Permeability scale effect in sandy aquifers: a few case studies

Effet d’échelle et perméabilité des aquifères sableux : quelques études de cas

Chapuis R.P.
École Polytechnique, Montréal, QC, Canada

ABSTRACT: In sandy aquifers, stratification results in a range of values for the hydraulic conductivity *K*, which can be evaluated at three scales. Since large-scale tests are more likely to meet preferential flow paths, they are also likely to yield larger *K* values than small-scale tests, which may be viewed as a scale effect. The small scale is that of soil samples: their quality must be assessed and their grain size distribution analyzed to check for mixes of sub-layers, before using reliable methods to predict the *K* values. The middle scale is that of field permeability tests for which it is important to respect the standards and perform verifications. The large scale is that of pumping tests. The paper presents a few case studies of sandy aquifers. Their stratification led to unimodal or multimodal grain size distributions. For all cases, the *K* distributions provided consistent images of the sandy aquifers. It was then concluded that, after a quality control of data and interpretations, there was no scale effect in the aquifers.

RÉSUMÉ : Dans les aquifères sableux, la stratification donne une gamme de valeurs pour la conductivité hydraulique *K* qui peut être évaluée à trois échelles. Les essais à grande échelle ayant plus de chances de tester des zones d’écoulement préférentiel, ils ont aussi plus de chances de donner des valeurs élevées de *K* que les essais à petite échelle, ce qui peut être vu comme un effet d’échelle. La petite échelle est celle des échantillons : leur qualité doit être évaluée et leur granulométrie analysée pour détecter les mélanges de strates, avant d’utiliser des méthodes fiables de prédiction de *K*. L’échelle moyenne est celle des essais de perméabilité in situ pour lesquels on doit respecter les normes et faire des vérifications. La grande échelle est celle des essais de pompage. L’article présente des études de cas d’aquifères sableux. Leur stratification a donné des granulométries unimodales ou multimodales. Pour tous les cas, les distributions de *K* ont fourni des images cohérentes des aquifères sableux. On a conclu, après un contrôle de qualité des données et des interprétations, qu’il n’y avait pas d’effet d’échelle dans ces aquifères.

KEYWORDS: aquifer, grain size distribution, monitoring well, permeability test, pumping test, scale effect

1 INTRODUCTION

In sandy aquifers, groundwater seepage is controlled by stratification, with coarse size sediments deposited at high water velocities and small size sediments settling at low water velocities, or in temporary ponds. Many methods can be used to assess the hydraulic conductivity, *K*, which can vary over orders of magnitude. It is often believed that since large-scale tests involve large volumes, which are more likely to meet preferential flow paths, they are likely to yield larger *K* values than small-scale tests (Bradbury and Muldoon 1990; Rovey 1998; Rovey and Niemman 1998). Thus, there should be a scale effect for the *K* value, some increase with the tested volume.

There is no consensus about this scale effect. Many studies tried to check or challenge theoretical opinions. They differed about testing techniques, investigated scales, and geologic media. Alas, the quality of each *K* value usually was not questioned even if poor quality data and interpretation are known to yield an artificial scale effect. Regrettably, the quality control of groundwater parameters, which must be methodically completed for engineered facilities, is not always done (Chapuis 1995). This paper examines quality control issues with data and interpretation, in order to exclude artificial scale effects.

The idea of scale effect was rejected by Butler and Healey (1998). They argued that scale effect results from artifacts linked to incomplete well development and low-*K* skins around well screens, but they did not study what produce a positive or negative skin. These skin phenomena and their effects on the apparent *K* value being related to safety issues, they are more studied in geotechnique (Chapuis and Chenaf 2010) than in geosciences.

Moreover, many studies have not examined how incorrect interpretation methods for slug tests and pumping tests can yield artificial scale effects. However, the quality control of slug test methods has been largely investigated in geotechnique (Chapuis et al. 1981; Chapuis 1988, 1998, 1999, 2001; Chapuis and Chenaf 2002, 2003). For pumping tests in unconfined aquifers, the large-scale *K* values obtained were shown to be incorrect if the interpretation was performed using current methods for unsteady-state (Akindunni and Gilham 1992).

