



Hydraulic conductivity explored by factor analysis of borehole geophysical data

Norbert Péter Szabó

Abstract A multivariate statistical method is presented for providing hydrogeological information on groundwater formations. Factor analysis is applied to borehole logs in Hungary and the USA to estimate the vertical distribution of hydraulic conductivity of rocks intersected by the borehole. Earlier studies showed a strong correlation between a statistical variable extracted by factor analysis and shale volume in primary porosity rocks. Hydraulic conductivity as a related quantity can be derived directly by factor analysis. In the first step, electric and nuclear logs are transformed into factor logs, which are then correlated to hydraulic properties of aquifers. It is shown that a factor explaining the major part of variance of the measured variables is inversely proportional to hydraulic conductivity. By revealing the regression relation between the above quantities, an estimate for hydraulic conductivity can be given along the entire length of the borehole. Synthetic modeling experiments and field cases demonstrate the feasibility of the method, which can be applied both in primary and secondary porosity aquifers. The results of factor analysis show consistence with those of the Kozeny-Carman method and hydraulic aquifer tests. The application of the statistical analysis of well logs together with independent ground geophysical and hydrogeological methods serves a more efficient exploration of groundwater resources.

Keywords Factor analysis · Factor log · Hydraulic conductivity · Hungary · USA

Introduction

Hydraulic conductivity quantifies the ease with which water can move through the intergranular pore and fracture spaces of the formation. In hydrogeological

problems, it is one of the most important petrophysical properties of rocks that can be generally measured in the laboratory using rock samples, and by aquifer tests performed in the field or at the aquifer scale, for instance from aquifer modeling. The study is mostly performed at the well log scale, which attempts to calculate hydraulic conductivity purely from well-logging measurements influenced by only the near vicinity of the borehole at a given depth. In porous media, the hydraulic conductivity is related theoretically to the grain-size, porosity and fracture characteristics. In primary porosity rocks, other textural properties of rocks are also taken into account such as cementation exponent or tortuosity factor. Geophysical interpretation approaches usually use some empirical method or statistical tool for the evaluation (Idrissy and De Smedt 2007; Ross et al. 2007; Odong 2013). Borehole geophysical measurements are part of the in-situ investigations that are used primarily to detect the variation of hydraulic conductivity along a borehole and to correlate it between neighboring boreholes. In oilfield applications, the direct determination of permeability as a related quantity is possible by means of the nuclear magnetic resonance (NMR) log. The surface geophysical application of its technique known as magnetic resonance sounding (MRS) is an emerging method in hydrogeology (Roy and Lubczynski 2003). Borehole NMR has been adapted from the oilfield for hydrogeological applications, using boreholes typical of environmental and hydrogeological investigations (Walsh et al. 2013). Although borehole NMR is very expensive, it has the added advantage of not only providing effective porosity, it can also be used to determine the pore-size distribution and pore-fluid characteristics to provide a better estimate of the hydraulic properties of rocks. The effective pore-radius based permeability prediction also has promising results both in the laboratory and in hydrocarbon exploration fields (Glover and Walker 2009). The indirect (in-situ) methods for hydraulic conductivity estimation are based on the determination of formation porosity and groundwater saturation (Timur 1968). Alger (1966) connected formation factor to effective grain-size, which allowed for calculation of hydraulic conductivity from borehole logs. As a continuation, Csókás (1995) suggested a well-logging technique to estimate hydraulic conductivity and other freshwater quality parameters for unconsolidated sediments, which requires the preliminary calculation of

Received: 16 June 2014 / Accepted: 29 January 2015
Published online: 26 March 2015

© Springer-Verlag Berlin Heidelberg 2015

N. P. Szabó (✉)
Department of Geophysics,
University of Miskolc, 3515, Miskolc, Egyetemváros, Hungary
e-mail: norbert.szabo.phd@gmail.com
Tel.: +3646361936

porosity, pore-water resistivity and true resistivity of the aquifer. The theory of freshwater assessment by means of well-logging information was summarized in Alger and Harrison (1989). In this framework, the reservoir parameters (including hydraulic conductivity) are connected to physical quantities that are measurable by well-logging probes. For extracting the unknown petrophysical parameters some deterministic or inverse modeling-based procedure is usually applied. A resistivity and porosity log-based approach can be found in Khalil et al. (2011), while example inversion applications can be found in Drahos (2005), Szabó and Dobróka (2013a).

Most types of well logs used in oilfields are commonly applicable in hydrogeology practice. Spontaneous potential and natural gamma-ray intensity logs are used for lithology identification and calculation of shale volume. Gamma-gamma and neutron-neutron intensity measurements give an accurate estimate of porosity. Resistivity tools are mainly sensitive to the water saturation that is calculated by them. In a regular case, the above well log types fulfill the requirements; however, there are some other advanced techniques that could give further information. Acoustic measurements are typically used for porosity estimation but, in shallow applications, nuclear logs work better. In secondary porosity rocks, the sonic log gives lower porosity reading than true porosity of the formation, because the acoustic waves avoid vugs and fractures as a result of the Fermat's principle. Full-wave sonic logs may provide more detailed information on porosity, elastic parameters and horizontal stress conditions in the vicinity of the borehole. The separation of Stoneley waves propagating in the borehole enables the determination of the permeability of porous formations. Permeability is generally estimated from the inversion or statistical processing of Stoneley transit-time data with an advantage that they do not require the prior knowledge of porosity (Buffin 1996). Recently a non-linear statistical model was suggested by Szabó and Kalmár (2013) for improving the description of the relation between the characteristics of Stoneley waves and permeability. Their findings suggest that the circumferential borehole acoustic or optical images used in fractured formations should be completed with full-wave sonic logs for a more accurate and reliable interpretation.

When the currently used data processing methods are evaluated, one can see that each of them have their own weaknesses. The assumed model can be ambiguous and the data sets may be noisy or not sensitive enough to give a good estimate to hydraulic properties. Moreover, the interpretation results are often in contradiction with those determined from core samples. The estimation error of permeability may reach one to one and a half orders of magnitude. To reduce the uncertainty, the borehole surveys are frequently expanded with ground geophysical measurements. In Perdomo et al. (2014), hydraulic parameters are estimated by a joint application of well-logging resistivity and direct current electric measurements. Slater (2007) combined surface induced-polarization measurements with borehole flowmeter data,

ground penetrating radar tomographic data and neutron-log-derived porosity information. Guérin (2005) presents the advantages of the application of electromagnetic methods. In the hydrogeological assessment of underground mines, quantitative information can be extracted from the joint inversion of borehole seismic and in-mine geoelectric data, used for instance, in detecting tectonic disturbances and fault zones in coal seam series, water inrush and in estimating the thickness of impervious layers (Dobróka et al. 1991).

In addition to the present measurement techniques, the introduction of an independent data processing method is of utmost importance. The simultaneous application of the new method and the existing ones can improve the accuracy and reliability of the estimation result. In this paper, an alternative approach based on multivariate statistical principles is presented, which processes all borehole logs together to give an estimate to the vertical distribution of formation hydraulic conductivity. Factor analysis is normally used to reduce the dimensionality of multivariate statistical problems, and to extract latent information from the data set that is non-measurable. The basic principle of the theory of factor analysis can be found in the paper of Lawley and Maxwell (1962). Several geophysical applications showed that the extracted factor variables correlate to some petrophysical properties of geological formations. Szabó (2011) introduced a factor analysis-based method to estimate the shale volume of sedimentary rocks. The statistical technique gave proper estimates for several domestic and some overseas deep wells in hydrocarbon fields (Szabó and Dobróka 2013b). The method was also tested in some shallow freshwater wells drilled in Hungary. Szabó et al. (2014) proposed a regression relationship between one of the factors and shale volume for eastern Hungary. Similar results were recently published by Asfahani (2014), who processed nuclear logs including natural gamma-ray intensity, density and neutron-porosity data and long- and short-normal electrical logs by factor analysis to characterize the large extended basaltic areas in southern Syria. It was concluded that the factor responsible for the largest part of variance of original logs (i.e. first factor) can be interpreted as a shale factor, which is useful for separating different lithological units. The hydraulic conductivity of groundwater formations is strongly related to shale volume in primary porosity rocks. In this study, it is assumed that the well log of the first factor correlates adequately with hydraulic conductivity, which is useful to extract the relevant parameter from borehole geophysical measurements. The feasibility of the statistical method is presented by synthetic modeling experiments and field studies, and the interpretation results are validated by independent laboratory and aquifer tests.

