Use of Tracing Tests to Study the Impact of Boundary Conditions on the Transfer Function of Karstic Aquifers

L. Duran, M. Fournier, N. Massei and J.-P. Dupont

Abstract The impact of the variations of multiple environmental parameters on the response of karstic systems was investigated after a campaign of tracing tests acquired in very different hydrologic conditions. Principal components analysis and hierarchical clustering were applied on both environmental variables and karstic system response variables (parameters of the RTD curves). Equations between the RTD parameters and the most relevant variables were established using a symbolic regression algorithm. This model giving RTDs parameters in function of boundary conditions is more accurate than the PCA analysis since it takes into account the nonlinearity of the relations between variables. It appeared that the variations of the RTD parameters depend mainly on the piezometric level downstream of the aquifer, the cumulated rainfall preceding the injection, and on the tide coefficient (suggesting sensitivity to the annual variations of tide, in this case of a karstic system under marine influence). So the RTDS parameters are controlled by the hydraulic conditions downstream of the system, including tide. The dispersivity was found to be very sensitive to the precipitation and tides variations at a daily scale.

1 Introduction

Karstic aquifers represent 25 % of the water resources worldwide (Ford and Williams 2007), and are vulnerable to pollution, especially in regard of the high flow velocities in the conduits. Therefore, comprehension of those systems and the study of their vulnerability is a major stake in the management of the water resource. Tracing tests are a privileged way to understand the transport mechanisms within a karst system and to characterize its flow dynamics (Bakalowicz 2005). Commonly, one or two tracing tests are carried out. Still, the response of

L. Duran (✉) · M. Fournier · N. Massei · J.-P. Dupont
M2C, UMR 6143, CNRS, Morphodynamique Continentale et Côtière, Université de Rouen, Bât. IRESE a, 76821 Mont Saint Aignan, France
e-mail: lea.duran1@univ-rouen.fr

B. Andreo et al. (eds.), *Hydrogeological and Environmental Investigations in Karst Systems*, Environmental Earth Sciences 1,
DOI 10.1007/978-3-642-17435-3_13

the system (characterized by the breakthrough curve (BTC) and the residence time distribution (RTD)) changes significantly depending on the hydrologic conditions during the tracing test. For example, the study of the apparent velocity of the tracer can give some clues about the structure of the system (Dörfliger 2010). In order to fully characterize the functioning of a karstic system, tracing tests should ideally be carried out in different hydrologic conditions such as: low, medium, and high seasons, during rainfall events, and with variations of the water table. The transport variations in low flows and high flows have been investigated (Göppert and Goldscheider 2008; Larocque et al. 1998), and the impact of other components of the surface system like vegetation, nature of soils has been studied by simulation (Doummar et al. 2012). The variations of the BTCs according to the tracer used have been modeled (Geyer et al. 2007). Nevertheless, the impact of multiple boundary conditions (including various downstream controls) on the response of a karstic aquifer has not been fully assessed yet.

This study is based upon a campaign of tracing tests acquired on the same karstic system, but in different hydrologic conditions. Some statistical analyses have been carried out on the responses of this karstic system in order to assess the relative importance of all of the environmental variables and to establish relations between the parameters the karstic system response and those of the boundary conditions.

2 Methods: PCA, HC, Symbolic Regression

The principal components analysis method (PCA), commonly used to interpret hydrogeological data (Bakalowicz 1997; Fournier et al. 2007; Helena et al. 2000; Moore et al. 2009), was used in this study to assess the links between the response of the aquifer and the boundary conditions of the system. The individuals of the PCA are the tracing tests conducted on the study site. The selected variables are of two types. The first type corresponds to environmental variables (or boundary conditions of the karst aquifer): upstream boundaries with the precipitation, downstream boundaries, with the piezometric level in the aquifer, the tide coefficient, the level of the river (see description of the site). The second type is composed by variables related to the tracing test: recovery rate, characteristic times, dispersivity, and RTD parameters.

Hierarchical clustering was performed on the data: an agglomerative method with Ward algorithm (maximizing inter-class inertia, in order to obtain compact, spherical clusters) has been used on Euclidean distance matrix. In order to find the optimal number of clusters, Partitioning around medoids has been performed on HC results. The Kruskal Wallis post hoc test and Tukey's post hoc tests were then used to determine which groups were significantly different (Saporta 2011).

