relationships between the logarithms of specific capacity S_c and empirical estimates T_{emp} of transmissivity, a linear regression model was created for 51 pairs of log*T* and logS_c values (Figure 2a). Theoretical equations for the empirical relationships between *T* and S_c in other aquifers have been established by different authors [3,6,7,9,10]. The test of normality for log*T* and logS_c values was carried out using the Shapiro–Wilk method [43].

2.4. Semivariogram

The semivariogram γ (short "variogram" hereafter) is a curve that describes the degree of spatial continuity as well as the degree of anisotropy in spatial variability of a given dataset *Z* at coordinates x_i. It represents half of the mean square of differences among the *n* pairs of data points in the study area with a distance *h* from each other:

$$\gamma(h) = \frac{1}{2n}\sum\nolimits_{i=1}^{n}(Z(x_i + h) - Z(x_i))^2 \tag{1}$$

In a typical situation, the variance increases with the distance among sampling points up to a maximum level (sill) that corresponds to the total variance [44]. Here, the theoretical

variogram is visually adjusted to the empirical variogram points from Equation (1), giving higher importance to shorter lag distances to optimize for spatial interpolation.

2.5. Ordinary Kriging

Kriging has been widely applied in many environmental studies, with ordinary kriging being the most prominent interpolation method given its simplicity and the quality of results provided [44–47]. It is a point estimation method, where the estimated value Z* at an unsampled point x_0 results from the weighted linear combination of *m* observations $Z(x_i)$ with weights λ_i found in the neighborhood.

$$Z^*(x_0) = \sum_{i=1}^{m} \lambda_i Z(x_i) \tag{2}$$

The kriging weights λ_i for a given location x_0 are determined based on the variogram by minimizing the kriging variance at x_0. The best value of *m* is determined from cross-validation using the root mean square error (RMSE). Here, the numerical grid mesh used for both kriging and stochastic simulations contains 8191 square cells, each with sides of 500 m.

2.6. Conditional Stochastic Simulation

We used the conditional sequential Gaussian simulation [45,46] of $\log T_{emp}$ after confirming normality according to [43]. A total of 60 simulations were carried out to assess a sufficient number of realizations. Each simulation consisted of a growing number from 10 to 600 realizations. For each simulation, the variance was calculated at every grid node using the respective number of realizations, and the spatial mean of these variances was determined. This mean variance increased with the number of realizations but stabilized when the number of realizations was sufficient to represent the whole process variability. We calculate the mean variance increment (MVI) from one simulation to the next and used the proximity of MVI to zero as a criterion to infer the sufficient number of simulations. From the distribution of transmissivity values simulated at each grid location, expectations, standard deviations, and probabilities of (non) exceedance of certain threshold values were computed.

3. Results

3.1. Empirical Relationship between Specific Capacity and Transmissivity

The values of *T* and S_c for the Salitre aquifer calculated with the recovery method [40] are shown in Table 2. S_c varies from 2 to 870 m²/d, while *T* ranges between 1 and 980 m²/d. This wide variation in *T* is a reflection of aquifer heterogeneity due to the complex fracture systems and karst conduits. The high level of aquifer heterogeneity is also well documented in independent field observations [15,16]. The high skewness in *T* and S_c values is also reflected by the differences between the mean and median for each parameter, which are greater than 100% for both.

Table 2. Summary statistics of S_c and *T* (both in m²/d) and their decimal logarithms for *N* = 51 wells (Figure 2a).

Attributes	Min [1]	Max [2]	Mean	Median	SD [3]	CV [4]
S_c	2	871	159	73	202	1.27
$\log S_c$	0.37	2.94	1.78	1.86	0.68	0.38
T	1	980	133	50	209	1.57
$\log T$	−0.05	2.99	1.55	1.71	0.81	0.52

[1] Minimum; [2] Maximum; [3] Standard deviation; [4] Coefficient of variation.

Figure 3a shows the correlation between *T* and S_c with a linear regression fit as well as the univariate distributions represented by the boxplots for each parameter. These boxplots show positive skewness, indicating the predominance of small values close to zero over

few large values above 500 m^2/d. After applying a logarithmic transformation with base ten (log$_{10}$ where the base "10" will be omitted from now on), a symmetrical distribution is achieved and both parameters have a good linear correlation (Figure 3b). The normality tests [43] for log$T$ and logS_c resulted in *p*-values of 0.17 and 0.19, respectively, and therefore, the null hypothesis of normal distributions was not rejected.
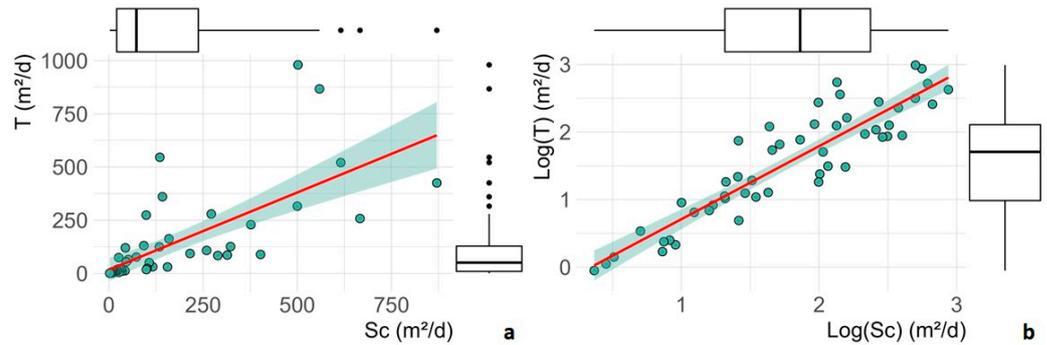
Figure 3. Scatter plots between T and S_c (green circles) with linear regression lines (red) and confidence intervals (green shaded). (**a**) Raw data displaying skewed distributions and (**b**) decimal logarithms with more symmetrical distributions.