Theory of the method

According to Darcy's law, hydraulic conductivity is the proportionality factor between Darcy's velocity of water

flow and hydraulic gradient, which in rocks with predominantly primary porosity, depends on the density and viscosity of pore-water, grain-size and pore-size distribution, porosity and water saturation. The spatial distribution of the related parameters can be derived from well-logging measurements. In the literature regarding the forementioned parameters, some empirical or approximate formula is normally used to calculate hydraulic properties such as hydraulic conductivity, transmissivity and storativity. The hydraulic conductivity K is directly proportional to intrinsic permeability expressing the measure of the aquifer's ability to transmit water through its pore spaces. The Kozeny-Carman equation is one of the most widely used formulas for the estimation of hydraulic conductivity given in units of cm/s (Bear 1972)

$$K = \frac{\rho_w g}{\mu} \frac{d^2}{180} \frac{\Phi^3}{(1-\Phi)^2}, \quad (1)$$

where d (cm) is the grain diameter, Φ (v/v) is the porosity of formation, ρ_w (g/cm³) is the density of pore-fluid, μ (g/cm/s) is the dynamic viscosity and g (cm/s²) is the normal acceleration of gravity. In Eq. (1), the dominant grain diameter d (cm) can be found from grain-size analysis (Juhász 2002)

$$d = \frac{d_{10} + d_{60}}{2} \sqrt{\frac{d_{10}}{d_{60}}}, \quad (2)$$

where d_{10} (cm) and d_{60} (cm) are the representative sample diameters at 10 and 60 % cumulative frequencies, respectively. As rock samples can be taken from boreholes and porosity can be estimated from well logs, the vertical distribution of hydraulic conductivity can be calculated continuously along the borehole. A new statistical method is presented in this paper that makes use of suitable well logs sensitive to hydraulic properties of rocks to give an estimate to hydraulic conductivity for the entire logging interval.

The method of factor analysis is applied to borehole logs in the following manner. Let the column vector \mathbf{d}^l contain the data of the l -th measured variable along the borehole. The readings of all data types are gathered in data matrix \mathbf{D}

$$D_{il} = d_i^l, \quad (3)$$

where $i = 1, 2, \dots, N$ is the total number of measuring points in the processed depth-interval and $l = 1, 2, \dots, L$ is the number of geophysical tools measuring different physical quantities in the investigated borehole. The input data must be standardized in the first step of the analysis

$$\hat{D}_{il} = \frac{(D_{il} - \bar{D}_l)}{\sqrt{\frac{1}{N-1} \sum_{i=1}^N (D_{il} - \bar{D}_l)^2}}, \quad (4)$$

where \bar{D}_l represents the arithmetic mean of the data measured by the l -th probe. Factor analysis reduces the N -by- L matrix in Eq. (4) to a lower dimension by the matrix decomposition

$$\hat{\mathbf{D}} = \mathbf{F}\mathbf{W}^T + \mathbf{E}, \quad (5)$$

where \mathbf{F} denotes the N -by- M matrix of factor scores, \mathbf{W} is the L -by- M matrix of factor loadings, \mathbf{E} is the N -by- L matrix of residuals, M is the number of factors extracted from a higher number of observed variables, that is, $M < L$ (T indicates the matrix transpose operator). The factor scores given in a column of matrix \mathbf{F} represent a well log of the extracted statistical variable. Matrix \mathbf{W} contains the weights of individual data corresponding to the extracted factors. Practically, the factor loadings represent the degree of correlation between each factor and measured data type. Since the factors are assumed to be linearly independent ($\mathbf{F}^T\mathbf{F}/N = \mathbf{I}$), the correlation matrix of the standardized data is

$$\mathbf{R} = \frac{1}{N} \hat{\mathbf{D}}^T \hat{\mathbf{D}} = \mathbf{W}\mathbf{W}^T + \mathbf{\Psi}, \quad (6)$$

where $\mathbf{\Psi} = \mathbf{E}^T\mathbf{E}/N$ is the diagonal matrix of specific variances (\mathbf{I} is the identity matrix). If the notation of communalities represented by the elements of the main diagonal of matrix $\mathbf{R}^c = \mathbf{W}\mathbf{W}^T$ is introduced, it can be realized that matrix $\mathbf{\Psi}$ represents the part of variance of the observations that are not explained by the common factors. Normally \mathbf{W} and $\mathbf{\Psi}$ are estimated by an iterative algorithm that minimizes the following type of objective functions

$$\Omega(\mathbf{W}, \mathbf{\Psi}) = \text{tr}(\mathbf{R} - \mathbf{W}\mathbf{W}^T - \mathbf{\Psi})^2 = \min, \quad (7)$$

where tr denotes the trace of the square matrix given in the argument. Function Ω must be minimized with respect to factor loadings and specific variances simultaneously. For solving the optimization problem the use of the maximum likelihood method is generally applied, which can give a robust solution (Jöreskog 1969). Assuming that \mathbf{W} and $\mathbf{\Psi}$ are known, the factor logs can be extracted by the maximization of the following log-likelihood function

$$\lg(P) = -\frac{1}{2} \left[\lg|2\pi\mathbf{\Psi}| + (\hat{\mathbf{D}} - \mathbf{F}\mathbf{W}^T) \mathbf{\Psi}^{-1} (\hat{\mathbf{D}} - \mathbf{F}\mathbf{W}^T)^T \right] = \max. \quad (8)$$

After solving Eq. (8), an unbiased estimate to factor scores can be given by the hypothesis of linearity (Bartlett 1953)

$$\mathbf{F}^T = (\mathbf{W}^T \mathbf{\Psi}^{-1} \mathbf{W})^{-1} \mathbf{W}^T \mathbf{\Psi}^{-1} \hat{\mathbf{D}}^T. \quad (9)$$

The optimal number of factors can be set by statistical tests (Bartlett 1950) or a non-iterative approach (Jöreskog 2007). The resultant factors are usually rotated for an easier interpretation. Since factor loadings are defined non-uniquely, an orthogonal transformation $\mathbf{W}\mathbf{W}^T = \mathbf{W}^*\mathbf{W}^{*T}$ can be applied to factor loadings, where $\mathbf{W}^* = \mathbf{W}\mathbf{V}$ holds for a suitably chosen M -by- M orthogonal matrix \mathbf{V} . In this study, the *varimax* algorithm suggested by Kaiser (1958) is used to generate rotated factors, which can be directly compared to hydraulic conductivity of groundwater formations.

The resultant factors can be related with formation parameters in regression analysis. Szabó et al. (2014) showed a strong exponential connection between the first factor (i.e. first column of matrix \mathbf{F}) and shale volume in groundwater formations, where the regression coefficients obtained were approximately the same for different areas in eastern Hungary. As hydraulic conductivity is inversely proportional to shale content in primary porosity rocks (Benson and Trast 1995; Sallam 2006; Shevnin et al. 2006), it is assumed that the first factor variable is also sensitive to hydraulic conductivity. In this study, a linear relationship between the first factor and the decimal logarithm of hydraulic conductivity is demonstrated

$$-\lg(K) = aF_1^* + b, \quad (10)$$

where a , b are site specific constants and F_1^* is the first factor scaled into an arbitrary interval. Equation (10) is confirmed by synthetic modeling experiments (see section ‘Synthetic modeling experiments’) and well-site studies (see section ‘Well-site applications’). The Pearson’s correlation coefficient $R = \text{cov}(K, F_1^*) / \sigma_K \sigma_{F_1^*}$ characterizes the strength of the linear connection between the given factor and the logarithm of hydraulic conductivity, where cov is the sample covariance operator, σ_K and $\sigma_{F_1^*}$ are the standard deviations of the correlated quantities, respectively. For statistical experiments using an exactly known model the misfit between the noisy and noiseless synthetic data is measured by the relative data distance

$$D_d = \sqrt{\frac{1}{NL} \sum_{i=1}^N \sum_{l=1}^L \left(\frac{d_{il}^{(m)} - d_{il}^{(c)}}{d_{il}^{(m)}} \right)^2}, \quad (11)$$

where $d_{il}^{(m)}$ and $d_{il}^{(c)}$ denote the l -th measured and calculated data in the i -th depth point, respectively. Multiplying D_d by 100, the measure of misfit is given in percent. The closeness of hydraulic conductivities estimated from different sources can be measured also by the Pearson’s correlation coefficient.