Once the more relevant variables were identified and the partitioning of the data studied, the next step was to find some relations between the environmental variables and the RTD variables. For that purpose, we used a software able to test a high number of possible relations between several variables: Eureqa (Version 0.98 beta).

Its functioning is based on a machine learning technique called Symbolic Regression to unravel the intrinsic relationships in data and explain them as simple equations. Using Symbolic Regression, Eureqa can create predictions (Schmidt and Lipson 2009). But even before obtaining the relationships between the variables, one of the interests is to identify which variables are relevant in the estimation of one particular parameter, as the algorithm calculates a lot of equations changing the nature and the number of variables implied.

3 Study Site and Tracing Tests Campaign

The Norville karst system is located in Normandy (France), near the Seine River. The site has been studied since 1999 and is described in various publications (Fournier et al. 2008; Massei 2001). It is a part of a national observation network on karstic systems (SNO karst).

The geology and geomorphology of the site are characteristic of the upper Normandy: the Seine River cuts deeply into chalk plateaus. An aquifer takes place in those formations than can be variously karstified. On the top of this Mesozoic chalk layer, the weathered chalk has formed a clay with flint layers, quite impervious. Swallow holes are penetrating the clay and chalk layers: the infiltration of water into the karstified chalk aquifer can be very quick. More locally, a N70E fault affects the site with a net slip of 120 m. South to the fault, the formations are the ones described before, while North to the fault, cretaceous layers of clay alternate with sand formations. Norville study site is a sinkhole-spring system: upstream, the small Bébec river drains a watershed of 10 km^2, before infiltrating into the ground when reaching the Triquerville fault (Fig. 1). The Bébec discharge has an important variability, from 5 l/s during low flow up to 400 l/s after storms events. Downstream, the perennial Hannetôt spring, at the bottom of the chalk cliffs, has been proved to be the resurgence of the Bébec River

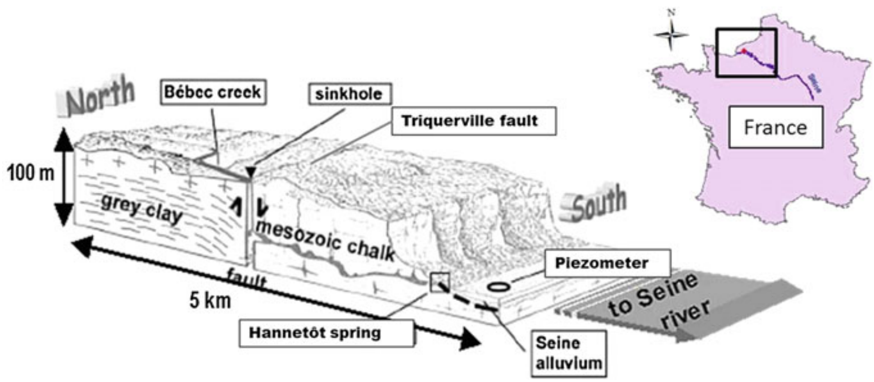

Fig. 1 Norville system (modified from Massei 2001)

with multiple tracing tests (Massei 2001). Karstic conduits are also likely to be found within the chalk aquifer between the spring and the Seine River, underneath alluvial aquifer; hence a hydraulic connection between the karstic aquifer and the Seine River cannot be excluded (Fournier et al. 2008). The variations of the level of a piezometer located south of the Seine, in the same chalk layer, are correlated to the variations at the spring.

The tracing tests are one of many tools available to assess the functioning of a karst system (Lepiller and Mondain 1986). In this study, 14 tracer tests were performed between 1999 and 2013, under various boundary conditions, including the precipitation, the piezometric level, the time within the hydrological cycle, and the tide conditions (being in the context of a karstic system under coastal influence). The tracer was injected into a perennial flow, in a sinkhole. Some of those tracing tests were made at less than 10 h interval in order to assess the role of the tide causing the BTCs to overlap. For that reason, it was sometimes necessary to separate these complex overlapping BTCs before interpreting them: we used PeakFit 4.0 (SPSS Inc.). The BTCs were then analyzed with the TRAC software, released by the BRGM (Klinka et al. 2012). The characteristic parameters of the BTCs were calculated, as well as the RTDs, and the dispersivity was estimated using the simulation tool. Through RTDs, different tracing tests within the same system or in different systems can be easily compared; the parameters of the RTD are linked to the dispersive parameters of the studied system. Here, the normalized RTDs (Fig. 2) present some important variations (area, time of the peak, tailing).