The correlation between log$T$ and logS_c (Figure 3b) shows a good linear fit as indicated by the coefficient of determination $R^2 = 0.85$ ($p = 2.2 \times 10^{-16}$) of the linear regression (red line). Such a good correlation between S_c and T for fractured and karst aquifers have also been reported for other regions [3,4,6,7,48]. The functions that express the linear regression model and its antilogarithmic transformation are shown in Equations (3) and (4), respectively.

$$\log(T_{emp}) = -0.37 + 1.08 \log(S_c) \tag{3}$$

$$T_{emp} = 0.42\, S_c^{1.08} \tag{4}$$

where T_{emp} is the empirical transmissivity estimated from the measured S_c values.

3.2. Transmissivity Estimation and Spatial Interpolation

Values of logT_{emp} and T_{emp} calculated from the S_c available for the SKA using Equations (3) and (4) are summarized in Table 3. The results reveal a large difference between the mean and median values of T_{emp} due to a skewed distribution. On the other hand, the differences between the mean and median for logT_{emp} values are smaller, which is also reflected by a *p*-value of 0.17, which does not reject the null hypothesis for normality according to the Shapiro–Wilk test.

Table 3. Summary statistics of S_c and T_{emp} (both in m^2/d) and their decimal logarithms for $N = 269$ wells (Figure 2b).

Attributes	Min [1]	Max [2]	Mean	Median	SD [3]	CV [4]
S_c	0.9	8448	617	158	1106	1.79
logS_c	−0.04	3.93	2.25	2.2	0.74	0.33
T_{emp}	0.4	7314	471	100	912	1.94
logT_{emp}	−0.42	3.86	2.05	2	0.8	0.39

[1] Minimum; [2] Maximum; [3] Standard deviation; [4] Coefficient of variation.

The computation of the experimental logT_{emp} semivariogram took into account the following azimuth directions in degrees: 0, 22.5, 45, 67.5, 90, 112.5, 135, 157.5. The semivariogram shows an exponential isotropic behavior with a range of 10 km and a nugget effect that accounts for approximately 60% of the total variance (Figure 4). The high value of the nugget effect can be explained by the uncertainties of measuring S_c, the transformation into

$\log T_{emp}$ estimates using Equation (3), as well as heterogeneities at the scale of a few tens or hundreds of meters (smaller than the typical separation distance among neighboring wells). These heterogeneities are linked to the complex fracture system and karst conduits in the aquifer, as demonstrated by [13,37]. The structural variability up to 10 km can be attributed to a more gradual variability of the mean intensity of the aquifer heterogeneities.
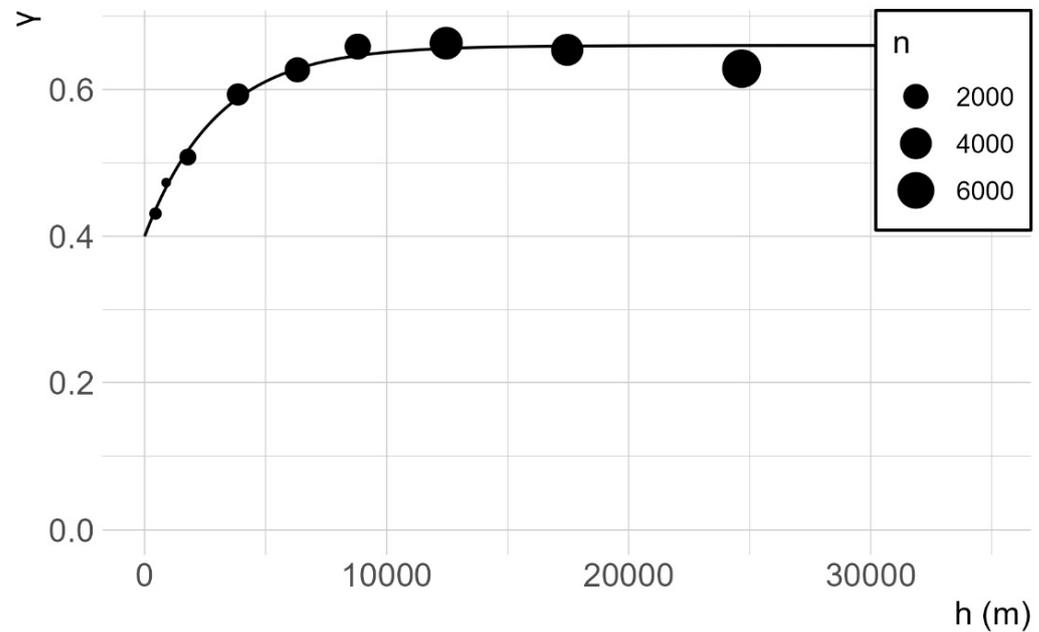
Figure 4. Theoretical semivariogram (solid line) overlying the experimental omnidirectional semivariogram (dots with size indicating the number n of data pairs used) of $\log T_{emp}$ in m^2/d as a function of lag distance h in m.

Cross validation from ordinary kriging indicates that the best neighborhood consists of m = 25 wells with a minimum RMSE of 0.75 for $\log T_{emp}$ when used in units of m^2/d (Figure 5, black line). This neighborhood was found as the minimum number of wells, above which no significant reductions in RMSE occurs. The minimum value of RMSE is relatively large, which is partially due to the significant nugget effect in the variogram, but it is also significantly smaller than the total range of $\log T_{emp}$ of approximately 4.3 encountered in the study are (Table 3).

