PREPROCESSING OF INITIAL DATA FOR CREATING HYDROGEOLOGICAL MODELS

A.Spalvins, J.Slangens, R.Janbickis

Riga Technical University, Latvia

Abstract
The quality of a hydrogeological model (HM) depends not only on the available initial data but also upon the interpolation of these data. It is important to decrease the influence of unreliable interpolation methods and possible human errors in a data set. The Geological Data Interpolation (GDI) programme has been developed by specialists of the Riga Technical University. GDI minimises the above-mentioned drawbacks of data preprocesing for HM. Basic principles of GDI are explained and some results reported.

1. Introduction

In a spatial (3D) steady state hydrogeological model (HM), the vector φ of piezometric levels is the solution of the following algebraic system of grid equations:

$$A \varphi = b \ , \quad A = A_{xy} + A_z - G \ , \quad b = - G \psi + \beta \qquad (1)$$

where the matrices A_{xy} , A_z , G represent, correspondingly, horizontal links (transmissivities) of aquifers, vertical connections (leakages) of semipervious strata (aquitards), elements connecting nodes of the HM grid with digitised boundary conditions ψ of the first kind; the vector β accounts for boundary conditions of the second kind (discharges/recharges of groundwater).
All the above-mentioned components of (1) are obtained by applying special interpolation methods. Even the solution φ of (1) may be interpreted as a 3D-interpolation.
There exist a wide scope of interpolation programmes which are available for obtaining various elements of the system (1). However, one ought to be careful, because some of them are not reliable enough. It is shown in this paper that such world-wide used methods as: Kriging (KR), Minimal Curvature (MC) and Inverse Distance (ID) may provide wrong results for elements of A_{xy} , A_z , φ .
In order to avoid drawbacks of the above-mentioned methods, the GDI programme has been developed by the team of the Environment Modelling Centre (EMC) of the Riga Technical University. The basic principles of GDI are explained in this paper and some of the results are reported.

2. Drawbacks of some interpolation methods

This section is based, predominantly, upon (SPALVINS A., SLANGENS J., 1994).
In the case of HM, the subdiagonal elements of A_{xy} , A_z are always equal to zero or nonnegative, because they are functions of thicknesses and permeabilities of geological strata. Never these two properties can take negative values.
Let us compare results of interpolating the thickness 2D-distribution in the case of a simple demonstration test by using four methods: KR, MC, ID and Boundary Problem (BP) grid solution. The BP method gives an approximation of the solution of the Laplace's equation by using the grid system of (1) for a homogenous environment.
The idea of the test is explained by Figure 1a which shows the BP solution. In a rectangular region two areas with zero and nonzero thicknesses σ are defined, accordingly. The nonzero distribution depends on $\sigma = 0$ and $\sigma = 100$, respectively, at fourth of the octagonal borderline and at the single point in the corner of the rectangle. The zero solution is defined by the nonflux condition $\partial\sigma / \partial n = 0$ on the borderline of the rectangle and $\sigma = 0$ on the mentioned octagonal borderline.
Similar zero areas on regional thickness maps are quite common (ATRUSKIEVICS J. et al. 1994) when hydrogeological layers are discontinuous. For the very reason, it is preferable that an interpolation method can obtain these zero zones trustworthy by using simple means of providing and accounting for initial data.
The following conclusions can be drawn regarding the BP solution surface of Figure 1a :

- the maximum/minimum principle is kept there automatically;
- the interpolation surface always carries a minimised amount of energy;
- if lines are used as information carriers, it is possible to account for initial data very easily and with a high reliability; in Figure 1a the conditions $\sigma = 0$ and $\partial\sigma / \partial n = 0$ are carried by the octagonal borderline and the cut along the edge of the rectangular, correspondingly.