Synthetic modeling experiments

Consider the petrophysical model with exactly known parameters such as effective porosity (POR), water

saturation in the immediate vicinity of the borehole flushed by mud (SX0) and that of the undisturbed formation away from the borehole including only pore-water (SW), shale volume (VSH), sand volume (VSD), dominant grain-size (D) and hydraulic conductivity (K). The model parameters vary vertically along the borehole. The model represents a shallow unconsolidated sedimentary geological formation made up of five inhomogeneous beds (model-well-1). In the near-surface region, not only freshwater but gas (normally air) may fill the pore space. The air saturation can be calculated by $SG = 1 - SW$, which can be divided into movable ($SGM = SX0 - SW$) and irreducible ($SGIR = 1 - SX0$) parts. The lithology from the top to the bottom is silty sand (42 % air and 58 % water), fine-grained sand (42 % air and 58 % water), shale (100 % water), fine-grained sand (35 % air and 65 % water) and shaly sand (100 % water). Typical grain-sizes are set from the literature (Wentworth 1922), while hydraulic conductivities are calculated by Eq. (1). In Well-1 bulk density (DEN), natural gamma-ray intensity (GR), spontaneous potential (SP), deep resistivity (RD), shallow resistivity (RS), neutron-neutron intensity (NN) logs are applied. The parameters of the petrophysical model are usually related to borehole geophysical measurements empirically. These mathematical relations called probe response equations can be used to predict data in a forward modeling procedure. The values of theoretical data would be measured along the borehole, if the geological structure was characterized by the assumed (exactly known) model parameters. Most types of borehole geophysical data can be expressed as a linear combination of the physical properties of rock matrix and fluid components weighted by the relative volumes of rock constituents. The following set of response functions can be used for the solution of forward problem in groundwater formations (Alberty and Hashmy 1984)

$$\begin{aligned} \text{DEN} = & \text{POR}[(\text{SX0} \cdot \text{DEMF}) + (1 - \text{SX0})\text{DEG}] \\ & + \text{VSH} \cdot \text{DESH} + \text{VSD} \cdot \text{DESD}, \end{aligned} \quad (12)$$

$$\text{GR} = \text{GRSD} + \frac{1}{\text{DEN}} (\text{VSH} \cdot \text{GRSH} \cdot \text{DESH} + \text{VSD} \cdot \text{GRSD} \cdot \text{DESD}), \quad (13)$$

$$\begin{aligned} \text{SP} = & \text{SPSD} + \text{VSH}(\text{SPSH} - \text{SPSD}) \\ = & \text{VSH} \cdot \text{SPSH} - C \cdot \lg\left(\frac{\text{RMF}}{\text{RW}}\right) (1 - \text{VSH}), \end{aligned} \quad (14)$$

$$\text{NN} = \text{POR} \cdot \text{NNF} + \text{VSH} \cdot \text{NNSH} + \text{VSD} \cdot \text{NNSD}, \quad (15)$$

$$\frac{1}{\sqrt{RD}} = \left[\frac{VSH^{(1-VSH/2)}}{\sqrt{RSH}} + \frac{(\sqrt{POR})^{BM}}{\sqrt{BA \cdot RW}} \right] (\sqrt{SW})^{BN}, \tag{16}$$

$$\frac{1}{\sqrt{RS}} = \left[\frac{VSH^{(1-VSH/2)}}{\sqrt{RSH}} + \frac{(\sqrt{POR})^{BM}}{\sqrt{BA \cdot RMF}} \right] (\sqrt{SX0})^{BN}, \tag{17}$$

$$POR + VSH + VSD = 1. \tag{18}$$

In Eqs. (12)–(17), there are additional quantities called zone parameters that express the physical properties of the solid and fluid parts of the groundwater formation. The detailed list of zone parameters practically chosen as constant in the forward modeling procedure can be found in Table 1. Equation (18) is the material balance equation for the rock environment, which is used to constrain the domain of model parameters in the interpretation procedure. Response Eqs. (12)–(17) can be used to generate synthetic borehole logs. In the test, these data are contaminated with some amount of random noise to produce quasi measured logs. By the processing of the noisy data it demonstrates how accurately the statistical procedure reconstructs the parameters of the exact model. The experiment evaluates the performance of the method, namely it characterizes its accuracy, stability and noise sensitivity.

The method of factor analysis is tested on synthetic well-logging data calculated on model-well-1. The workflow of the method is detailed in Szabó et al. (2014). Borehole logging data are calculated to each depth by substituting the actual values of model parameters (POR, SX0, SW, VSH, VSD) to Eqs. (12)–(17). As a result, six types of well logs (GR, SP, DEN, NN, RS, RD) in 250 depth levels are given, where the total number of data is 1,500. The synthetic data set is contaminated by random noise by adding a random number to each data generated from Gaussian probability distribution with zero mean and a scale parameter proportional to the noise level. The values of zone parameters are listed in Table 1. The correlation matrix in Table 2 shows that data variables represented by noisy synthetic data are relatively strongly correlated. The first two factors (i.e. first and second columns of matrix **F**) are calculated, because they explain 90.8 % and 9.2 % of total variance of measured data, respectively. Table 3 contains the loadings of the factors calculated by Eq. (8) for the case of the synthetic data set including 5 % Gaussian distributed noise. The largest

Table 1 Groundwater-zone parameters chosen for synthetic modeling experiments

Zone parameter	Definition	Symbol	Value (unit)
Gamma ray intensity	Shale	GRSH	160 API
	Sand	GRSD	10 API
Spontaneous potential	Shale	SPSH	0 mV
	Sand	SPSD	12.3 mV
	Temperature constant	C	70
Thermic neutron intensity	Shale	NNSH	4.0 kcpm
	Sand	NNSD	7.5 kcpm
Gamma-gamma (density)	Pore-fluid	NNF	3.0 kcpm
	Mud-filtrate	DEMF	0.9982 g/cm ³
	Shale	DESH	2.5 g/cm ³
	Sand	DESD	2.65 g/cm ³
Electric resistivity	Gas (air)	DEG	1.2 · 10 ⁻³ g/cm ³
	Mud-filtrate	RMF	8 ohmm
	Pore-water	RW	12 ohmm
Textural parameters	Shale	RSH	2 ohmm
	Cementation exponent	BM	1.5
	Saturation exponent	BN	2.0
Hydraulic parameters	Tortuosity factor	BA	1.0
	Dynamic viscosity	MU	0.019 g/cm/s
	Acceleration of gravity	G	981 cm/s ²

weights on the first factor are given by lithology sensitive logs (GR and SP), while those of the second one are obtained by resistivity logs (RS and RD). These results are consistent with earlier studies (Szabó 2011; Szabó and Dobróka 2013b; Szabó et al. 2014). Asfahani (2014) also concluded that the first factor expresses the presence of clay in basaltic formations, and can be termed as a clay factor, while the second factor can be used to separate different lithological units according to their resistivity response; therefore, it can be termed as a resistivity factor. Two uncorrelated factors are calculated with the factor loadings by Eq. (9), which are then rotated and scaled to the interval of 0 and 100 (in earlier studies the same interval was chosen for the sake of comparability). The linear connection between the first factor and hydraulic conductivity in Fig. 1 is based on Eq. (10). The regression model is $\lg(K) = -0.033F_1^* - 2.72$, where coefficients *a*, *b* are obtained highly reliably ($a_{\min} = -0.034$, $a_{\max} = -0.032$, $b_{\min} = -2.75$, $b_{\max} = -2.70$ with 95 % confidence bounds). The Pearson’s correlation coefficient between $\lg(K)$ and

Table 2 Correlation matrix of synthetic borehole logging data in case of exactly known inhomogeneous model

	SP	GR	DEN	NN	RS	RD
SP	1	0.97	0.56	-0.84	-0.82	-0.80
GR	0.97	1	0.56	-0.86	-0.84	-0.82
DEN	0.56	0.56	1	-0.46	-0.51	-0.49
NN	-0.84	-0.86	-0.46	1	0.71	0.69
RS	-0.82	-0.84	-0.51	0.71	1	0.98
RD	-0.80	-0.82	-0.49	0.69	0.98	1

Table 3 Factor loadings derived from synthetic wellbore data in case of inhomogeneous model for Well-1

Factor	$W^{(SP)}$	$W^{(GR)}$	$W^{(DEN)}$	$W^{(NN)}$	$W^{(RS)}$	$W^{(RD)}$
1st	0.97	0.98	0.55	-0.86	-0.80	-0.78
2nd	-0.06	-0.08	-0.09	0.02	0.58	0.62

F_1^* is $R = -0.98$, which indicates a strong correlation and inverse proportionality between the variables. In Fig. 2, the input logs, the petrophysical model and the estimated hydraulic conductivity logs are illustrated, where GR log primarily correlates with shale volume of the formations. Since the resistivity of mud-filtrate (RMF) is lower than that of the pore-water (RW), a sand aquifer is at higher potential than shale (reverse SP). Shales are not invaded by mud; therefore, RS and RD logs overlap. The porous formations are flooded by a more conductive mud filtrate than the original pore-water. Normally in water-saturated rocks, the mud causes lower RS values than RD. The RD curve ascends more in the first, second and fourth layers where the pore space is partially saturated with air. The air effect appears as a separation between the nuclear logs, which provide porosity estimates for Eq. (1). The composition and saturation of the modelled rocks are represented by the last two tracks of Fig. 2. Beside the two factors, the well logs of the hydraulic conductivity estimated from Eq. (1) and Eq. (10) are plotted in track 7. The K(KOZENY) and K(FA) logs show a close agreement with a correlation coefficient of $R = 0.98$.