The ordinary kriging estimates for $\log T_{emp}$ values as a function of the 25 nearest neighbors as well as the theoretical semivariogram from Figure 4 are shown in Figure 6a. The results yield $\log T_{emp}$ values (with T_{emp} in m^2/d) that range between 1.1 and 2.8, displaying the highest values throughout the western boundary and in isolated areas of the central regions of confluences from the surrounding drainage network. These areas are fed by surface water deriving from the metasedimentary rocks of the Chapada Diamantina Group and account for the allogenic recharge of the SKA [35,49]. On the other hand, Figure 6b shows that the highest ordinary kriging standard deviations of $\log T_{emp}$ are at the contact zones between the SKA carbonate rocks and the underlying basement rocks since these areas have a low density of groundwater wells (fewer data points).
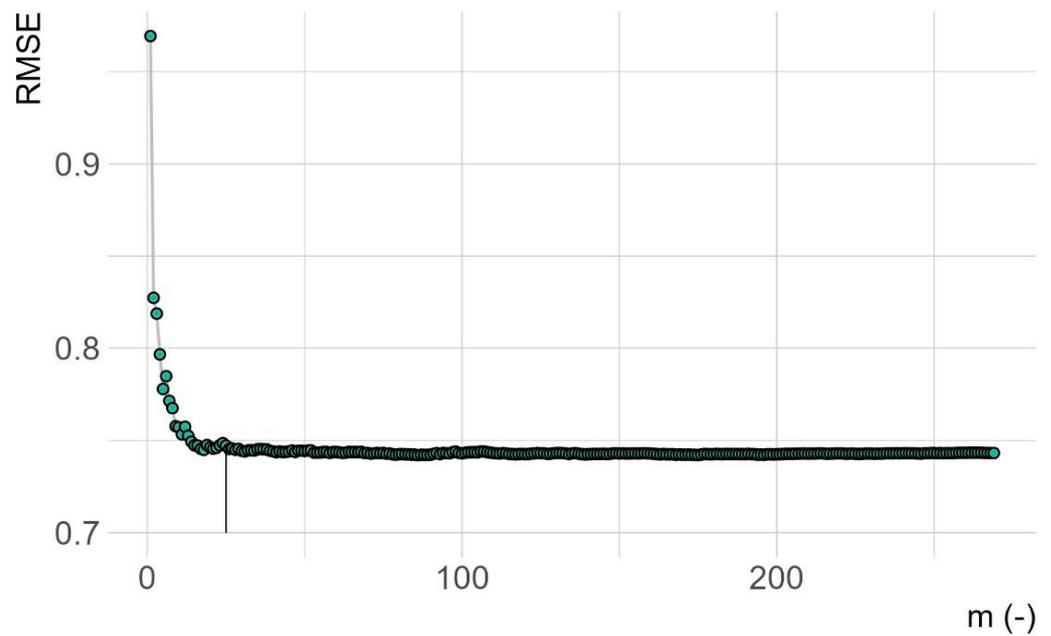
Figure 5. Variation in RMSE from cross-validated $\log T_{emp}$ values in m^2/d as a function of the number *m* of nearest neighbors used in ordinary kriging (vertical black line indicating optimum value chosen of *m* = 25).
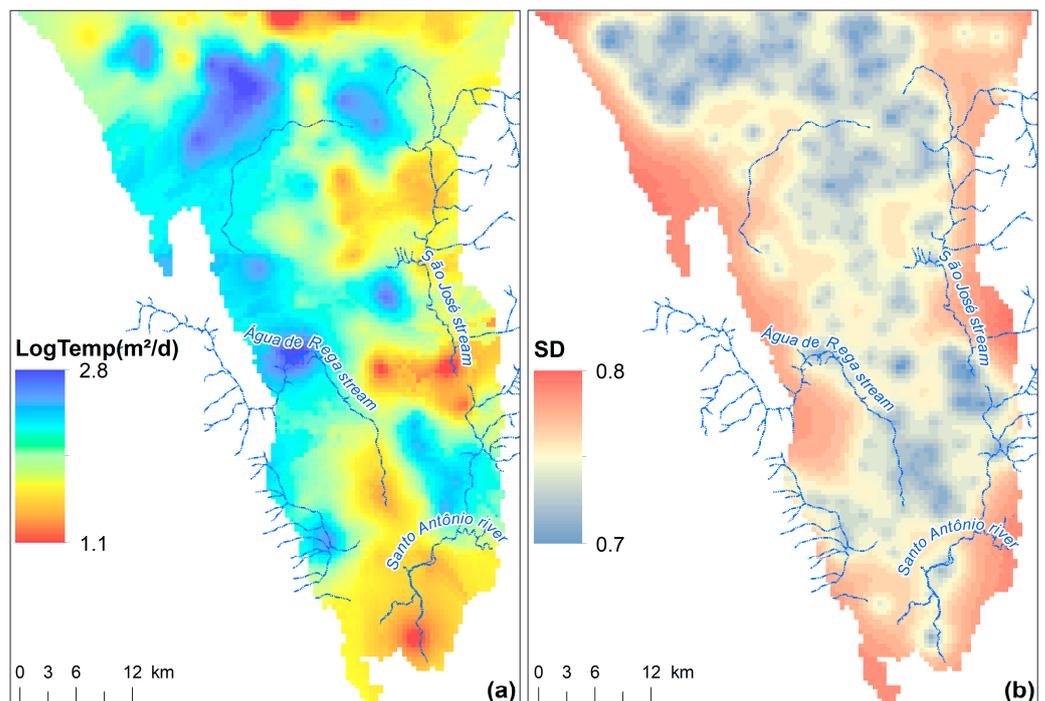
Figure 6. Spatial representation of (**a**) ordinary kriging estimates of $\log T_{emp}$ and (**b**) respective ordinary kriging standard deviations SD.

3.3. Stochastic Simulation of Transmissivity

The results of the stochastic simulations yielded near-zero MVI values for a number of $r \approx 500$ realizations (Figure 7), which is therefore considered a sufficient number of realizations for $\log T_{emp}$ stochastic simulations. The averages of conditional realizations of $\log T_{emp}$ are presented in Figure 8a, while the associated standard deviations are shown in Figure 8b. Given the nature of the methods, the ordinary kriging interpolations (Figure 6a) and the mean of the conditional sequential Gaussian simulation for 500 realizations show