In Figures 1b , 1c, 1d the KR, MC, ID solutions are shown, accordingly. They were provided by the SURFER programme (SURFER, 1997). No information was given about the rectangular border, because these methods could not account for any fluxes. On the other hand, stripping of initial data there helped very much to accentuate some hidden defects of the above methods.
No one of the three methods provided the right solution in the zero area. The KR and MC algorithms even produced negative values there, thus, ignoring the maximum/minimum principle. The ID result (Figure 1d) is much better, because it respects this fundamental condition. Nevertheless, $\sigma > 0$ in the zero zone for ID contradicts the minimal energy rule.
One can minimise the solution error in the zero zone by fixing some additional $\sigma = 0$ conditions there. However, as it is shown in (ATRUSKIEVICS J., 1994), the KR method manages to produce negative values even then.
One may succeed in correcting the wrong KR, MC and ID solutions in simple cases when the right answer is available. Practically no improvement of results and their quality check-up are possible in complex interpolation cases. Because of this reason, the EMC team has developed the GDI method which is based on the above mentioned BP solution techniques which have already been applied in (1) as well.

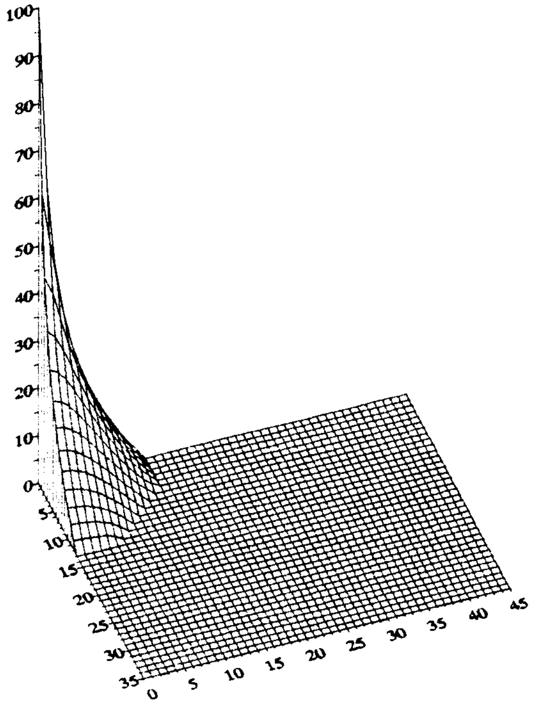

Figure 1a. BP method

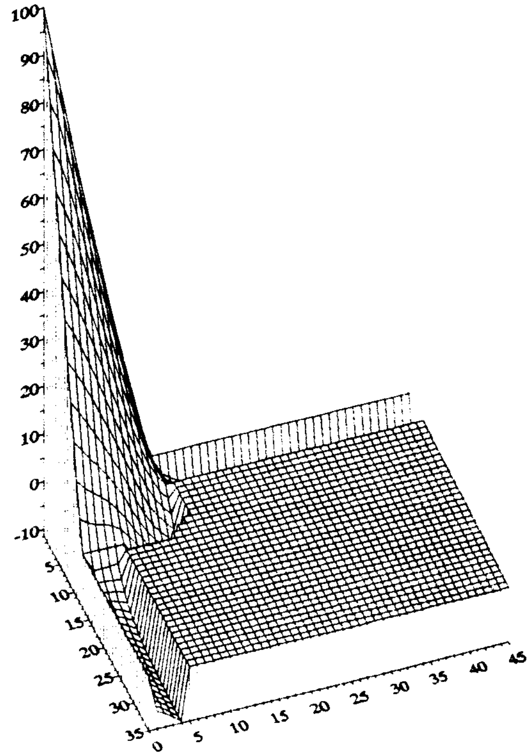

Figure 1b. KR method

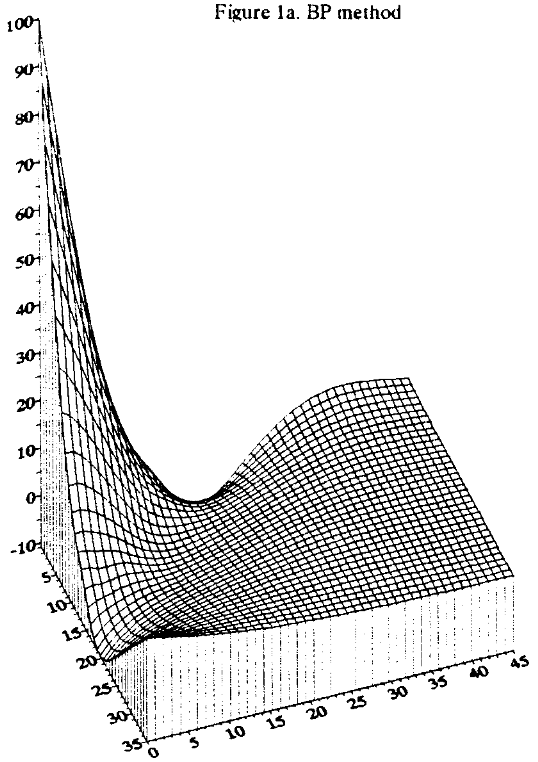