Factor analysis as a data processing method is necessarily affected by the propagation of errors. The factor logs are estimated with a certain accuracy depending on the noise level of input data. A set of additional

synthetic tests are performed using different level of uncertainties of wellbore data calculated on model-well-1 to examine the noise sensitivity of the factor analysis algorithm. In Table 4 the test results for data sets with different data distances based on Eq. (11) are listed. Beside Gaussian noise up to 10 %, some non-Gaussian typed (asymmetric) data distribution are also generated by adding outliers to the Gaussian distributed data (five times higher amount of noise is added randomly to the 10 % of the given Gaussian distributed data). Two factors are calculated in each case. The magnitude and sign of factor loadings do not vary significantly, as they show a low-rate decay for lithology dependent logs with increasing data distances. The coefficients for the regression connection between the hydraulic conductivity and the first-scaled factor show high accuracy and little variation. Estimation errors are directly proportional to data noises and are relatively small even in highly noisy cases. The Pearson's correlation coefficients show strong and steady relationship between the two variables. Non-Gaussian tests show that factor analysis is properly resistant against data distributions different from Gaussian type. The regression connection between the factor and hydraulic conductivity can also be recognized beside extreme noises. The synthetic tests show a reliable and stable statistical algorithm.

Well-site applications

Validation with Kozeny-Carman model

The suggested statistical method works properly when applied to real data sets. The following example shows a

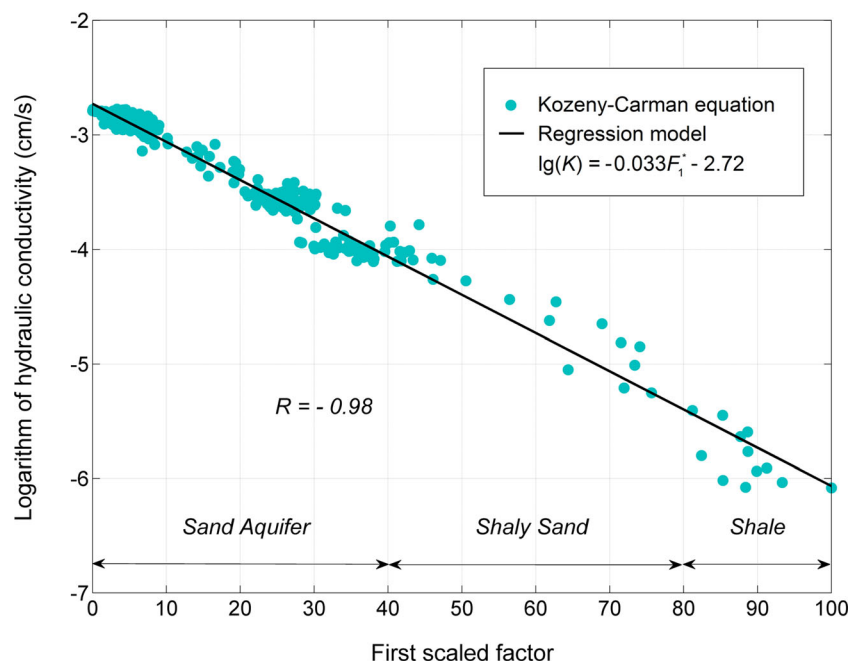


Fig. 1 Hydraulic conductivity versus factor scores derived from synthetic well-logging data contaminated by 5 % Gaussian distributed noise for the model-well-1

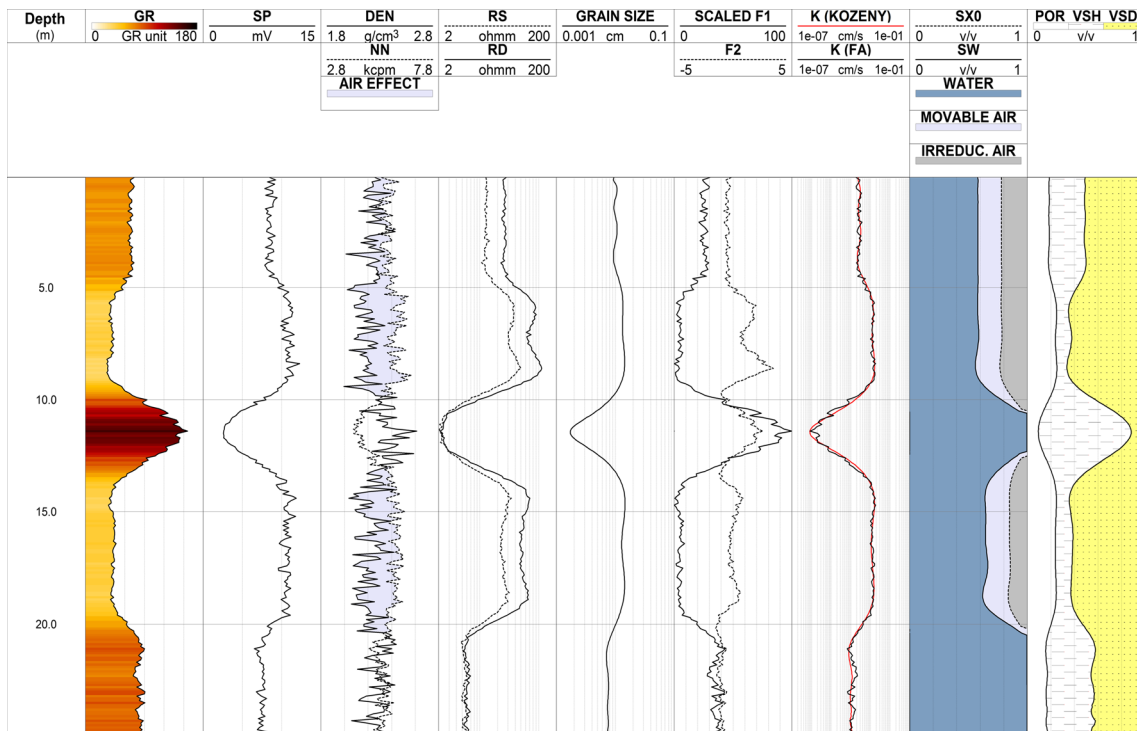


Fig. 2 Synthetic borehole logs calculated for model-well-1. *GR* is natural gamma-ray intensity, *SP* is spontaneous potential, *DEN* is density, *NN* is neutron-neutron intensity, *RS* and *RD* are shallow and deep resistivity logs, respectively. Dominant grain diameter is in track 5. The well logs of the extracted factors are in track 6. Hydraulic conductivity (*K*) logs are estimated from Kozeny-Carman equation (*red curve*) and factor analysis (*FA, black curve*), separately (track 7). *SX0* is water saturation in the invaded zone, *SW* is water saturation in the virgin zone (track 8), *POR* is porosity, *VSH* is shale content, *VSD* is sand volume (track 9)

Hungarian case study. In the Pannonian Basin Province of Central Europe, a thick Tertiary sedimentary sequence including hydrocarbon and thermal-water reservoirs was deposited over the Mesozoic and older basement rocks. The overlying unconsolidated Quaternary formations mainly consist of freshwater-bearing gravel or sand aquifers confined by shales. The shallow part of well B-1 is investigated, which is located in Baktalórántháza, Szabolcs-Szatmár-Bereg County, North-East Hungary (Fig. 3). The aim of the survey and earlier results of factor analysis can be found in Szabó et al. (2014).