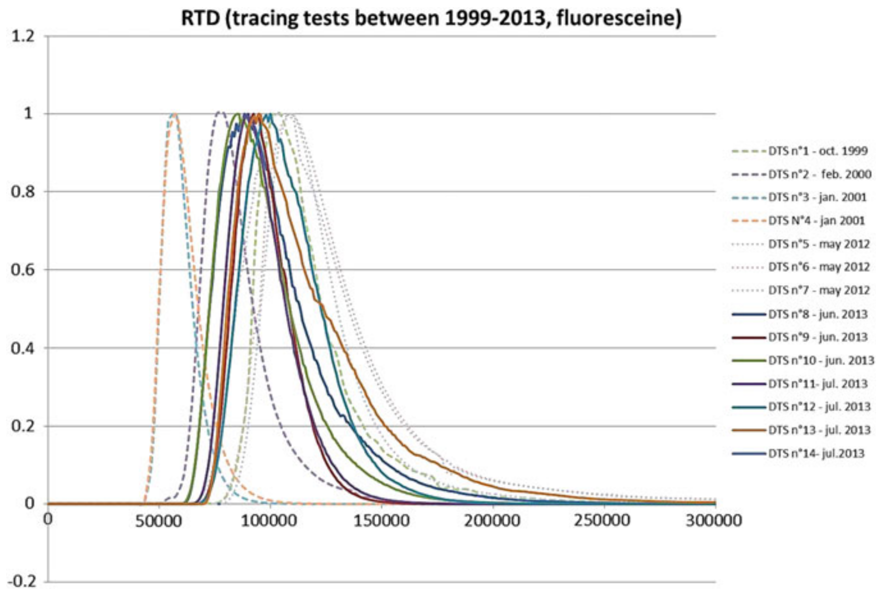

Fig. 2 Normalized residence time distribution curves

Table 1 Main results of the PCA (variables contributing positively and negatively)

Main results of PCA	Axis I	Axis II	Axis III	Axis IV	Axis V
Inertia (%)	53.7	14.17	10.34	6.7	6.05
Variables +	RTD parameters, BTC times	–	–	–	mar
Variables −	pf, qmoy, vmod, p7j, %rest	$p + 24$, $p - 24$	–	–	–

pf piezometric level within the chalk aquifer; *mar* tide coefficient; *P7j* cumulated rainfall during the week before the injection; $P - 24$, $P + 24$, rainfalls the day before and after the injection; Qmoy, mean discharge (l/s)

4 Results and Discussion

4.1 PCA Analysis

The results of the PCA analysis are presented in Table 1 and Fig. 3.

The percent of inertia explained by the different axis as well as the information given by the different components conducted to keep axis I, II, and V, explaining 73.89 % of the inertia. On axis I, several environmental variables (piezometric level, rainfall) contribute strongly negatively, when response parameters contribute positively. The main information on axis II is that the rainfalls of the 24 h before and after the injection contribute negatively. The tide coefficient (mar) contributes to axis V for 75 %. The fact that most of the environmental variables

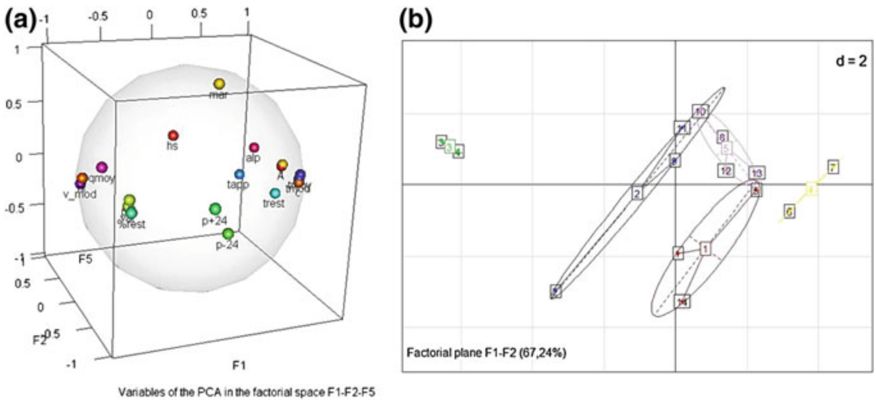

Fig. 3 Variables of the PCA **a** and tracing tests **b** in the factorial space F1–F2–F5. (*hf* piezometric level within the chalk aquifer; *mar* tide coefficient; *P7j* cumulated rainfall during the week before the injection; *P − 24*, rainfall the day before the injection; *P + 24*, rainfall the day after the injection; *Qmoy*, mean discharge (l/s); A, c and l: area, center and width of RTD; *vmod* modal velocity; *alp* dispersivity, trest, *%rest* recovery rate, *trest* time of recovery, *tapp* time of apparition)