consistent results. Similar to the pattern obtained through ordinary kriging (Figure 6b), areas with lower sampling density also have the highest simulated standard deviations (Figure 8b). However, the range of the simulated standard deviations (0.6 to 0.9) is three times that of the ordinary kriging (0.7 to 0.8), which is based merely on the variogram and not on the data itself, as it is the case in the stochastic simulation.
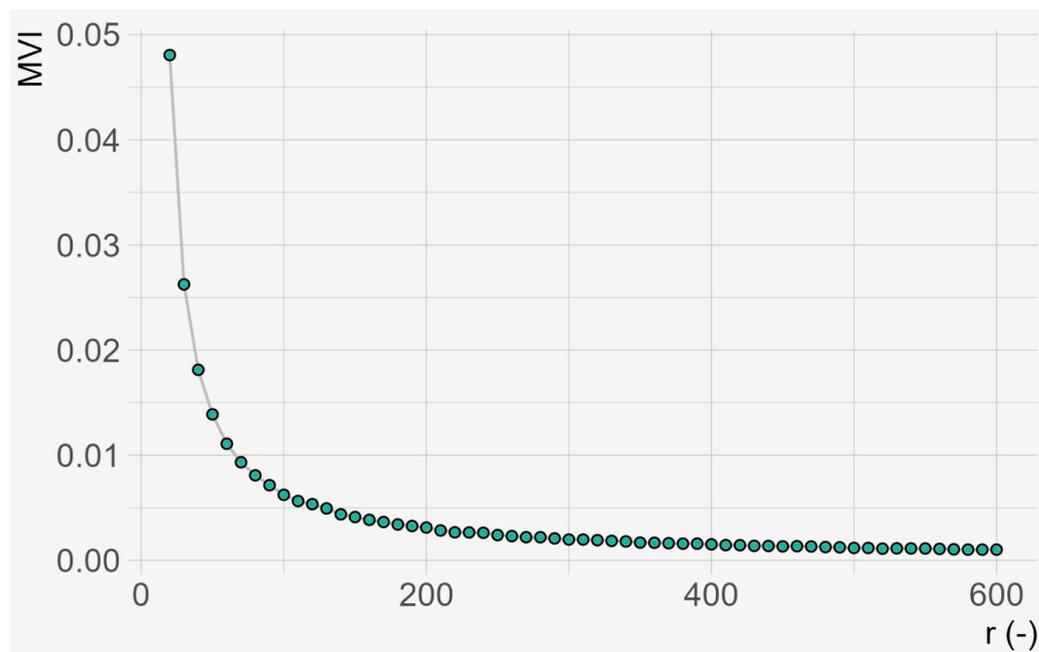
Figure 7. Decrease of the mean variance increment (MVI) with a growing number of realizations *r* used for the stochastic simulation.

Figure 8c shows the probabilistic zoning for $\log T_{emp}$ values smaller than 2 (i.e., $T_{emp} < 100$ m^2/d) and, therefore, reveal the south-central, north-central, and northeast zones of the aquifer as likely low-transmissivity areas. The spatial domains displaying the highest probabilities of large $\log T_{emp}$ values greater than 2.5 (i.e., $T_{emp} > 316$ m^2/d) are shown in Figure 8d. Those areas with higher water potential are associated with convergences in the drainage network receiving allogenic recharge to the aquifer and a larger potential for limestone dissolution (conduit formation). Areas with low probabilities of both $T_{emp} < 100$ m^2/d and $T_{emp} > 316$ m^2/d (shades of green in Figure 8c,d simultaneously) display a high probability of intermediate transmissivity.
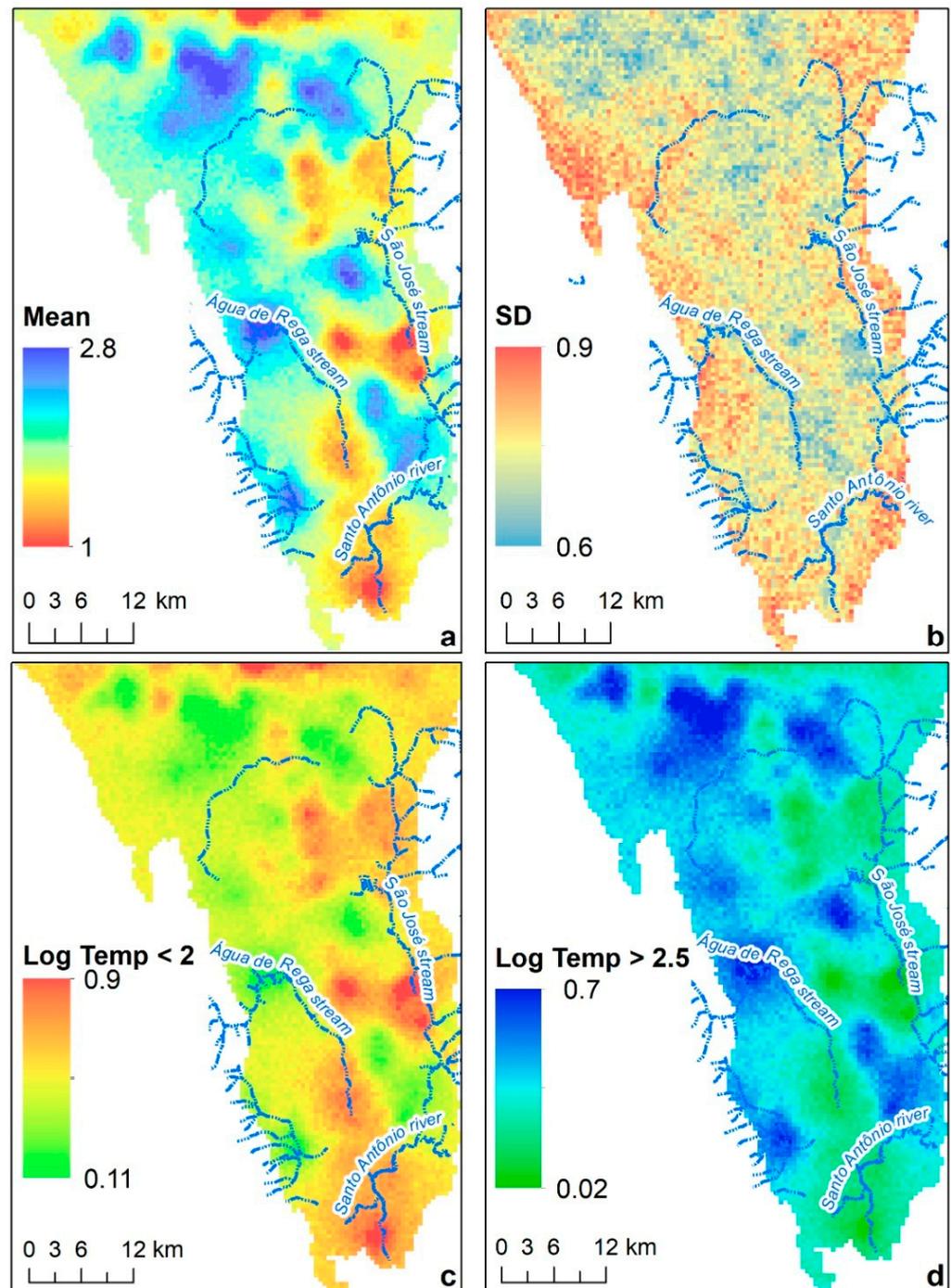
Figure 8. Spatial representation of results from 500 stochastic realizations of $\log T_{emp}$ (with T_{emp} in m^2/d). (**a**) Means of $\log T_{emp}$ at each location, (**b**) standard deviations SD, (**c**,**d**) probability of $\log T_{emp}$ values lower than 2 (i.e., $T_{emp} < 100$ m^2/d) and higher than 2.5 (i.e., $T_{emp} > 316$ m^2/d), respectively.