In this study, factor analysis is applied to borehole logs to evaluate the hydraulic conductivity of geological

formations between the interval of 105 and 486 m, where 176 core samples were previously collected. The processed well logs include: natural gamma-ray intensity (*GR*), spontaneous potential (*SP*), shallow resistivity (*RS*), gamma-gamma intensity (*GG*) and neutron-neutron intensity (*NN*). The total number of data is 19,075 which are processed together in one interpretation procedure. The strength of correlation between the measured variables is moderate, as is seen in Table 5. The correlation coefficients are even smaller than they are assumed in the calculations of the synthetic case (Table 2). Singular value decomposition (*SVD*) of the data covariance matrix shows that the total variance of input data can be explained by

Table 4 Results of noise sensitivity tests in case of inhomogeneous model and noisy synthetic well-logging data sets

Data noise	D_d (%)	$W^{(SP)}$	$W^{(GR)}$	$W^{(DEN)}$	$W^{(NN)}$	$W^{(RS)}$	$W^{(RD)}$	a_{min}	a	a_{max}	b_{min}	b	b_{max}	$r^{(F_1^*, \lg K)}$
Gaussian	1.0	0.99	0.99	0.88	-0.99	-0.83	-0.80	-0.031	-0.030	-0.030	-2.74	-2.73	-2.72	-0.99
	2.0	0.99	0.99	0.78	-0.97	-0.82	-0.78	-0.033	-0.032	-0.031	-2.74	-2.72	-2.71	-0.99
	3.0	0.98	0.98	0.68	-0.93	-0.80	-0.76	-0.032	-0.031	-0.030	-2.75	-2.73	-2.71	-0.99
	4.0	0.97	0.98	0.53	-0.90	-0.77	-0.74	-0.034	-0.033	-0.032	-2.75	-2.73	-2.71	-0.98
	5.0	0.97	0.98	0.55	-0.86	-0.80	-0.78	-0.034	-0.033	-0.032	-2.75	-2.72	-2.70	-0.98
	10.0	0.89	0.92	0.31	-0.57	-0.66	-0.60	-0.037	-0.035	-0.033	-2.60	-2.54	-2.49	-0.93
Gaussian with outliers	1.66	0.98	0.98	0.77	-0.98	-0.77	-0.74	-0.032	-0.032	-0.031	-2.78	-2.76	-2.74	-0.99
	3.22	0.98	0.98	0.63	-0.92	-0.83	-0.80	-0.030	-0.029	-0.029	-2.77	-2.75	-2.74	-0.98
	5.28	0.95	0.98	0.32	-0.74	-0.77	-0.73	-0.035	-0.034	-0.033	-2.58	-2.55	-2.52	-0.98
	7.01	0.93	0.96	0.28	-0.71	-0.71	-0.68	-0.035	-0.034	-0.033	-2.70	-2.66	-2.63	-0.96
	7.98	0.93	0.94	0.49	-0.61	-0.74	-0.71	-0.039	-0.037	-0.036	-2.34	-2.29	-2.25	-0.96
	15.64	0.80	0.87	0.03	-0.39	-0.63	-0.54	-0.043	-0.040	-0.037	-2.38	-2.30	-2.22	-0.87



Fig. 3 Location map of the investigated well sites, in a view of the northern hemisphere of the Earth from above the North Pole

two lithological factors. The first factor is responsible for 82 % of data variance, while the second factor explains the 18 % of observed information. In the previous study, the first factor log was identified as a shale indicator. Now a comparison is made between the first factor and hydraulic conductivity coming from grain-size analysis and porosity information. The factor loadings for well B-1 are listed in Table 6, where the GR and RS logs bear the heaviest influence on the first factor. Factor scores are scaled by the same procedure as in section ‘Synthetic modeling experiments’. Figure 4 shows the regression relationship between the first scaled factor and the decimal logarithm of hydraulic conductivity calculated by Eq. (1). Dominant grain-sizes are calculated by Eq. (2) from d_{10} and d_{60} values of grain-size distribution curves measured on core samples. Porosity in Eq. (1) was calculated from the neutron log using Eq. (15). The zone parameters for the neutron response can be obtained from the neutron-neutron vs. gamma-gamma crossplot (NNSD = 7.5 kcpm, NNSH = 4.0 kcpm, NNF = 1.0 kcpm). The regression

function suggested in Eq. (10) approximates well the relationship between the two variables, where the Pearson’s correlation coefficient ($R = -0.79$) indicates an unequivocal inverse proportionality. This index number is highly dependent on the level of data noise and the uncertainty of positioning the places of core sampling. The regression function is $\lg(K) = -0.046F_1^* - 3.38$, where coefficients $a = [-0.052, -0.041]$ and $b = [-3.66, -3.11]$ are estimated with 95 % level of significance.

The input borehole logs and the interpretation results are illustrated in Fig. 5. The GR log indicates the boundary between the Pleistocene and Pannonian (late Miocene) complex at around 240 m. The latter dominantly consists of clayey sands, silts and marls, while the former is formed mainly by coarse-grained sands and gravels. At the top of the lithology column Holocene flood sediments can be found. A GR-log-based method proposed by Larionov (1969) can be applied to give an independent estimate to shale volume (indicated by VSH in the last track). The formation porosity (POR) is extracted from the neutron log by Eq. (15), while the

Table 5 Correlation matrix of borehole logging data observed in well B-1

	RS	SP	GR	GG	NN
RS	1	-0.01	-0.74	-0.33	-0.06
SP	-0.01	1	-0.13	-0.42	-0.14
GR	-0.74	-0.13	1	0.36	0.05
GG	-0.33	-0.42	0.36	1	0.02
NN	-0.06	-0.14	0.05	0.02	1

Table 6 Factor loadings derived from borehole geophysical data collected from well B-1

Factor	$W^{(RS)}$	$W^{(SP)}$	$W^{(GR)}$	$W^{(GG)}$	$W^{(NN)}$
1st	-0.79	-0.41	0.91	0.36	0.03
2nd	0.21	0.85	-0.15	-0.24	-0.18

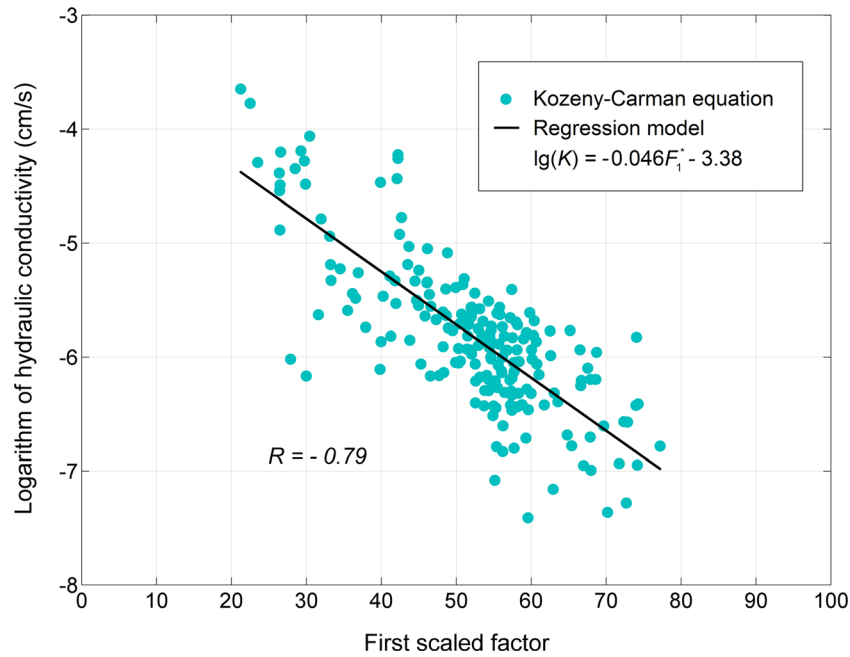


Fig. 4 Regression relation between the hydraulic conductivity calculated from core measurements and the scores of the first factor derived from borehole logs measured in well B-1

volume of sand is calculated by Eq. (18). As the entire interval is fully water-saturated, thus the water saturation in the flushed and the virgin zone is 100 % ($SX0 = SW = 1$). In the fifth track, the well log of the first scaled factor is plotted. Hydraulic conductivity $K(CORE)$ values indicated by red

circles are calculated by Eq. (1) in the places of core sampling. The same quantity derived from factor analysis is represented by the black curve of $K(FA)$. The hydraulic conductivity logs show a close fit to confirm the feasibility of the statistical method.