Therefore, when studying scale effect, some caution must be observed to avoid using scale effect as a final excuse, or as a fudge factor, when the heterogeneity of the tested material could have been more thoroughly investigated and when errors involved in sampling, testing and interpretation methods could have been taken into account.

Note that properly taking into account scale effect is important for numerical analyses, since an aquifer numerical model cannot be as detailed as the physical reality. Most often, the grids of numerical models cannot contain enough elements to model the detail of real features. This is why up-scaling techniques are needed to define some equivalent *K* value for grid elements (Renard and de Marsily 1997; Zhang et al. 2011).

In this paper, the results of three sites are briefly examined. The small scale, about 10^{-3} m^3, is that of samples recovered in boreholes for which the *K* value was evaluated using predictive methods. The middle scale, about 1 m^3, is that of field permeability tests in monitoring wells. The large scale, about 10^3 m^3, is that of pumping tests. Now, the problems linked to the collected data at three scales in sandy aquifers are examined in detail, starting with the soil samples taken in boreholes.

2 SMALL-SCALE *K* VALUES (SAMPLES)

Many soil samples can be taken in boreholes, usually with a split spoon. Quality issues relative to soil sampling have been the topic of many geotechnical researches. Five sample classes are defined by considering the relationships between sampling methods, quality of sample and quality of laboratory tests. All borehole samples in sandy aquifers are of class-3 or class-4 quality. For information, the class-4 quality is obtained with the hollow stem auger, rotary, percussion, cable tool and sonic drilling methods (Baldwin and Gosling 2009). These methods strongly influence not only the quality of samples, but also the quality of permeability tests, and the quality of the MW installation (Chapuis and Sabourin 1989; Chesnaux et al. 2006; Chesnaux and Chapuis 2007). In sandy aquifers, a tube sampler with a clear plastic liner can be used. This tool does not provide class-1 or -2 samples. It roughly preserves the grain size distribution curve (GSDC), with major margin disturbance (thick-walled sampling) plus some mixing between adjacent sub-layers. It does not preserve the water content w, void ratio e, and K in situ values. For that reason, this sampler provides class-4 samples, and not intact ones as claimed in a few papers.

Several methods can be used to predict the K value of a soil sample. Chapuis (2012a) listed 45 methods and assessed their capacity against large data sets for laboratory permeability tests performed on homogenized fully saturated specimens. All tests were not plagued by one of the 14 most frequent mistakes when performing such tests. For sandy aquifers, the in situ porosity n can be assessed using the method of Chapuis (2012b) and the K values can be predicted with the method of Chapuis (2004), which yields good predictions for natural soils in the ranges $0.003 \leq d_{10} \leq 3$ mm and $0.3 \leq e \leq 1$. The range for the effective diameter d_{10} was recently extended up to 150 mm (Côté et al. 2011; Chapuis et al. 2012).

If the soil sample is homogenous, its GSDC is smooth. This is not the case for most borehole samples in sandy aquifers. Therefore, when studying the GSDCs, caution must be taken to avoid confusing homogenous samples (single layer) with those made by mixing 2 or 3 small layers. The analysis proceeds with a modal decomposition (Chapuis 2010; Chapuis et al. 2013), which provides the GSDC and percentage of each layer in the composite sample. The equivalent horizontal K value (stratified sample) is then obtained using the composition rule.

3 MEDIUM-SCALE *K* VALUES (SLUG TESTS IN MWS)

The middle scale, about 1 m^3, is that of permeability tests (slug tests) performed in monitoring wells.

It is important to use the standard methods to interpret the slug test data. In Canada, CAN/BNQ 2501-135 is the standard for an overdamped response (CAN/BNQ 1988, 2008), but there is no standard for an underdamped response. ASTM, however, has standards for the underdamped response (ASTM 2012a) and for the critically damped response (ASTM 2012b).