4. Discussion

4.1. Regression

The regression coefficients in Equation (4) resulted as $a = 0.42$ (multiplier) and $b = 1.08$ (exponent) with $R^2 = 0.85$ (log–log coefficient of determination). While b and R^2 are very similar to that in other studies in karst aquifers, a is relatively low (Table 4). However, its value is still consistent with a previous study in the SKA [16], which included a larger portion of the aquifer but did not attempt to achieve any spatial mapping or uncertainty assessment. Moreover, a difference in a by a factor of two to other aquifers is still much

smaller than the full range of T here, spanning over three orders of magnitude (factor of approximately 1000; Table 4). In fractured rock aquifers, values of a as low as 0.12 have been reported [48,50].

Table 4. Summary of data for S_c and T with regression and variogram results (where available) for related and the present studies in karst aquifers.

Reference	S_c		T		Regression		Variogram	
	N [1] (-)	Min [2]–Max [3] (m²/d)	N (-)	Min [2]–Max [3] (m²/d)	T_{emp} (m²/d)	R^2 (-)	Nugget/Sill (-)	Range (km)
[8]	124	5.75–1141.4	45	8.41–1940	$0.85S_c^{1.07}$	0.95	0.4	0.6
[10]	-	-	422	50–2200	-	-	0.1	40
[7]	71	14.6–12,948	71	1–100,000	$0.76S_c^{1.08}$	0.89	-	-
[50]	14	65.9–47,455	14	100–100,000	$1.23S_c^{1.05}$	0.8	-	-
[16]	1334	0.63–3738	213	0.64–3490	$0.5S_c^{1.05}$	0.84	-	-
This study	269	0.9–8448	51	1–980	$0.42S_c^{1.08}$	0.85	0.6	10

[1] Number of wells; [2] Minimum; [3] Maximum.

To further investigate possible reasons for the small value of a in the SKA, we resort to the classical Theis solution with adjustments for well losses s_l and a partial penetration factor s_p. A respective relationship between well drawdown s_w and T has been established as [48,50]

$$T = \frac{Q}{4\pi(s_w - s_l)}\left[ln\left(\frac{2.25Tt}{r_w^2 S} + 2s_p\right)\right] \tag{5}$$

where S is the storage coefficient and r_w the well radius. Equation (5) can be reformulated into $T = aS_c$, where $S_c = Q/s_w$ and

$$a = \frac{1}{4\pi\left(1 - \frac{s_l}{s_w}\right)}\left[ln\left(\frac{2.25Tt}{r_w^2 S} + 2s_p\right)\right] \tag{6}$$

The multiplier a also grows with T, which is related to the fact that the empirical relationship in Equation (4) becomes non-linear, with an exponent b somewhat larger than one. For some given average transmissivity T, Equation (6) also shows that a decreases with smaller pumping times t used for estimating S_c. Since these pumping times are only around 12 h in our study area (Section 2.2) and the range of T is lower (Table 4), the adjusted value of a also becomes smaller. Intuitively, this compensates for the effect of overestimating S_c when pumping times are short because the "steady-state" drawdown is underestimated. In addition, wells in the SKA are mostly fully penetrating and left as open boreholes (no filter screen) below a top casing (Section 2.2), such that s_l and s_p are close to zero and do not raise the value of a due to the effects of well losses and partial penetration. Finally, larger well radii r_w may also decrease a (to the square in the denominator of the argument of the logarithm), which is also a potential factor in the SKA, where uncased borehole diameters in the strongly karstified rock may be larger than the nominal drilling diameters and highly variable with depth (including intersections with large flow conduits). Equation (6) with values of $T = 50$ m²/d (median in Table 2), $t = 0.5$ d (typical pump test time in SKA), $r_w = 0.5$ m (enlarged borehole cavity), and $S = 0.25$ (specific yield of a highly karstified aquifer) results in a value of $a \approx 0.5$, for example. Note that r_w and/or S may even be larger at intersections with large flow conduits, thus further lowering a.

The regression relationship in Equations (3) and (4) is defined for the data of S_c in the range of 2 to 871 m²/d (standard deviation $SD = 202$ m²/d; Table 2), while the transmissivity estimation based on this regression relationship is over the range of S_c from 0.9 to 8448 m²/d ($SD = 1106$ m²/d; Table 3). Only two values are smaller than 2 m²/d, but approximately 20% of the values for S_c in Table 3 are larger than 871 m²/d. This means

that the regression relationship is used to extrapolate values of T_{emp} within that range, and associated uncertainties are not accounted for in the results. Another type of uncertainty that is not accounted for in our approach is due to the prediction error of the regression model (see shaded confidence interval and dispersion of data point around regression line in Figure 3b). Thus, the uncertainty estimates and probabilities of (not) exceeding certain T thresholds in Figure 8 have to be regarded as lower bounds.

4.2. Variogram and Kriging

The relative nugget effect (60%) and range (10 km) of the variogram in Figure 4 are consistent with previous studies in karst aquifers (Table 4). The combination of regression and ordinary kriging applied here is a simplified approach with respect to cokriging, for example, which could directly consider a scarcely sampled primary variable (here, $\log T$) and one (or several) more abundantly sampled secondary variable(s) (here, $\log S_c$). To build a full cokriging system, variograms of the primary and the secondary variables, as well as a cross-variogram between both variables, need to be inferred from the data and have to meet certain compatibility requirements. In our study area, however, the variogram of $\log T$ from 51 well recuperation tests does not present any spatial correlation at all (pure nugget effect; chart not shown). Moreover, the numbers of data pairs for short lag distances are very small, hence, rendering those variogram points unreliable.