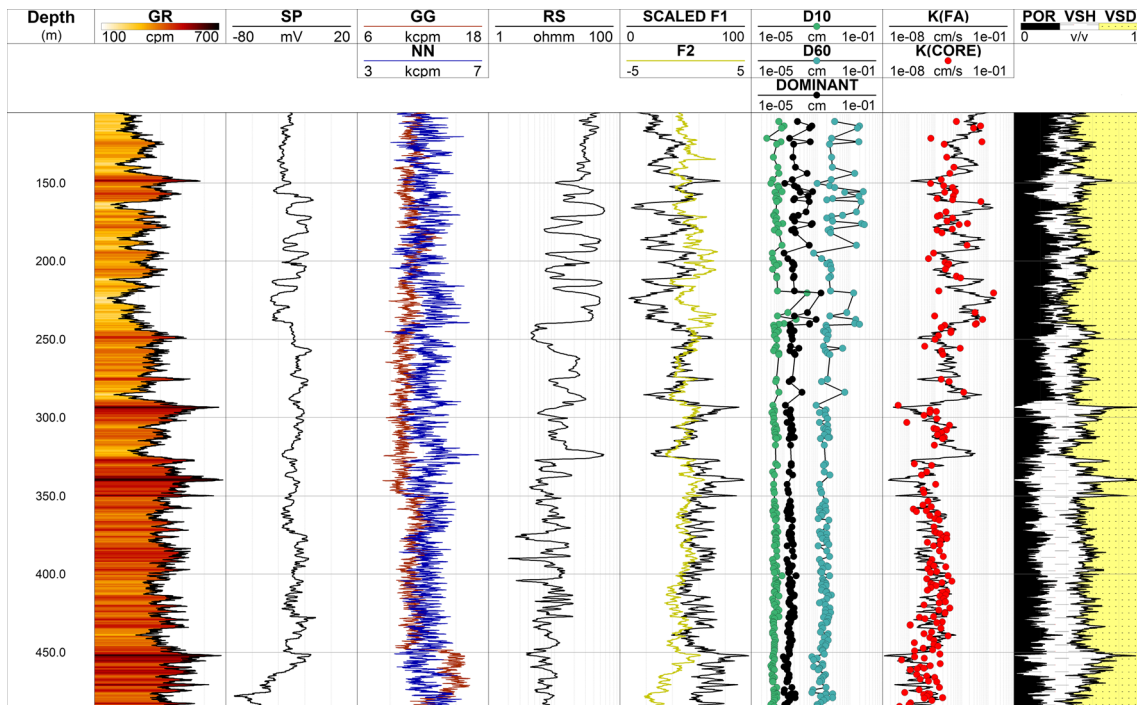


Fig. 5 Borehole logs measured from well B-1 (tracks 1–4). *GR* is natural gamma-ray intensity, *SP* is spontaneous potential, *GG* is gamma-gamma intensity, *NN* is neutron-neutron intensity, *RS* is shallow resistivity log. The well log of the first factor is in track 5. *D10* and *D60* are representative grain diameters at 10 and 60 % cumulative frequencies, *DOMINANT* is dominant grain diameter (track 6). Hydraulic conductivity (*K*) logs are estimated from grain-size analysis made on core samples (red circles) and factor analysis (black curve), separately (track 7). *POR* is porosity, *VSH* is shale content, *VSD* is sand volume (track 8)

Table 7 Correlation matrix of borehole geophysical data acquired in well FL-800

	GR	TEMP	RES16	RES64	LAT	NN-N	NN-F
GR	1	-0.03	-0.65	-0.50	-0.69	-0.52	-0.66
TEMP	-0.03	1	0.36	0.51	0.28	0.12	0.17
RES16	-0.65	0.36	1	0.92	0.98	0.46	0.56
RES64	-0.50	0.51	0.92	1	0.85	0.40	0.51
LAT	-0.69	0.28	0.98	0.85	1	0.47	0.57
NN-N	-0.52	0.12	0.46	0.40	0.47	1	0.59
NN-F	-0.66	0.17	0.56	0.51	0.57	0.59	1

Validation with aquifer test

The drilling of Well FL-800 was part of a project of US Geological Survey and US Environmental Protection Agency to characterize the lithostratigraphy and regional hydrogeology of the Ordovician Sinnipee Group situated in Waupun, Fond du Lac County, Wisconsin, USA (Fig. 3). The aim of the study was to acquire useful information to manage and protect the groundwater supply. For this task, detailed laboratory measurements made on core samples, borehole geophysical and hydrogeological surveys were performed at the investigation site (Dunning and Yeskis 2007). The Ordovician Sinnipee Group is considered a bedrock aquifer, which consists of the Platteville, Decorah and Galena Dolomite formations (from the bottom to the top). They are overlain only by a thin layer of unconsolidated deposits of Quaternary age. The principal lithology of the Sinnipee Group is dolomite and shaly dolomite in the logged interval. Core measurements indicate a primary porosity of 2–4 % in the dolomite, while primary porosity is up to 10 % in shaly intervals. The dolomite is quite massive and in some levels the fractures and bedding-plane partings are responsible for secondary porosity and permeability. These features were identified by the acoustic borehole televiewer, single-hole directional reflection and cross-hole radar tomography measurements. Borehole geophysical results showed that water is transmitted primarily along bedding-plane partings. In some isolated intervals, horizontal hydraulic conductivity was estimated by hydraulic tests, which proved to be highly consistent with heat-pulse flow-meter (flow velocity and volume) data. In this study, horizontal hydraulic conductivities calculated from slug tests using the method of Hvorslev (1951) are directly compared to the results of factor analysis.

Borehole logging measurements in well FL-800 suitable for factor analysis includes: natural gamma-ray intensity (GR), short- and long-spaced neutron-neutron intensity (NN-N for near receiver and NN-F for far receiver), electric resistivity (RES-16: short normal, RES-64: long normal, LAT: lateral) and temperature (TEMP) logs. The processed interval is between elevation 910 and 750 feet (277.4 and 228.6 m) above mean sea

level referenced to National Geodetic Vertical Datum of 1929 in the United States of America (NGVD 29). The total number of data is 11,298. Table 7 shows the correlation matrix of the borehole logs, which indicates moderate correlation between the measured variables on the average. Three factor logs are calculated, where SVD results show that 79.1 % of total variance is explained by the first factor, 14.9 % by the second factor and 6 % by the third factor. The factor loadings for well FL-800 are listed in Table 8, which shows some difference compared to the case of primary porosity rocks. The first factor appears to be principally a resistivity factor, while the second one is affected more by lithology and water content. In these fractured formations the first factor seems to be directly proportional to resistivity, while it is inversely related to natural gamma-ray log. After scaling the first factor as in section ‘Synthetic modeling experiments’, the logarithm of hydraulic conductivities obtained from slug tests can be plotted with respect to the factor scores. The regression relationship is illustrated in Fig. 6, which shows that the inverse connection is even stronger ($R = -0.90$) than in section ‘Validation with Kozeny-Carman model’. The regression function is $\lg(K) = -0.05F_1^* + 2.84$, where the coefficients estimated with 95 % significance bounds are $a = [-0.08, -0.02]$ and $b = [1.47, 4.21]$.

The boundary of Galena Dolomite and Decorah formations runs at 810 elevation in feet (246.9 m) and the top of the Platteville Formation is at around 790 elevation in feet (240.8 m). The GR log provides information on shale content, of which maximum is ~40 % in the Decorah Formation. The highest permeability zones are in the argillaceous dolomite above 870 elevation in feet (265.2 m), which was obtained from flow-meter tests. The acoustic televiewer images showed earlier also a high density of subvertical fractures and bedding-plane partings in the same interval. In the Decorah Formation, the larger porosity of shale causes a relative higher permeability than in the massive dolomite of Platteville Formation. The permeable zones are also indicated by the separation of resistivity logs of different penetrations. The well logs of the factors are plotted in the fifth track in Fig. 7. The logarithm of hydraulic

Table 8 Factor loadings derived from borehole geophysical data measured in well FL-800

Factor	$W^{(GR)}$	$W^{(TEMP)}$	$W^{(RES16)}$	$W^{(RES64)}$	$W^{(LAT)}$	$W^{(NN-N)}$	$W^{(NN-F)}$
1st	-0.47	0.11	0.81	0.59	0.87	0.21	0.20
2nd	-0.67	0.03	0.39	0.32	0.41	0.63	0.83
3rd	-0.02	0.59	0.44	0.73	0.29	0.11	0.17

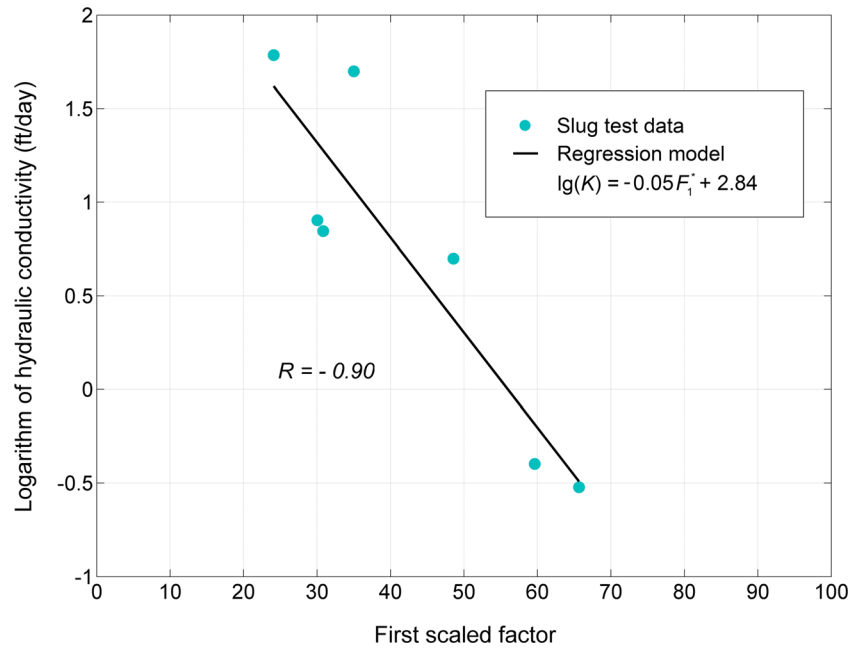


Fig. 6 Regression relation between the hydraulic conductivity calculated from aquifer tests and the scores of the first factor estimated from well logs observed in well FL-800

conductivity LG_K(ST) values represented by red circles are obtained from slug test results, while the continuous curve of LG_K(FA) is estimated from factor analysis. The overall fit between the results is sufficient to confirm the validity of the method.