are opposite to the RTD parameters on axis I suggests that they are strongly related; the objective is to characterize those links. In the individuals' space, some clear distinctions between the tracing tests appear: tracing tests number 3 and 4 contribute negatively to axis I, when tracing tests 1, 2, and 14 contribute negatively to axis II. In order to understand the structure of the data, some clustering and partitioning have been conducted.

4.2 Clustering and Interpretation of the Groups

The partitioning with medoids has been tested with a number of groups k from 2 to 10. The best partitioning is the one with the highest average silhouette width; here we obtained a width of 0.51 for $k = 5$. The next step is the hierarchical clustering with $k = 5$ (Fig. 4, left). The resulting groups contain from two elements to four elements, and can be drawn in the factorial plane (Fig. 4, right).

In order to test the significance of those groups, the Kruskal-Wallis test was applied. The null hypothesis was rejected for variables hf, %rest, tapp, tmod, tmoy, qmoy, alp, a, c, and l. This implies that for those variables, at least one group was significantly different from the other. The simple ANOVA test gave the same results. The ANOVA with regression on component gave a p-value of 0.0014, indicating that the partitioning was significant. Then the Tukey post hoc test was conducted to identify more precisely the groups which are significantly different. Independently, we attempted to assess the common conditions for the tracing tests in each group. The first group corresponds to tracing tests 3 and 4, with very high level in the piezometric level (extreme conditions) and precipitation conditions. Group 2 (tracing tests 2, 9, and 11) could be characterized by a rather low tide and a medium piezometric level. Group 3 (tracing tests 5, 1, and 14), would also concern tracing tests at a medium piezometric level, but with high tide conditions. Group 4 (tracing tests 6 and 7), corresponds to a low piezometric level. And at last, group 5 (tracing tests 8, 10, 12 and 13) would gather tracing tests with a high tide coefficient, or at default a high tide condition.

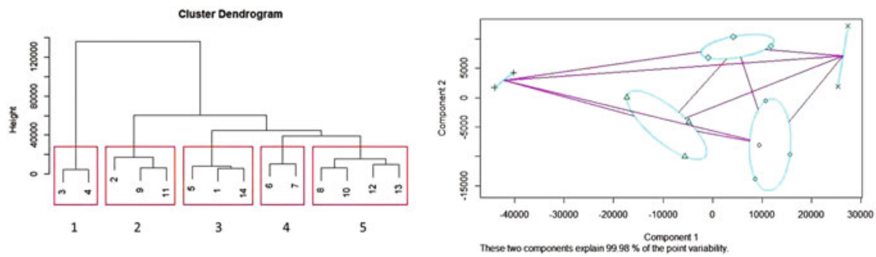

Fig. 4 Partitioning around medoids and cluster dendogram

4.3 Role of the Environmental Parameters

The clustering is particularly significant for all the variables characterizing the response to the karstic systems, like the RTD parameters. The PCA highlights the importance of the environmental parameters: the piezometric level in the chalk aquifer (downstream of the spring), as well as the spring discharge, are the major contributors to the first component of the PCA, and are anti-correlated to the parameters of the RTD curves. On axis II, the cumulated rainfall during the 24 h before the injection of tracer and the 24 h after the injection are the main contributors, and therefore it can be assumed that they play an important role in the response of the system. At last, the tide coefficient is by far the most important contributor to axis V; the tide is likely to be an important downstream boundary condition of the system.

4.4 Relations Between Environmental Variables and RTD Parameters

A lot of different configurations have been tested, by changing the number of variables included in the algorithm, and the forms of the mathematical equations. Each parameter of the RTD curve was expressed in function of all the environmental parameters; and the dispersivity in function of those same parameters. The possibilities of relations between A, the area of the RTD, were investigated. The piezometric level appears in 17 models over the 18 models calculated; it indicates that this variable is relevant to characterize the area, which is consistent with the results of the PCA. The second variable that appears in most models is the tide coefficient (in 14 models over 18). Then come the accumulated rainfall on the week before the injection and the rainfall 24 h before the injection. The level within the Seine River and the mean discharge are not well represented. The type of model obtaining the best R^2 Goodness of Fit and correlation coefficient ($R^2 = 0.96$ and $C = 0.987$) is:

$$A = a + b * \mathrm{qmoy} + d * P7j - c * \mathrm{mar} - f * \mathrm{Pant} - g * pf - h * P7j * P_{\mathrm{post}} - i * hs^2$$