Figure 1c. MC method

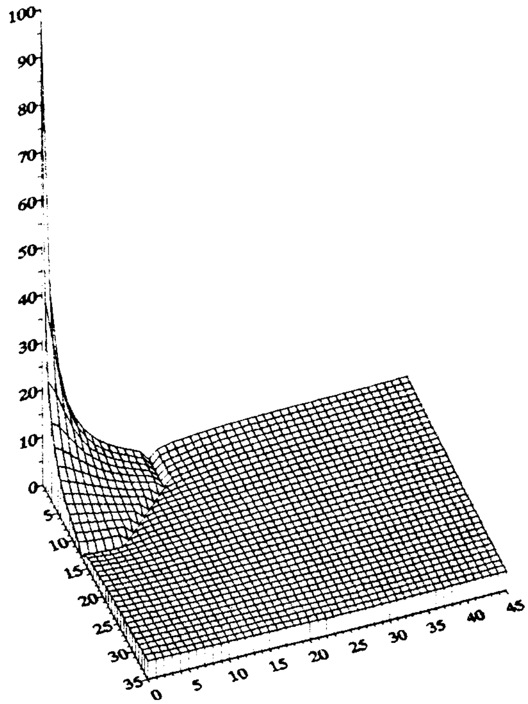

Figure 1d. ID method

Figure 1. A comparison of thickness distributions prepared by BP, KR, MC and ID methods.
The right solution for the zero thickness zone must be such as in Figure 1a

3. Basic principles of the GDI programme

In the GDI programme, any parameter σ to be interpolated (thickness, permeability, ψ -surface, etc.) is the solution of the following associated 2D-boundary problem which is in conformity with (1):

$$\partial (\upsilon \, \partial\sigma / \partial x) \partial x + \partial (\upsilon \, \partial\sigma / \partial y) \partial y = 0 \qquad (2)$$

where the boundary conditions of the first kind are provided by initial data σ_{in}. They may be applied as points and lines. Both are described in a grid-independent way. Two kinds of lines are used: carriers of σ_{in} ; cutlines for providing the nonflow $\partial\sigma / \partial n = 0$ condition.
The variable nonnegative conductivity parameter $\upsilon = \upsilon(xy) \geq 0$ provides a control upon the shape of the σ-surface of (2). If υ = const, then (2) turns into the Laplace's equation (as in the BP solution of Figure 1a where all the principal kinds of boundary conditions used in GDI are represented). As nonzero boundary conditions of the second kind are not used in (2), the maximum/minimum principle for σ holds in the case of any υ-distribution applied.
An approximation of (2) on the HM of (1) grid xy-plane gives the following algebraic system for finding σ in the plane chosen:

$$V\sigma = f, \qquad f = f(\sigma_{in}) \qquad (3)$$

where V is the interpolation matrix affected both by the parameter υ and the topology of σ_{in} and $\partial\sigma / \partial n$ carriers; f is the vector of fluxes caused by σ_{in}.
The idea of introducing in (2) variable υ was caused by an intuitive feeling that sharp-peaked solution surfaces at the vicinity of single σ_{in} points (as in Figure 1a for the BP solution) are not the best possible interpolation. It ought to be improved by transforming the acute tops into flat ones. For example, it is possible in the ID method (SURFER, 1997). For the BP solution the similar effect can be achieved by introducing high conductivity υ_n zones with respect to the data point. In GDI, it is accomplished with the aid of a factorial type function:

$$\upsilon_n = (n + 1)! \qquad (4)$$

where n is the number of transformed conductivity zones. The width of each zone is one grid step h . If n = 0 , then no transformation takes place (υ_0 = 1). When n > 1, then the data point gets encircled by n zones. For example, if n = 4, then values of conductivities for the transformed zones are : 120 ; 24 ; 6 ; 2 . The highest value is at the data point zone.
In Figures 2a , 2b the transformed BP solutions of Figure 1a are shown if n = 4 ; 9 at the point σ = 100 , correspondingly. The acute top σ = 100 has turned into an almost even summit. If $n \geq 4$, then its radius $\sim (n - 3)h$.
In GDI, it is possible to control, for point data, the υ parameter generally or individually. If line data are applied as a set of point data, then the width of the influence path of any line can be controlled too in a similar manner.
In Figure 3, as an example, the landscape elevation map is shown. It is built as the ψ-surface by applying various lines as information carriers: the long line profile of a river, water divides, etc. Single data points are scarcely used there as they are usually incorporated in data lines.
It is not our intention to discuss GDI here in detail, because they are explained in (SPALVINS A., SLANGENS J., 1994), (SPALVINS A., SLANGENS J., 1995). We will only classify and