In this section, the feasibility of factor analysis has been demonstrated for two different geological settings. An acceptable fit between the estimates of hydraulic conductivity derived separately from factor and core analysis was achieved in well B-1 (Fig. 8a). In the

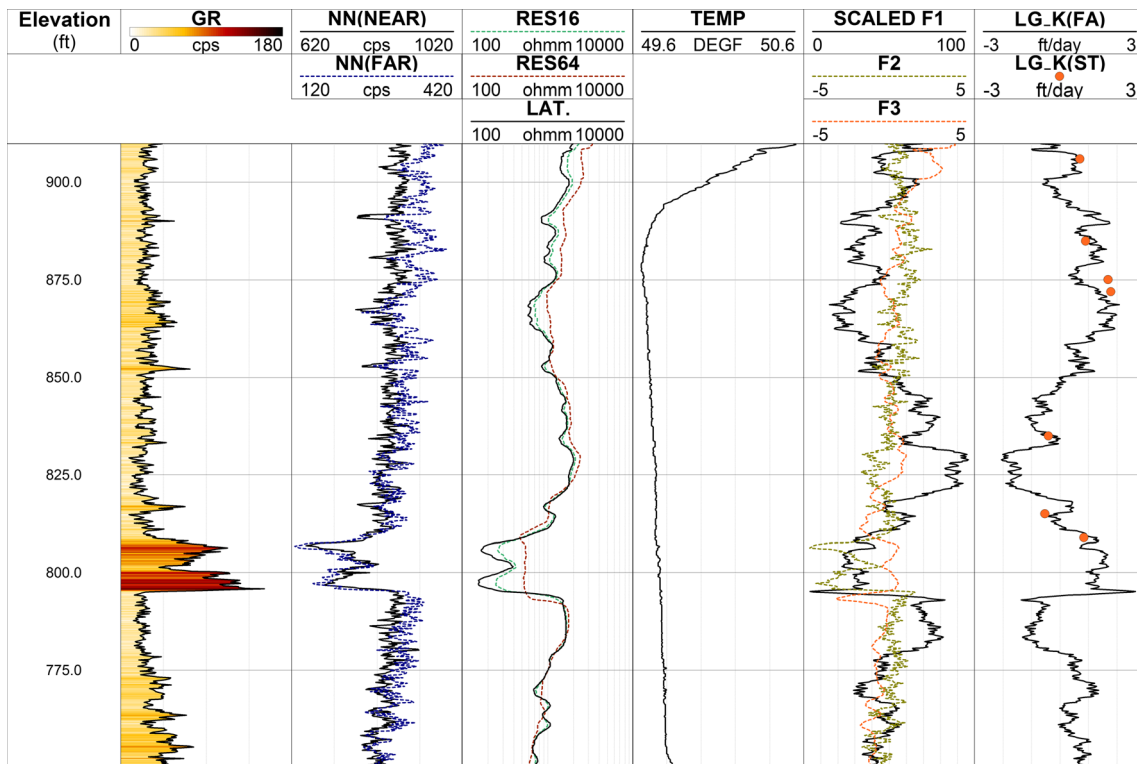


Fig. 7 Borehole logs measured from well FL-800 (tracks 1–4). GR is natural gamma-ray intensity, NN(NEAR) and NN(FAR) are short- and long spaced neutron-neutron intensity, RES-16 and RES-64 short and long normal resistivity, LAT is lateral resistivity, TEMP is temperature log. The extracted three factors are in track 5. Hydraulic conductivity (K) logs estimated from slug tests (red circles) and factor analysis (black curve), separately (track 6). Conversion: 1 m=3.2808 ft, 1 cm/s=2834.65 ft/day

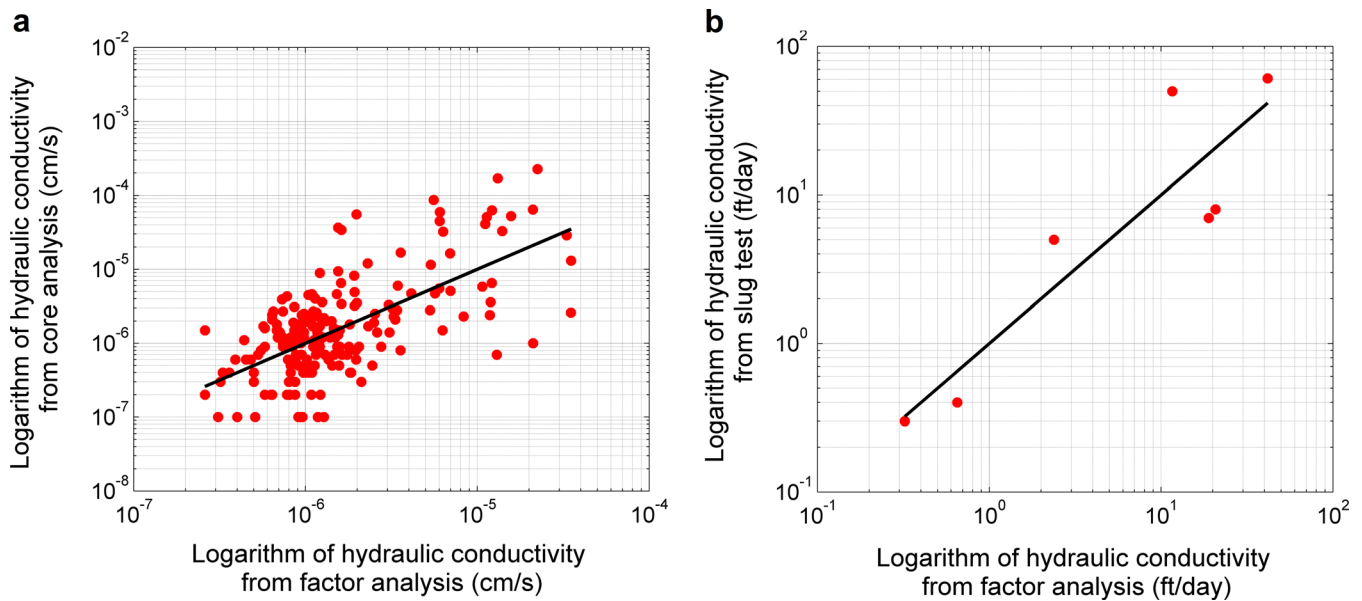


Fig. 8 Regression relation between hydraulic conductivity estimations. **a** The decimal logarithm of hydraulic conductivity was estimated from factor and laboratory analysis in well B-1. **b** The decimal logarithm of hydraulic conductivity was estimated from factor analysis and slug tests in well FL-800. Conversion: 1 cm/s=2,834.65 ft/day

Pleistocene, the matching points of the crossplot show bigger discrepancies from the black straight line representing the equation $\lg K(\text{FA}) = \lg K(\text{CORE})$, while the misfit is smaller in the lower hydraulic conductivity Miocene formation. In well FL-800, a closer fit between the estimates of factor analysis and slug tests is indicated for the entire depth interval (Fig. 8b), where the equality $\lg K(\text{FA}) = \lg K(\text{ST})$ is represented by black straight line in the figure. The accomplished synthetic and field experiments infer that the statistical method gives a reliable evaluation of hydraulic conductivity of groundwater formations.

Conclusions

A multivariate statistical approach applied to borehole logging data is suggested for the estimation of hydraulic conductivity. The factor analysis-based method has been proven to be a powerful tool in solving hydrogeophysical problems, for instance in earlier studies factor analysis was used to evaluate the shaliness of groundwater formations (Szabó 2011; Szabó and Dobróka 2013b; Szabó et al. 2014). As a continuation, a new statistical procedure is introduced in this paper to extract the vertical distribution of formation hydraulic conductivity purely from borehole geophysical observations. As a result, it provides continuous in-situ information to interpolate hydraulic conductivity between cored or pumped intervals. Synthetic modeling experiments using noisy data sets show that the estimation results are consistent, accurate and outlier resistant. Since the estimation error of permeability or hydraulic conductivity can normally reach one order of magnitude, any additional information

given from a different source may increase the reliability of the hydrogeological interpretation. The estimation error can be effectively reduced by the suggested method, because of a large number of observed variables and statistical sample. The method of factor analysis jointly processes suitable data within the logging interval allowing some shorter intervals of missing data. The process requires only a few seconds of CPU time using a quad-core processor-based workstation.