For overdamped slug tests, the velocity graph method helps to establish the correct piezometric level (PL) and K value for the test. It also helps to detect several phenomena during the test. Even if the aquifer is unconfined, and even if the MW is correctly installed, there are several reasons why the test data must be corrected by a systematic error on the assumed PL, of a few centimetres (Chapuis 2009a, b). The velocity graph gets rid of any systematic error, which may be due to incorrect calibration of a pressure transducer (PT), waiting time, PT line slippage, piezometric modification, faulty MW installation, and unknown PL. However, it cannot make a distinction between these six errors.

For underdamped slug tests, it is preferable to fit the test data using a least squares method, instead of a visual fit, and the verification of three physical conditions must be done for each tests, otherwise large errors can be made (Chapuis 2012c).

4 LARGE-SCALE *K* VALUES (PUMPING TESTS)

For the large scale of pumping tests, about 10^3 m^3, precautions must be taken when installing the pumping well and MWs, and also when interpreting the pumping test data. The common theories for unsteady-state are based on some wishful thinking about drainage, unsaturated seepage and a misleading concept of specific yield (e.g., Akindunni and Gilham 1992; Chapuis et al. 2005a). For MWs, it is commonly admitted that two thirds of them are improperly installed (Nielsen and Schalla 2005).

5 THE SITES

5.1 *The Lachenaie site*

The site is located 50 km north-east of Montreal. The sand unconfined aquifer has been used for field training and research. The GSDCs could be correctly fitted using a unimodal lognormal distribution. The little variability for the mean and the standard deviation indicate homogeneity (Fig. 1). For the pumping test, the steady-state drawdown data were used, the interpretation methods being proven to be reliable (Chapuis et al. 2005a, b). In this aquifer, the average K values at the three scales are very close, and thus there is no scale effect (Fig. 2).

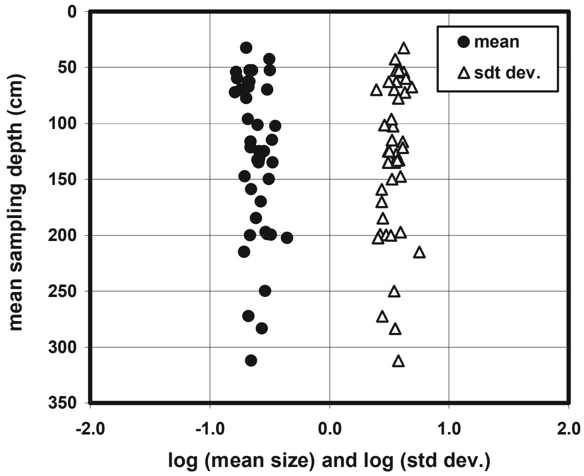

Figure 1. Lachenaie: modal decomposition of the sand GSDCs.

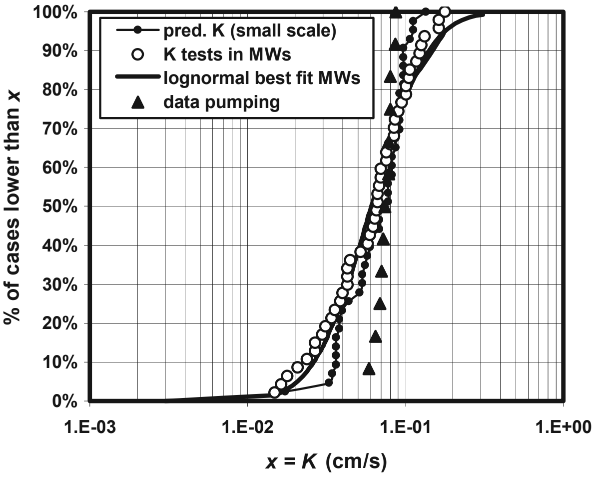

Figure 2. Lachenaie: comparison of the K values obtained at the small, medium, and large scales.