The same investigation was conducted for the center of the RTD curve C. The piezometric level was also the most represented variable (in 19 over 21 models and 19 occurrences). But the second most recurrent variables were the precipitation of the week before and the day before and the mean discharge. The best fitted model obtained a R^2 goodness of fit of 0.992 and a correlation coefficient of 0.996. The corresponding equation is:

$$C = a + b * \mathrm{Pant} + d * P7j^2 - \mathrm{qmoy} - c * hs - f * pf - g * \mathrm{Pant}^2$$

The piezometric level was the principal variable in most of the models for the width of the RTD. The second one was the tide coefficient, followed by the mean discharge and precipitation preceding the injection. The equation obtaining the best R^2 goodness of fit and correlation coefficient (0.97 and 0.99) is:

$$l = a + b * \mathrm{qmoy} - d * pf - c * \mathrm{mar} + i * \mathrm{mar} * \mathrm{qmoy} + g * \mathrm{mar} * hs \\ - h * \mathrm{Pant} - i * hs - j * P7j + k * pf * P7j$$

As for the dispersivity, the most recurrent variable in all models was the cumulated rainfalls during the week preceding the injection (in 22 models over 24, with 42 occurrences). The second one was the level of the Seine River. The tide coefficient and the piezometric level in the chalk aquifer have an intermediary importance, appearing respectively in 14 and 15 models. The rainfalls at 24 h and the mean discharge were less significant with few occurrences. With a R^2 goodness of fit of 0.98 and correlation coefficient of 0.99, the selected equation for alpha is:

$$\alpha = a + b * P7j + c * hs + d * pf + e * \mathrm{mar} + hs * (f * \mathrm{Post} - g * pf) \\ + h * \mathrm{qmoy}$$

4.5 Discussion

The fact that the majority of models for *A*, *l,* and *c* contain the piezometric level *pf* as a key variable is consistent with the results of the PCA. The tide coefficient is essential to calculate the area and the width of the RTD. The importance of the precipitation was also highlighted, especially influencing the center of the RTD. The mean discharge appears as an important contributor to axis I in the PCA, whereas the algorithm used to estimate the relations between the environmental parameters and the parameters of the RTD does not enhance this discharge as essential. As for the dispersivity, the results of the algorithm indicate that the precipitations during the week preceding the injection contribute strongly to its variations. Unlike the parameters of the RTD, the dispersivity is sensitive to the variations of the level of the Seine River that is to say of the tide, but at a daily scale, not in its annual variations.

The equations obtained are to be taken with caution, since they are calculated on a relatively small sample (14 tracing tests). These equations have to be tested on later tracing tests. Moreover, even though the tracing tests have been conducted in order to cover the maximum diversity of environmental conditions, some are not well represented. For example, the tracing test $n°$ 14 is the only one with an important rainfall occurring during the injection. That explains its position in the factorial plane F1–F2, contributing negatively to axis II, which corresponds to $P + 24$ (rainfall in the day following the injection). So this variable could have a non-negligible effect, but could be underestimated because of the tracing tests sample.

Nevertheless, this is a first model of the response of karstic system (RTDs) in function of environmental parameters. Moreover, it takes into account the non-linearity of the relations between the variables, by opposition to the PCA which indicates only the possible linear relations between them. Another important result is that the shape of the RTDs is controlled by downstream conditions, including tide.

5 Conclusion

In this study, we investigated the links between various environmental parameters and some characteristic parameters of the karstic system response. A campaign of tracing tests was conducted in different hydrologic conditions (high flows, low flows, tide variations, rainfalls…) on the same karstic system. For each tracing test, the parameters of the BTC and the RTD curves and the dispersivity were assessed. A PCA indicated that most relevant variables were the piezometric level downstream of the karstic system, the cumulated rainfall, and the mean discharge. These were anti-correlated to the parameters of the normalized RTD curve. The precipitation around the injection was also relevant, as well as the tide coefficient. We studied the structuration of the data by using the partitioning around medoids and hierarchical clustering. These groups were linked to corresponding environmental conditions.