The paper also gives an overview on the practical issues on the application of the method. Numerical results confirm the validity of the regression formula defined in Eq. (10). The regression coefficients must be set for each well using core or aquifer test data. Even with the lack of this information, the formula gives a first approximation of hydraulic conductivity. On the basis of earlier results, the method is tested first in shallow clastic aquifers (section ‘Validation with Kozeny-Carman model’), where a strong inverse proportionality is shown between the first factor and hydraulic conductivity. Another field case is presented to demonstrate the feasibility of the method also in fractured sedimentary rocks (section ‘Validation with aquifer test’). Both studies confirm the validity of regression Eq. (10) where some deviations in the regression constants can be experienced because of the different measurement units of hydraulic conductivity. In sedimentary rocks with intergranular porosity, the first factor is mainly sensitive to shale content, where the factor loadings show positive correlation with the lithology logs (GR and SP). In these rocks the permeability is inversely proportional to shale content. In secondary or mixed porosity systems, where the mineral composition and textural (or structural) properties are more complex, the first factor may be influenced by other properties different

from lithology (e.g. pore-fluid content and secondary permeability). In section ‘Validation with aquifer test’, a massive dolomite complex is studied, in which shale volume is directly proportional to permeability. The loading of the first factor for the GR log appears to be negative and smaller than those of the resistivity logs. The interpretation of factors related to fractured rocks should be further analyzed in forthcoming research. The presented results show that reliable information on the distribution of hydraulic conductivity can be extracted from borehole geophysical data sets, which may improve the solution of hydrogeological problems. The method can be extended to multiple (two- or three-dimensional) borehole applications as it was shown earlier in Szabó et al. (2012) to support ground geophysical and hydrogeological surveys even in regional scales.

Acknowledgements The research was supported by the Hungarian Scientific Research Found (project number PD 109408). The author is thankful for the support of the János Bolyai Research Fellowship of the Hungarian Academy of Sciences. Special thanks go to Charles P. Dunning PhD for permission to use the data set of Well FL-800, and Professor Mihály Dobróka for his valuable scientific advice and inspiring co-operation.

References

- Alberty MW, Hashmy KH (1984) Application of ULTRA to log analysis. Paper Z, SPWLA 25th Annual Logging Symposium, New Orleans, LA, 10–13 June 1984, pp 1–17
- Alger RP (1966) Interpretation of electric logs in fresh water wells in unconsolidated formations. SPWLA 7th Annual Logging Symposium, Tulsa, OK, 9–11 May 1966, pp 1–25
- Alger RP, Harrison CW (1989) Improved fresh water assessment in sand aquifers utilizing geophysical well logs. *Log Anal* 30(1):31–44
- Asfahani J (2014) Statistical factor analysis technique for characterizing basalt through interpreting nuclear and electrical well logging data (case study from southern Syria). *Appl Radiat Isot* 84:33–39
- Bartlett MS (1950) Tests of significance in factor analysis. *Br J Psychol* 3(2):77–85
- Bartlett M S (1953) Factor analysis in psychology as a statistician sees it. *Nordisk Psykologi's Monograph Series 3*, Almqvist and Wiksell, Uppsala, Sweden, pp 23–34
- Bear J (1972) *Dynamics of fluids in porous media*. Dover, New York
- Benson CH, Trast JM (1995) Hydraulic conductivity of thirteen compacted clays. *Clay Clay Miner* 43(6):669–681
- Buffin A (1996) Permeability from waveform sonic data in the Otway basin. SPWLA 37th Annual Logging Symposium, New Orleans, LA, 16–19 June 1996, pp 1–11
- Csókás J (1995) Determination of water discharge and quality using geophysical well logs (in Hungarian). *Magyar Geofiz* 35(4):176–203
- Dobróka M, Gyulai Á, Ormos T, Csókás J, Dresen L (1991) Joint inversion of seismic and geoelectric data recorded in an underground coal mine. *Geophys Prospect* 39:643–665
- Drahos D (2005) Inversion of engineering geophysical penetration sounding logs measured along a profile. *Acta Geodetica Geophys Hung* 40:193–202
- Dunning CP, Yeskis DJ (2007) Lithostratigraphic and hydrogeologic characteristics of the Ordovician Sinnipee Group in the vicinity of Waupun, Fond du Lac County, WI, 1995–96. *US Geol Surv Sci Invest Rep* 2007–5114, 51 pp
- Glover PWJ, Walker E (2009) Grain-size to effective pore-size transformation derived from electrokinetic theory. *Geophysics* 74(1):E17–E29
- Guérin R (2005) Borehole and surface-based hydrogeophysics. *Hydrogeol J* 13:251–254
- Hvorslev MJ (1951) Time lag and soil permeability in ground-water observations. *Bulletin 36*, Waterways Experimentation Station, US Army Corps of Engineers, Vicksburg, MI
- Idrissy EHE, De Smedt F (2007) A comparative study of hydraulic conductivity estimations using geostatistics. *Hydrogeol J* 15:459–470
- Jöreskog KG (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34(2):183–202
- Jöreskog KG (2007) Factor analysis and its extensions. In: Cudeck R, MacCallum RC (eds) *Factor analysis at 100: historical developments and future directions*. Erlbaum, Hillsdale, NJ
- Juhász J (2002) *Hydrogeology (in Hungarian)*. Akadémiai Kiadó, Budapest
- Kaiser HF (1958) The varimax criterion for analytical rotation in factor analysis. *Psychometrika* 23:187–200
- Khalil MA, Ramalho EC, Monteiro Santos FA (2011) Using resistivity logs to estimate hydraulic conductivity of a Nubian sandstone aquifer in southern Egypt. *Near Surf Geophys* 9(4):349–355
- Larionov VV (1969) *Radiometry of boreholes (in Russian)*. Nedra, Moscow
- Lawley DN, Maxwell AE (1962) Factor analysis as a statistical method. *Statistician* 12:209–229
- Odong J (2013) Evaluation of empirical formulae for determination of hydraulic conductivity based on grain-size analysis. *Int J Agric Environ* 1:1–8
- Perdomo S, Ainchil JE, Kruse E (2014) Hydraulic parameters estimation from well logging resistivity and geoelectrical measurements. *J Appl Geophys* 105:50–58
- Ross J, Ozbek M, Pinder GF (2007) Hydraulic conductivity estimation via fuzzy analysis of grain size data. *Math Geol* 39:765–780
- Roy J, Lubczynski M (2003) The magnetic resonance sounding technique and its use for groundwater investigations. *Hydrogeol J* 11:455–465
- Sallam OM (2006) Aquifers parameters estimation using well log and pumping test data, in arid regions: step in sustainable development. *The 2nd International Conference on Water Resources and Arid Environment*, Riyadh, Saudi Arabia, 26–29 November 2006, pp 1–12
- Shevnin V, Delgado-Rodríguez O, Mousatov A, Ryjov A (2006) Estimation of hydraulic conductivity on clay content in soil determined from resistivity data. *Geofisica Int* 45(3):195–207
- Slater L (2007) Near surface electrical characterization of hydraulic conductivity: from petrophysical properties to aquifer geometries—a review. *Surv Geophys* 28:169–197
- Szabó NP (2011) Shale volume estimation based on the factor analysis of well-logging data. *Acta Geophys* 59:935–953
- Szabó NP, Dobróka M (2013a) Float-encoded genetic algorithm used for the inversion processing of well-logging data. In: Michalski A (ed) *Global optimization: theory, developments and applications*. Mathematics Research Developments, Computational Mathematics and Analysis Series, Nova Science, New York
- Szabó NP, Dobróka M (2013b) Extending the application of a shale volume estimation formula derived from factor analysis of wireline logging data. *Math Geosci* 45(7):837–850
- Szabó NP, Kalmár CS (2013) Nonlinear regression model for permeability estimation based on acoustic well-logging measurements. *Geosci Eng* 2(4):27–46
- Szabó NP, Dobróka M, Drahos D (2012) Factor analysis of engineering geophysical sounding data for water saturation estimation in shallow formations. *Geophysics* 77(3):WA35–WA44
- Szabó NP, Dobróka M, Turai E, Szűcs P (2014) Factor analysis of borehole logs for evaluating formation shaliness: a

- hydrogeophysical application for groundwater studies. *Hydrogeol J* 22:511–526
- Timur A (1968) An investigation of permeability, porosity, and residual water saturation relationships. SPWLA-1968-J, SPWLA 9th Annual Logging Symposium, New Orleans, LA, 23–26 June 1968
- Walsh D, Turner P, Grunewald E, Zhang H, Butler JJ, Reboulet E, Knobbe S, Christy T, Lane JW, Johnson CD, Munday T, Fitzpatrick A (2013) A small-diameter NMR logging tool for groundwater investigations. *Groundwater* 51:914–926
- Wentworth CK (1922) A scale of grade and class terms for clastic sediments. *J Geol* 30:377–392