5.2 *The Shannon site*

Shannon is a small town about 30 km north-west of Quebec City. For the TCE-contamination case, a lot of information was given in the defendants' expert reports, but without a quality control, which led to contradictions. The quality control and a synthesis were done in Chapuis (2009c, 2010, 2013a, b). There were about 1000 MWs for this huge contamination case.

The aquifer stratification could be considered or not when analyzing the GSDCs. When it was not, the distribution of predicted *K* values could not explain the high large-scale *K* values of pumping tests (Fig. 3). When it was, after using a modal decomposition of each GSDC, the distribution of predicted *K* values yielded a large-scale *K* value very close to that of pumping tests (Chapuis 2010, 2013b).

For the slug tests in MWs, Chapuis (2010) showed that the defendants' expert reports gave *K* values that were obtained without following the standards and without making the required verifications. They were about three times smaller than the *K* values obtained when following the standards and making the verifications. When the standards were not respected, the distribution of the slug test *K* values could not explain the large-scale *K* values of pumping tests (Fig. 4). When the standards were followed, the slug test *K* distribution yielded a large-scale *K* value very close to that of pumping tests (Fig. 4).

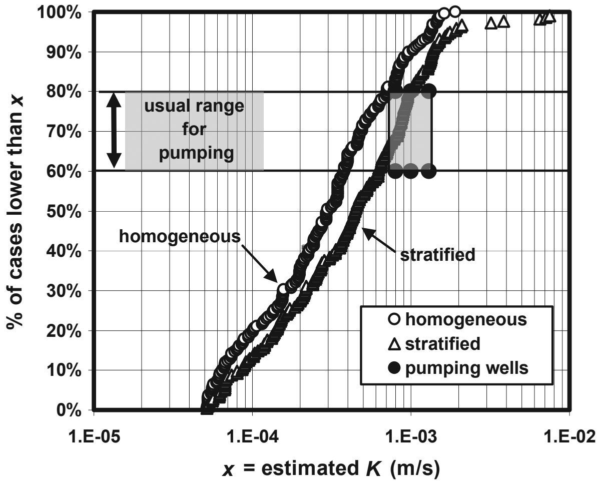

Figure 3. Shannon: *K* values predicted using the GSDCs, assuming either homogeneous or stratified samples (modal decomposition) and large scale pumping tests.

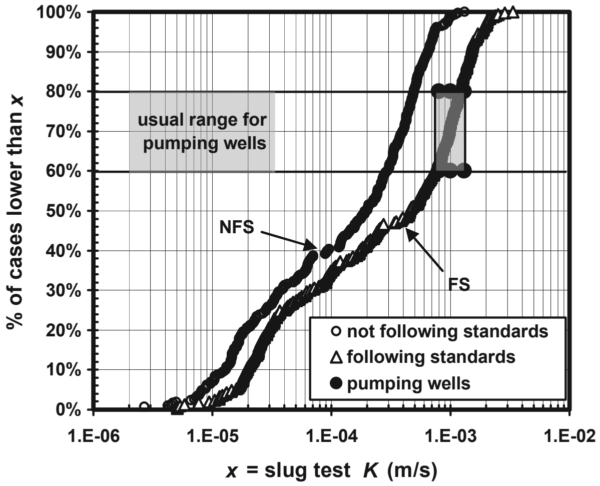

Figure 4. Shannon: *K* values obtained with slug tests in monitoring wells (following or not the standards), and large scale pumping tests.

How to perform the modal decomposition of a GSDC, and that of predicted or measured *K* values, is explained elsewhere (Chapuis 2013b; Chapuis et al. 2013). These papers also explain how to predict, for a *K* distribution, the large-scale *K* value which would be given by a pumping test, in order to logically compare the data at the three scales. A closed-form equation is also provided for the soil specific surface, more general than that of Chapuis and Légaré (1992). According to the detailed study following the quality control for the Shannon aquifer, all *K* distributions provided a coherent image of the hydraulic properties in the aquifer. Therefore, there was no scale effect.

