Secondary data value for geostatistical estimation of flow and transport parameters in low-permeability media

M. Huysmans[1], A. Dassargues[1,2]

1 Applied Geology and Mineralogy, Department of Geology-Geography, Katholieke Universiteit Leuven, Belgium
2 Hydrogeology Group, Departement GeomaC, Université de Liège, Belgium
Corresponding author: marijke.huysmans@geo.kuleuven.be

ABSTRACT : This study investigates the relative data value of different types of secondary information to estimate the flow and transport parameters in low-permeability media. Data from four boreholes in two different clay formations are analyzed. Cross-validation is used to estimate the relevance of the different types of secondary data. The flow and transport parameters (hydraulic conductivity, diffusion coefficient and diffusion accessible porosity) are first estimated by kriging using primary information only and then estimated by co-kriging using primary information and one type of secondary information (grain size, gamma ray and resistivity). The kriging residuals, i.e. the difference between the estimated and the measured values, of the different kriging and co-kriging variants are compared to evaluate the relevance of each type of secondary information. This analysis shows that in the different clays and the different boreholes, different types of secondary information result in the largest improvement of the estimates of the flow and transport parameters. It is not possible to determine the "best" type of secondary information for improving the estimation of the flow and transport parameters of low-permeability media. This probably depends on local factors such as the quality of the loggings and the amount of nearby primary and secondary data.

KEYWORDS : geostatistics, hydrogeology, low-permeability media.

1. Introduction

Understanding and modeling flow and transport in low-permeability media is essential for many environmental applications: nuclear waste storage, carbon-dioxide sequestration, landfilling of municipal and industrial waste, etc. Heterogeneity of the main flow and transport parameters (hydraulic conductivity, diffusion coefficient and diffusion accessible porosity) can have a significant effect on groundwater flow and solute transport in low-permeability media. Determining the heterogeneity of the flow and transport parameters in such media is however a difficult task since measurements are usually scarce in low-permeability media, e.g., due to experimental difficulties. Secondary data such as geophysical and geological information can therefore be very useful to complement and compensate the often limited knowledge of the main flow and transport parameters. This study investigates the relative data value of different types of secondary information (grain size, gamma ray and resistivity) to estimate the flow and transport parameters in low-permeability media. The aim of this study is to determine which type of secondary data improves the estimation of the flow and transport parameters in low-permeability media the most. Future investigations of low-permeability media could use this knowledge to focus on measuring this "best" type of secondary information.

2. Data analysis

Data from four boreholes in two different clay formations are analyzed: the Mol-1 borehole, the Zoersel borehole and the Weelde borehole through the Boom Clay (Belgium) and the Doel borehole through the Ieper Clay (Belgium). The resulting data set consists of the following data:

- Hydraulic conductivity measurements: 52 in the Mol-1 borehole, 34 in the Zoersel borehole, 28 in the Weelde borehole and 25 in the Doel borehole
- Diffusion coefficient and diffusion accessible porosity measurements: 41 in the Mol-1 borehole and 25 in the Doel borehole
- Grain size measurements: 71 in the Mol-1 borehole, 152 in the Zoersel borehole and 160 in the Weelde borehole
- Gamma ray and resistivity logs in the four boreholes

3. Variography

Semivariograms and cross-variograms of all the primary and secondary variables are calculated using VARIOWIN (Pannatier, 1996). In order to determine the common model and range of the variograms required for a linear model of coregionalisation, all semivariograms and cross-variograms are initially modeled separately using the VARIOWIN "Model" module. The average range of the individually modeled variograms is used as the common range for all models in the linear model of coregionalisation. The nugget and sill values are fitted by the optimization program LCMFIT2 (Pardo-Iguzquiza and Dowd, 2002).

4. Cross-validation

Cross-validation is used to estimate the relevance of the different types of secondary data. The flow and transport parameters are first estimated by kriging using primary information only and then estimated by co-kriging using primary information and one type of secondary information (grain size, gamma ray and resistivity). The kriging residuals, i.e. the difference between the estimated and the measured values, of the different kriging and co-kriging variants are compared to evaluate the relevance of each type of secondary information.

4.1. Cross-validation of hydraulic conductivity

Tables 1 to 5 show the statistical properties of the co-kriging residuals. Every type of secondary information improves the estimates of K_v in each borehole (i.e. smaller standard deviation of residuals and larger correlation coefficient), except for resistivity in the Zoersel borehole. This is logical since resistivity and K_v are only weakly correlated (r = 0.43) in the Zoersel borehole. In the different boreholes, different types of secondary information improve the K_v estimates the most: gamma ray in the Mol-1 borehole, grain size d_{40} in the Zoersel borehole and resistivity in the Weelde and Doel boreholes.