A simplified and less restrictive cokriging approach is collocated cokriging, which does not require a variogram of the secondary variable or a cross-variogram. It only requires a correlation coefficient between primary and secondary variables (as implied here in the regression relation) and a variogram of the primary variable. However, for the present situation, where the variogram of the primary variable does not present any spatial correlation, such a full or collocated cokriging would collapse to the approach actually taken in this study for computing estimates (or "transformed values") for the $\log T_{emp}$ of $\log T$ at locations where $\log S_c$ is known [46] (pp. 309, 339). When further using ordinary kriging for estimating $\log T_{emp}$ at all grid locations based on given values of $\log T_{emp}$ at all wells, this becomes a type of "kriging with transformed values" [50]. The same line of reasoning applies to the conditional stochastic simulation based on values of $\log T_{emp}$ only.

5. Conclusions

This work collected hydraulic transmissivity T and specific capacity S_c data from 51 and 269 wells, respectively, installed throughout the Salitre Karst Aquifer (SKA) in the semi-arid northeast of Brazil. Values of $\log T_{emp}$ were estimated using a linear regression between $\log T$ and $\log S_c$ measurements from the same wells. The univariate distributions (histograms) of $\log T_{emp}$ and $\log S_c$ passed the test of normality. The semivariogram displays an exponential and isotropic behavior with a high nugget effect. This large randomness was expected due to the high aquifer heterogeneity as a consequence of the expressive fracture system with high degree of karstification and also due to errors of single-well test data processing (i.e., a Theis recovery without observation wells). However, there is also structural variability within a range of approximately 10 km. This is likely the result of a more gradual and regional variability of the average fracturing and karstification intensities.

Ordinary kriging indicates an optimal number of 25 closest neighboring wells for interpolation. Using the same neighborhood, a conditional sequential Gaussian simulation was performed with 500 realizations, confirming the best estimates of $\log T_{emp}$ from ordinary kriging and also returning spatial uncertainty estimates in the form of standard deviations as well as the probabilities of exceeding or not exceeding certain transmissivity threshold values. The regions with the largest hydraulic transmissivities were identified in the vicinity of the main rivers in the region. There, the allogenic recharge of the aquifer, which is associated with the drainage network fed by the mountain highs carved on the siliciclastic metasedimentary rocks of the Chapada Diamantina, lead to stronger degrees of karstification. The low-transmissivity regions occur mainly in higher elevated aquifer

zones with a lower degree of karstification as well as along the edges of the karst domain, where the aquifer thickness decreases significantly.

We recognize that the data available are relatively uncertain due to the single-well pump test procedure in addition to the strong short-range spatial variability due to the inherent hydrogeological variability in the karst aquifer. However, to the best of our knowledge, the data used are the only available for the region. Despite these limitations, the variogram still presents 40% of the structural variability (besides the 60% nugget), which manifests in the spatial distribution of high and low transmissivity areas mapped in Figures 6 and 8. Although the kriged and simulated SDs are as high as that of the sample $\log T$ data in some areas (particularly along the aquifer borders, where wells are sparser), the probability maps in Figure 8c,d appear informative by identifying significant portions of the study area with large likelihoods of either high or low transmissivities.

This work produced a first set of spatially distributed transmissivity estimates for the study area that are in general agreement with geological properties and processes known in the region. Although uncertainties are large, this represents an advance in the hydrogeological knowledge of the SKA as well as in the establishment of transmissivity zones for Precambrian karst aquifers. The results may be useful for the parametrization of regional numerical models (e.g., Modflow) in terms of heterogeneous hydraulic transmissivity. Furthermore, these models can be run repetitively for different realizations of transmissivity, which allows for the propagation of uncertainties associated with the final results of a drawdown, for example. Therefore, the results herein obtained can contribute to an improved management of groundwater resources in the SKA, and they can help to establish better policies for groundwater exploitation based on technical criteria. Further data acquisition and modeling are needed, however, to extend the study area over the remaining (northern) part of the SKA and to allow for a direct inclusion of observed transmissivities T from pump tests (preferentially including observation wells to improve T estimates with respect to the single-well method applied so far) besides the empirical estimates T_{emp} in the stochastic simulation.

References

1. Lin, Y.P.; Tan, Y.C.; Rouhani, S. Identifying spatial characteristics of transmissivity using simulated annealing and kriging methods. *Environ. Geol.* **2001**, *41*, 200–208. [CrossRef]
2. Richard, S.K.; Chesnaux, R.; Rouleau, A.; Coupe, R.H. Estimating the reliability of aquifer transmissivity values obtained from specific capacity tests: Examples from the Saguenay-Lac-Saint-Jean aquifers, Canada. *Hydrol. Sci. J.* **2016**, *61*, 173–185. [CrossRef]

3. Valigi, D.; Cambi, C.; Checcucci, R.; Di Matteo, L. Transmissivity Estimates by Specific Capacity Data of Some Fractured Italian Carbonate Aquifers. *Water* **2021**, *13*, 1374. [CrossRef]

4. Bastos Leal, L.R.; Silva, H.P. *Hydrologic Model and Management of the Aquifer-River Sistem on the Verde and Jacaré Rivers—Semi-arid Region of the Bahia State, Brazil*; Final Report No. 002/02; SRH/UFBA: Salvador, Brazil, 2004. (In Portuguese)

5. Ford, D.C.; Williams, P.W. *Karst Hydrogeology and Geomorphology*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2007.