We established a model giving the expression of RTDs parameters in function of environmental variables. It appears that the downstream control is essential: the piezometric level and the tide coefficient are the most relevant variables. The cumulated rainfall and the mean discharge were considered as necessary variables in the estimation of the RTD parameters. It suggests that the mean discharge alone cannot provide enough information to estimate the RTD parameters and that the precipitation preceding the tracing tests is essential in the form of the RTD. The results also suggest that the dispersivity is very sensitive to the precipitation, and to the tide. The piezometric level was also an important variable, unlike the mean discharge. Equations were selected for each variable, but need to be cross-checked and validated through other tracing tests.

As perspectives, these analyses could be completed with more tracing tests, in order to confirm the trends brought out (especially investigating the rainfall parameter, with tracing tests during storms). The results could be compared to other karstic system. In the case of coastal karstic systems, it would enable to validate the hypothesis of the influence of annual variation of the tide on the response of the system, as well as the daily component of the tide on the dispersivity. For continental aquifers, it would be interesting to assess the joint role of the piezometric level and the precipitations in the response.

References

Bakalowicz M (1997) Water geochemistry: water quality and dynamics. In: Standford J, Gibert J, Danielopol D (eds) Ground-water ecology. Academic Press, San Diego, pp 97–127

Bakalowicz M (2005) Karst groundwater: a challenge for new resources. Hydrogeol J 13(1):148–160. doi:10.1007/s10040-004-0402-9

Dörfliger N (2010) Guide méthodologique: Les outils de l'hydrogéologie karstique pour la caractérisation de la structure et du fonctionnement des systèmes karstiques et l'évaluation de leur ressource

Doummar J, Sauter M, Geyer T (2012) Simulation of flow processes in a large scale karst system with an integrated catchment model (Mike She)—Identification of relevant parameters influencing spring discharge. J Hydrol 426–427:112–123. doi:10.1016/j.jhydrol.2012.01.021

Ford D, Williams P (2007) Karst Hydrogeology and Geomorphology. (John Wiley & Sons, Ed.). MacMaster, Canada; Wiley, Aukland, New Zealand, p 578. doi:10.1002/9781118684986

Fournier M, Massei N, Bakalowicz M, Dupont J (2007) Use of univariate clustering to identify transport modalities in karst aquifers. C. R. Geosci 339(339):622–631. doi:10.1016/j.crte.2007.07.009

Fournier M, Massei N, Mahler BJ, Bakalowicz M, Dupont JP (2008) Application of multivariate analysis to suspended matter particle size distribution in a karst aquifer, 2345 (October 2007), 2337–2345. doi:10.1002/hyp

Geyer T, Birk S, Licha T, Liedl R, Sauter M (2007) Multitracer test approach to characterize reactive transport in karst aquifers. Ground Water 45(1):36–45. doi:10.1111/j.1745-6584.2006.00261.x

Göppert N, Goldscheider N (2008) Solute and colloid transport in karst conduits under low- and high-flow conditions. Ground Water 46(1):61–68. doi:10.1111/j.1745-6584.2007.00373.x

Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000) Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal components analysis. Water Resour 34(3):807–816

Klinka T, Gutierrez A, Thiéry D (2012) Validation du logiciel TRAC: Aide à l ' interprétation de traçages en milieu poreux Rapport final Validation du logiciel TRAC: Aide à l ' interprétation de traçages en milieux poreux Rapport final. BRGM/RP-59425-FR, p. 58

Larocque M, Mangin A, Razack M, Banton O (1998) Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France). J Hydrol 205(3–4):217–231. doi:10.1016/S0022-1694(97)00155-8

Lepiller M, Mondain P (1986) Les traçages artificiels en hydrogéologie karstique. Mise en oeuvre et interprétation. Hydrogéologie 1:33–52

Massei N (2001) Transport de partcules en suspension dans l'aquifère crayeux karstique et à l'interface craie/alluvions. University of Rouen

Moore PJ, Martin JB, Screaton EJ (2009) Geochemical and statistical evidence of recharge, mixing, and controls on spring discharge in an eogenetic karst aquifer. J Hydrol 376(3–4):443–455. doi:10.1016/j.jhydrol.2009.07.052

Saporta, G. (2011). Probabilités, analyse des données et statistique (Editions T), pp 243–266

Schmidt M, Lipson H (2009) Distilling free-form natural laws from experimental data. Science (New York, N.Y.) 324(5923):81–85. doi:10.1126/science.1165893