Table 1. Statistical properties of the K_v kriging residuals of the Boom Clay in the Mol-1, Zoersel and Weelde boreholes

	Mol-1	Zoersel	Weelde
Arithmetic mean of residuals	-0.01	-0.02	-0.02
Standard deviation of residuals	0.32	0.42	0.51
Correlation coefficient of measured and estimated log K_v values	0.64	0.86	0.35

Table 2. Statistical properties of the Boom Clay K_v co-kriging residuals using K_v and grain size d_{40} information in the Mol-1, Zoersel and Weelde boreholes

	Mol-1	Zoersel	Weelde
Arithmetic mean of residuals	0.00	-0.04	-0.03
Standard deviation of residuals	0.29	0.30	0.46
Correlation coefficient of measured and estimated log K_v values	0.72	0.95	0.48

Table 3. Statistical properties of the Boom Clay K_v co-kriging residuals using K_v and gamma ray information in the Mol-1, Zoersel and Weelde boreholes

	Mol-1	Zoersel	Weelde
Arithmetic mean of residuals	-0.01	0.01	-0.03
Standard deviation of residuals	0.25	0.39	0.49
Correlation coefficient of measured and estimated log K_v values	0.81	0.87	0.39

Table 4. Statistical properties of the Boom Clay K_v co-kriging residuals using K_v and resistivity information in the Mol-1, Zoersel and Weelde boreholes

	Mol-1	Zoersel	Weelde
Arithmetic mean of residuals	0.01	0.01	-0.01
Standard deviation of residuals	0.26	0.46	0.44
Correlation coefficient of measured and estimated log K_v values	0.78	0.81	0.56

Table 5. Statistical properties of the K_v kriging residuals of the Ieper Clay in the Mol-1 in the Doel borehole using K_v and, respectively, no secondary data, grain size d_{40}, gamma ray and resistivity information

	no secondary data	d_{40}	GR	RES
Arithmetic mean of residuals	0.00	0.00	0.00	0.00
Standard deviation of residuals	0.38	0.38	0.37	0.36
Correlation coefficient of measured and estimated log K_v values	0.55	0.55	0.58	0.60

4.2. Cross-validation of diffusion coefficient

Tables 6 and 7 show the statistical properties of the co-kriging residuals. Every type of secondary information improves the estimates of the diffusion coefficient in each borehole (i.e. smaller standard deviation of residuals and larger correlation coefficient). In the different boreholes, different types of secondary information improve the D_e estimates the most: grain size in the Mol-1 borehole and resistivity in the Doel borehole.

Table 6. Statistical properties of the Boom Clay D_e kriging residuals using D_e and, respectively, no secondary data, grain size d_{40}, gamma ray and resistivity information in the Mol-1 borehole

	no secondary data	d_{40}	GR	RES
Arithmetic mean of residuals	-2.60E-12	-1.59E-12	-4.51E-12	-3.20E-12
Standard deviation of residuals	7.79E-11	7.13E-11	7.21E-11	7.32E-11
Correlation coefficient of measured and estimated D_e values	0.51	0.62	0.61	0.60

Table 7. Statistical properties of the Ieper Clay D_e kriging residuals using D_e and, respectively, no secondary data, grain size d_{40}, gamma ray and resistivity information in the Doel borehole

	no secondary data	d_{40}	GR	RES
Arithmetic mean of residuals	-8.98E-13	-7.98E-12	6.46E-13	-5.52E-12
Standard deviation of residuals	8.97E-11	6.85E-11	6.79E-11	6.08E-11
Correlation coefficient of measured and estimated D_e values	0.40	0.74	0.74	0.80

4.3. Cross-validation of diffusion accessible porosity

Tables 8 and 9 show the statistical properties of the co-kriging residuals. In the Mol-1 borehole, including secondary information does not result in a significantly better estimation of the diffusion accessible porosity, which is caused by the lack of correction between η and the secondary data. In the Doel borehole, every type of secondary information improves the estimates of η and including grain size results in the largest improvement.

Table 8. Statistical properties of the Boom Clay η kriging residuals using η and, respectively, no secondary data, grain size d_{40}, gamma ray and resistivity information in the Mol-1 borehole

	no secondary data	d_{40}	GR	RES
Arithmetic mean of residuals	0.00	0.00	0.00	0.00
Standard deviation of residuals	0.02	0.02	0.02	0.02
Correlation coefficient of measured and estimated η values	-0.09	0.24	0.02	-0.02

Table 9. Statistical properties of the Ieper Clay η kriging residuals using η and, respectively no secondary data, grain size d_{40}, gamma ray and resistivity information in the Doel borehole

	no secondary data	d_{40}	GR	RES
Arithmetic mean of residuals	0.00	0.00	0.00	0.00
Standard deviation of residuals	0.02	0.02	0.02	0.02
Correlation coefficient of measured and estimated η values	0.72	0.87	0.82	0.84

5. Discussion and conclusion

In the different clays and boreholes, different types of secondary information result in the largest improvement of the estimates of the flow and transport parameters. It is not possible to determine the "best" type of secondary information for improving the estimation of the flow and transport parameters of low-permeability media. This probably depends on local factors such as the quality of the loggings and the amount of nearby primary and secondary data.

Acknowledgments: The authors wish to acknowledge the Fund for Scientific Research – Flanders for providing a Research Assistant scholarship to the first author. We also wish to thank ONDRAF/NIRAS (Belgium agency for radioactive waste and enriched fissile materials) and SCK-CEN (Belgian Nuclear Research Centre) for providing the necessary data for this study.

REFERENCES

Pannatier, Y., 1996. *VARIOWIN: Software for Spatial Data Analysis in 2D*. Springer-Verlag, New York, NY.

Pardo-Igúzquiza, E. and Dowd, P.A., 2002. FACTOR2D: a computer program for factorial cokriging. *Comput. Geosci.*28(8): 857-875.