5.3 *The Sorel site*

The Sorel site, 100 km north-east of Montreal, has been used for many years for field training of students in groundwater engineering and geophysics. The site is part of the floodplain at the confluence of the Richelieu River and the St-Lawrence River. Down to about 5 m deep, the stratigraphy includes many layers of fine sand (deposited in low velocity water) and silty clay (deposited in ponds). Over 300 soil samples were recovered in over 40 boreholes.

The soil samples provided clearly bimodal GSDCs and *K* values (Chapuis et al. 2013). The split-spoon sampler could recover 30 or more individual layers of silty clay and fine sand, which were uniform in color. The GSDC modal decomposition provided results such as those of Fig. 5 for a few boreholes in the vicinity of the pumping well. The fine sand and silty clay were fairly homogeneous (Fig. 5). According to the modal decompositions, the portion between 1.9 and 3.1 m deep had more clayey silt than the upper and lower portions. The screens of the pumping well and nearby MWs were installed in the portion between 3.1 and 4.4 m this confined aquifer.

The horizontal *K* distribution curve was obtained from the modal decomposition of GSDCs and the *K* composition rule. The predicted *K* distribution was in good agreement with the pumping test *K* values, whereas the slug test *K* values were somewhat below the pumping test *K* values (Chapuis et al. 2013). Due to the fine stratification of fine sand and silty clay sub-layers, the development of monitoring wells was not effective. Therefore, the slug tests have slightly underestimated the horizontal medium-scale *K* value due to smearing between layers during drilling and MW installation. Therefore, there was no scale effect for the Sorel highly stratified aquifer.

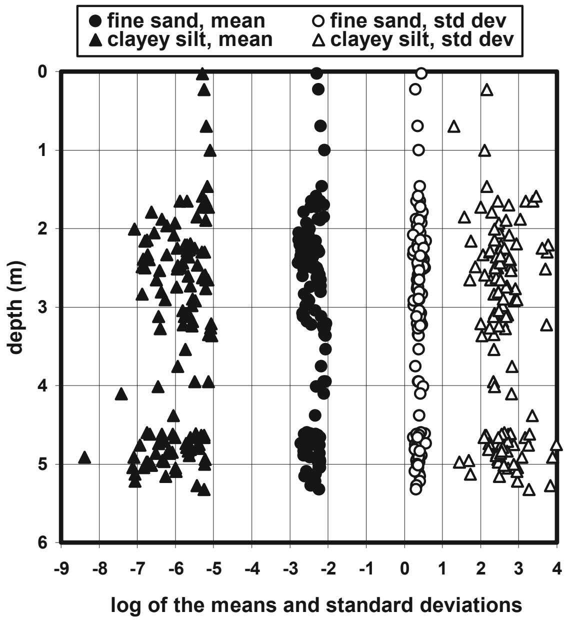

Figure 5. GSDCs modal decomposition for stratified samples of Sorel, showing fairly homogeneous layers of fine sand and clayey silt.

6 CONCLUSIONS

This paper studies the permeability of sandy aquifers at three scales. The aquifers are stratified or not, which leads to multimodal or unimodal distributions for grain size distribution curves. The small scale is that of soil samples: their quality must be assessed and their GSDC analyzed to check for mixes of sub-layers before using reliable methods to predict the *K* values. The middle scale is that of field permeability tests for which it is important to respect the standards and perform verifications. The large scale is that of pumping tests, which must be interpreted for steady-state.

The results presented here have shown that, when stratification is adequately considered, slug tests are interpreted according to standards, and the resulting *K* distributions are taken into account, the conclusion is that there is no scale effect.

Therefore, using a quality control approach for analyzing the GSDCs and interpreting field test data is essential for cross-checks, and for avoiding the creation of artificial scale effects.