6. El-Naqa, A. Estimation of transmissivity from specific capacity data in fractured carbonate rock aquifer, central Jordan. *Environ. Geol.* **1994**, *23*, 73–80. [CrossRef]

7. Mace, R.E. Determination of transmissivity from specific capacity tests in a karst aquifer. *Groundwater* **1997**, *35*, 738–742. [CrossRef]

8. Fabbri, P.; Piccinini, L. Assessing transmissivity from specific capacity in an alluvial aquifer in the middle Venetian plain (NE Italy). *Water Sci. Technol.* **2013**, *67*, 2000–2008. [CrossRef] [PubMed]

9. Galvão, P.; Halthan, T.; Hirata, R. Transmissivity of aquifer by capture zone method: An application in the Sete Lagoas karst aquifer, MG, Brazil. *Ann. Braz. Acad. Sci.* **2017**, *89*, 91–102. [CrossRef]

10. Al-Murad, M.; Zubari, W.K.; Uddin, S. Geoestatistical characterization of transmissivity: An example of Kuait Aquifers. *Water* **2018**, *10*, 828. [CrossRef]

11. Barbosa, S.; Almeida, J.; Chambel, A. A geoestatiscal methodology to simulate the transmissivity in a highly heterogeneous rock body based on borehole data and pumping tests. *Hydrogeol. J.* **2019**, *27*, 1969–1998. [CrossRef]

12. Lin, Y.P.; Chen, Y.W.; Chang, L.C.; Yeh, M.S.; Huang, G.H.; Petway, J.R. Groundwater Simulations and Uncertainty Analysis Using MODFLOW and Geostatistical Approach with Conditioning Multi-Aquifer Spatial Covariance. *Water* **2017**, *9*, 164. [CrossRef]

13. Guerra, A.M. Processos de Carstificação e Hidrogeologia do Grupo Bambuí na Região de Irecê-Bahia. Ph.D. Thesis, Institute of Geosciences, University of São Paulo, São Paulo, Brazil, 1986. (In Portuguese with English Abstract).

14. Silva, H.M. Geographic Information System of the Karst Aquifer of the Microregion of Irecê, Bahia: Subsidy for the Integrated Management of the Water Resources of the Verde and Jacaré River Basins. Master's Thesis, Institute of Geosciences, Federal University of Bahia, Salvador, Brazil, 2005. (In Portuguese with English Abstract).

15. Salles, L.Q.; Galvão, P.; Bastos-Leal, L.R.; Pereira, R.G.F.A.; Purificação, C.G.C.; Laureano, F.V. Evaluation of susceptibility for terrain collapse and subsidence in karst areas, municipality of Iraquara, Chapada Diamantina (BA), Brazil. *Environ. Earth Sci.* **2018**, *77*, 593. [CrossRef]

16. Gonçalves, T.S.; Leal, L.B. Water potential in the Salitre karst Aquifer in the region of Irecê, Bahia. *Águas Sub.* **2018**, *32*, 191–199, (In Portuguese with English Abstract). [CrossRef]

17. Barbosa, E.P.; Travassos, L.E.P. Caves, stories, history and populartraditions in the semi-desert (sertão) of Bahia, northeastern Brazil. *Acta Carst.* **2008**, *37*, 331–338. Available online: https://ojs.zrc-sazu.si/carsologica/article/view/155/145 (accessed on 26 February 2024).

18. Cruz, F.W., Jr. Geomorphological Aspects and Geospeleology of the Karst of the Iraquara Region, North Center of the Chapada Diamantina. Master's Thesis, Institute of Geosciences, University of São Paulo, São Paulo, Brazil, 1998. (In Portuguese with English Abstract).

19. Auler, A.; Farrant, A.R. A brief introduction to karst and caves in Brazil. *Proc. Univ. Bristol Spelaeol. Soc.* **1996**, *20*, 187–200. Available online: https://www.ubss.org.uk/resources/proceedings/vol20/UBSS_Proc_20_3_187-200.pdf (accessed on 26 February 2024).

20. Trajano, E.; Gallão, J.E.; Bichuette, M.E. Spots of high diversity of troglobites in Brazil: The challenge of measuring subterranean diversity. *Biodivers. Conserv.* **2016**, *25*, 1805–1828. [CrossRef]

21. Silva, M.S.; Ferreira, R.L. The first two hotspots of subterranean biodiversity in South America. *Subt. Biol.* **2016**, *19*, 1–21. [CrossRef]

22. Gallão, J.E.; Bichuette, M.E. Taxonomic distinctness and conservation of a new high biodiversity subterranean area in Brazil. *Ann. Braz. Acad. Sci.* **2015**, *87*, 209–217. [CrossRef]

23. Bichuette, M.E.; Rantin, B.; Hingst-Zaher, E.; Trajano, E. Geometric morphometrics throws light on evolution of the subterranean catfish *Rhamdiopsis krugi* (Teleostei: Siluriformes: Heptapteridae) in eastern Brazil. *Biol. J. Linn. Soc.* **2015**, *114*, 136–151. [CrossRef]

24. Bastos Leal, L.R.; Dutton, A.R.; Luz, J.G.; Barbosa, J.S.F. Hydrogeology and hydrochemistry of a Precambrian karst aquifer in a semi-arid region from Bahia, Brazil. In Proceedings of the Geological Society of America, Annual Meeting 2006, Philadelphia, PA, USA, 22–25 October 2006; Volume 38, p. 286.

25. Feitosa, F.A.C.; Diniz, J.A.O.; Kirchheim, R.E.; Kiang, C.H.; Feitosa, E.C. Assessment of Groundwater Resources in Brazil: Current Status of Knowledge. In *Groundwater Assessment, Modeling, and Management*; CRC Press: Boca Raton, FL, USA, 2016. Available online: https://www.routledgehandbooks.com/doi/10.1201/9781315369044-5 (accessed on 5 August 2023).