7 ACKNOWLEDGEMENTS

The author thanks the National Research Council of Canada, BFI Ltd (Lachenaie), the Régie de l'eau de Sorel-Tracy and Aquatech (Sorel), the Shannon Citizen's Committee, Charles Veilleux and Associates, and the FARC (Shannon).

8 REFERENCES

Akindunni F.F. and Gillham R.W. 1992. Unsaturated and saturated flow in response to pumping of an unconfined aquifer: Numerical investigation of delayed drainage. Ground Water 30, 873–884.

ASTM 2012a. D5785: Standard test method (analytical procedure) for determining transmissivity of confined nonleaky aquifers by underdamped well response to instantaneous change in head (slug test). Annual CDs of standards, vol. 04.09, ASTM, West Conshohocken, Penn.

ASTM 2012b. D5881: Standard test method (analytical procedure) for determining transmissivity of confined nonleaky aquifers by critically damped well response to instantaneous change in head (slug test). Annual CDs of standards, vol. 04.09, ASTM, West Conshohocken, Penn.

Baldwin M. and Gosling D. 2009. BS EN ISO 22475-1: Implications for geotechnical sampling in the UK. Ground Engineering, August 2009, 28–31.

Bradbury K.R. and Muldoon M.A. 1990. Hydraulic conductivity determinations in unlithified glacial and fluvial materials. ASTM STP 1053, 138–151.

Butler J.J. and Healey J.M. 1998. Relationship between pumping test and slug-test parameters: scale effect or artefacts. Ground Water 36(2), 305–313.

CAN/BNQ 1988. Canadian Standard CAN/BNQ 2501-135/1988: Soils – determination of permeability by the Lefranc method.

CAN/BNQ 2008. Canadian Standard CAN/BNQ 2501-135/2008: Soils – determination of permeability by the Lefranc method.

Chapuis R.P. 1995. Controlling the quality of ground water parameters: some examples. Can Geotech J 32(1), 172–177.

Chapuis R.P. 1988. Determining whether wells and piezometers give water levels or piezometric levels. In Ground Water Contamination: Field Methods, ASTM STP 963, 162–171

Chapuis R.P. 1998. Overdamped slug test in monitoring wells: Review of interpretation methods with mathematical, physical, and numerical analysis of storativity influence. Can Geotech J 35(5), 697–719.

Chapuis R.P. 1999. Borehole variable-head permeability tests in compacted clay liners and covers. Can Geotech J 36(1), 39–51.

Chapuis R.P. 2001. Extracting the local piezometric level and hydraulic conductivity from tests in driven flush-joint casings. Geotech Testing J 24(2), 209–219.

Chapuis R.P. 2004. Predicting the saturated hydraulic conductivity of sand and gravel using effective diameter and void ratio. Can Geotech J 41(5), 787–795.

Chapuis R.P. 2009a. Permeability or hydraulic conductivity tests in a monitoring well: Why are piezometric level corrections required? Geotech News 27(2), 46–49.

Chapuis R.P. 2009b. Interpreting slug tests with large data sets. Geotech Testing J 32(2), 139–146.

Chapuis R.P. 2009c. Recours collectif - Résidents de Shannon - Expertise sur les conditions hydrogéologiques - Rapport préliminaire en 3 volumes, 970 p.

Chapuis R.P. 2010. Recours collectif - Résidents de Shannon – Expertise sur les conditions hydrogéologiques - Rapport d'expertise, 156 p.

Chapuis R.P. 2012a. Predicting the saturated hydraulic conductivity of soils: A review. Bull Eng Geology Envir 71(3), 401–434.

Chapuis R.P. 2012b. Estimating the in situ porosity of sandy soils sampled in boreholes. Engng Geology 141–142, 57–64.

Chapuis R.P. 2012c. Improved curve fitting methods for underdamped slug tests. Geotech Testing J 35(5), 752–761.

Chapuis R.P. 2013a. TCE-contaminated groundwater in Shannon, Quebec: 2. Evaluating the hydraulic conductivity with permeability tests in observation wells. Bull Eng Geol Env, submitted

Chapuis R.P. 2013b. TCE-contaminated groundwater in Shannon, Quebec: 3. Evaluating the hydraulic conductivity at three scales. Bull Eng Geol Env, submitted

Chapuis R.P. and Chenaf D. 2002. Slug tests in a confined aquifer: Experimental results in a large soil tank and numerical modeling. Can Geotech J 39(1), 14–21.

Chapuis R.P. and Chenaf D. 2003. Variable-head field permeability tests in driven casings: Physical and numerical modeling. Geotech Testing J 26(3), 245–256.

Chapuis R.P. and Chenaf D. 2010. Driven field permeameters: Reinventing the wheel? Geotechnical News 28(1), 37–42.

Chapuis R.P. and Légaré P.P. 1992. A simple method for determining the surface area of fine aggregates and fillers in bituminous mixtures. In *Effects of Aggregates and Mineral Fillers on Asphalt Mixture Performance*, ASTM STP 1147, 177–186.

Chapuis R.P. and Sabourin L. 1989. Effects of installation of piezometers and wells on groundwater characteristics and measurements. Can Geotech J 26(4), 604–613.

Chapuis R.P., Paré J.J., and Lavallée J.G. 1981. Essais de perméabilité à niveau variable. Proc. 10th ICSMFE, Stockholm, Balkema, Vol. 1, 401–406.

Chapuis R.P., Chenaf D., Acevedo N., Marcotte D. and Chouteau M. 2005a. Unusual drawdown curves for a pumping test in an unconfined aquifer at Lachenaie, Quebec: Field data and numerical modeling. Can Geotech J 42, 1133–1144.

Chapuis R.P., Dallaire V., Marcotte D., Chouteau M., Acevedo N. and Gagnon F. 2005b. Evaluating the hydraulic conductivity at three different scales within an unconfined aquifer at Lachenaie, Quebec. Can Geotech J 42, 1212–1220.

Chapuis R.P., Weber S. and Duhaime F. 2012. Intrinsic permeability of materials ranging from sand to rock-fill using natural air convection tests: Discussion. Can Geotech J 49(11), 1319–1322.

Chapuis R.P., Dallaire V. and Saucier A. 2013. Getting information from modal decomposition of grain size distribution curves. Geotech Testing J, submitted.

Chesnaux R. and Chapuis R.P. 2007. Detecting and quantifying leakage through defective borehole seals: A new methodology and laboratory verification. Geotech Test J 30(1), 17–24.

Chesnaux R., Chapuis R.P. and Molson J.W. 2006. A new method to characterize hydraulic short-circuits in defective borehole seals. Ground Water 44(5), 676–681.

Côté J., Fillion M.H. and Konrad J.M. 2011. Intrinsic permeability of materials ranging from sand to rock-fill using natural air convection tests. Can Geotech J 48, 679–690.

Nielsen D.M. and Schalla R. 2005. Design and installation of ground-water monitoring wells. Chapter 10, *Practical Handbook of Environmental Site Characterization and Ground–Water Monitoring*, 2nd edition, CRC Taylor & Francis.

Renard P. and de Marsily G. 1997. Calculating equivalent permeability: a review. Adv Water Resources 20(5-6), 253–278.

Rovey C.W. II 1998. Digital simulation of the scale effect in hydraulic conductivity. Hydrogeology, 6(2), 216–225.

Rovey C.W. II and Niemann W.L. 1998. Wellskins and slug tests: where's the bias? J Hydrology 243(1-2), 120–132.

Zhang Y., Liu B.Z. and Gable C.W. 2011. Homogenization of hydraulic conductivity for hierarchical sedimentary deposits at multiple scales. Transp Porous Med 87, 717–737.