26. Funch, R.R.; Harley, R.M. Reconfiguring the boundaries of the Chapada Diamantina National Park (Brazil) using ecological criteria in the context of a human-dominated landscape. *Land. Urban Plan.* **2007**, *83*, 355–362. [CrossRef]

27. Ribeiro, R.R.R.; Sulaiman, S.N.; Sieber, S.; Trejo-Rangel, M.A.; Campos, J.F. Integrated assessment of drought impacts on rural areas: The case of Chapada Diamantina Region in Brazil. *GeoHazards* **2021**, *2*, 442–453. [CrossRef]

28. Pedreira, A.J. The Espinhaço Supergroup in the Central-Eastern Chapada Diamantina, Bahia: Sedimentology, Stratigraphy and Tectonics. Ph.D. Thesis, Institute of Geosciences, University of São Paulo, São Paulo, Brazil, 2004. (In Portuguese with English Abstract)

29. Misi, A.; Veizer, J. Neoproterozoic carbonate sequences of the Una Group, Irecê Basin, Brazil: Chemostratigraphy, age and correlations. *Precamb. Res.* **1998**, *89*, 87–100. [CrossRef]

30. Misi, A.; Kaufman, A.J.; Azmy, K.; Dardenne, M.A.; Sial, A.N.; de Oliveira, T.F. Neoproterozoic successions of the Sao Francisco craton, Brazil: The Bambuí, Una, Vazante and Vaza Barris/Miaba groups and their glaciogenic deposits. *Geol. Soc. Lond. Memoir.* **2011**, *36*, 509–522. [CrossRef]

31. Brito Neves, B.B.; Fuck, R.A.; Martins, M. The Brasiliano collage in South America: A review. *Braz. J. Geol.* **2014**, *4*, 493–518. [CrossRef]

32. Kuchenbecker, M.; Reis, H.L.S.; Fragoso, D.G.C. Caracterização estrutural e considerações sobre a evolução tectônica da Formação Salitre na porção central da Bacia de Irecê, norte do Cráton do São Francisco (BA). *Geonomos* **2011**, *19*, 42–49. [CrossRef]

33. Furtado, C.P.Q.; Medeiros, W.E.; Borges, S.V.; Lopes, J.A.G.; Bezerra, F.H.R.; Lima-Filho, F.P.; Maia, R.P.; Bertotti, G.; Auler, A.S.; Teixeira, W.L.E. The influence of subseismic-scale fracture interconnectivity on fluid flow in fracture corridors of the Brejões carbonate karst system, Brazil. *Mar. Petrol. Geol.* **2022**, *141*, 105689. [CrossRef]

34. Furtado, C.P.Q.; Borges, S.V.F.; Bezerra, F.H.R.; Castro, D.L.; Maia, R.P.; Teixeira, W.L.E.; Souza, A.M.; Auler, A.S.; Lima-Filho, F.P. The fracture-controlled carbonate Brejões Karst System mapped with UAV, LiDAR, and electroresistivity in the Irecê Basin—Brazil. *J. South Amer. Earth Sci.* **2022**, *119*, 103986. [CrossRef]

35. Auler, A.S.; Smart, P.L. The influence of bedrock-derived acidity in the development of surface and underground karst: Evidence from the Precambrian carbonates of semi-arid northeastern Brazil. *Earth Surf. Process. Landforms.* **2003**, *28*, 157–168. [CrossRef]

36. Ennes-Silva, R.A.; Bezerra, F.H.R.; Nogueira, F.C.C.; Fabrizio Balsamo, F.; Klimchouk, A.; Cazarin, C.L.; Auler, A.S. Superposed folding and associated fracturing influence hypogene karst development in Neoproterozoic carbonates, São Francisco Craton, Brazil. *Tectonophysics* **2016**, *666*, 244–259. [CrossRef]

37. Salles, L.Q.; Bastos-Leal, L.R.; Pereira, R.G.F.A.; Laureano, F.V.; Gonçalves, T.S. Influence of hydrogeological aspects of karst aquifers on evolution of landscape in the Chapada Diamantina, Bahia, Brazil. *Rer. Bras. Geomorf.* **2018**, *19*, 93–106, (In Portuguese with English Abstract). [CrossRef]

38. Beraldo, V.J. Environmental Isotopic and Hydrochemical Study of the Surface and Groundwater of the Irecê Region, Bahia. Master's Thesis, Institute of Geosciences, Federal University of Bahia, Salvador, Brazil, 2005. (In Portuguese with English Abstract).

39. Ramos, S.O.; Araújo, H.A.; Basos Leal, L.R.; Luz, J.A.G.; Dutton, A.R. Temporal variation of the water table of the karst aquifer of Irecê—Bahia: Contribution to the use and management of groundwater in the semi-arid region. *Braz. J. Geol.* **2007**, *37*, 227–233. [CrossRef]

40. Theis, C.V. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. Eos Trans. *Am. Geoph. Union* **1935**, *16*, 519–524. [CrossRef]

41. SIAGAS. Sistema de Informação de Águas Subterrâneas, Home Page. Available online: http://siagasweb.cprm.gov.br/layout/visualizar_mapa.php (accessed on 17 July 2023).

42. Sistema Nacional de Informações sobre Saneamento, Ministério do Desenvolvimento Regional, Brasil. *Diagnóstico dos Serviços de Água e Esgoto—2023*; Sistema Nacional de Informações sobre Saneamento: Brasília, Brasil, 2023; ISBN 9788578110796.

43. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]

44. Isaaks, E.H.; Srivastava, R.M. *Applied Geostatistics*; Oxford University Press: New York, NY, USA, 1989.

45. Theodossiou, N.; Latinopoulos, P. Evaluation and optimization of groundwater observation networks using the Kriging methodology. *Environ. Model Softw.* **2006**, *21*, 991–1000. [CrossRef]

46. Chiles, J.P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*; John Wiley & Sons: New York, NY, USA, 1999.

47. Rossi, M.E.; Deutsch, C.V. *Mineral Resource Estimation*; Springer: Dordrecht, The Netherlands; Berlin/Heidelberg, Germany; New York, NY, USA; London, UK, 2014; ISBN 978-1-4020-5717-5. [CrossRef]

48. Huntley, D.; Nommensen, R.; Steffey, D. The use of specific capacity to assess transmissivity in fractured-rock aquifers. *Ground Water* **1992**, *30*, 396–402. [CrossRef]

49. Laureano, F.V.; Karmann, I.; Granger, D.E.; Auler, A.S.; Almeida, R.P.; Cruz, F.W.; Strícks, N.M.; Novello, V.F. Two million years of river and cave aggradation in NE Brazil: Implications for speleogenesis and landscape evolution. *Geomorphology* **2016**, *27*, 63–77. [CrossRef]

50. Mace, R.E. *Estimating Transmissivity Using Specific-Capacity Data*; Geological Circular 01-2; Bureau of Economic Geology: Austin, TX, USA, 2001.