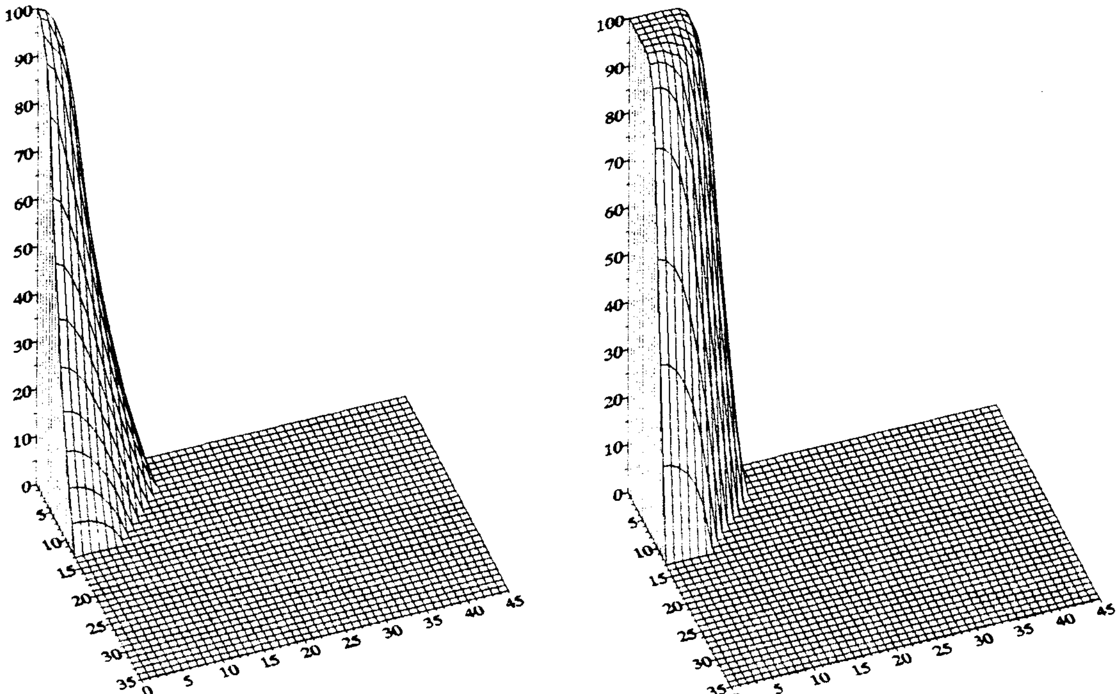

Figure 2a. n=4

Figure 2b. n=9

Figure 2. A comparison of the BP method nonzero thickness distribution surfaces the shape of which is controlled by encircling the point $\sigma = 100$ by n=4;9 transformed conductivity zones. The acute top at this point, shown in Figure 1a, has changed into an almost even summit

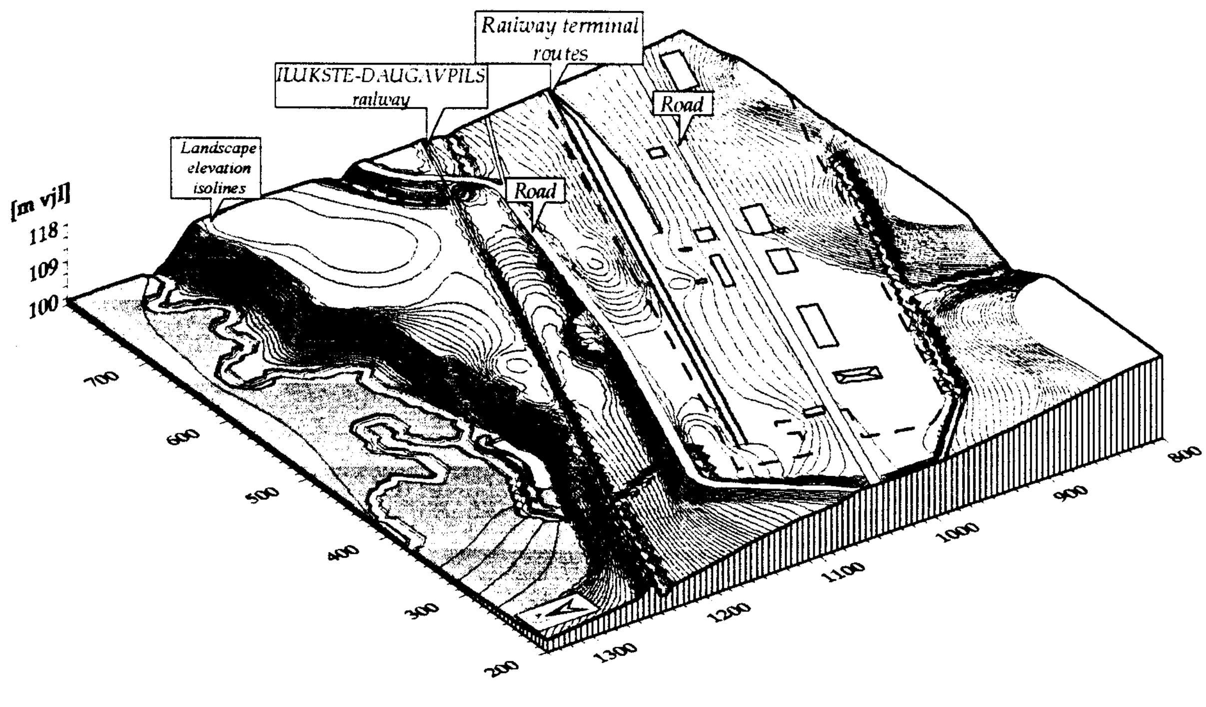

Figure 3. An example of the GDI generated Ψ-distribution in the form of a digital landscape elevation map. This map was obtained by applying mostly data lines as information carriers

comment on the general problems which were solved in order to implement the above explained basic ideas of GDI in software tools:

- check-up of initial data; conversion of initial data into grid-independent information carriers (points, lines);
- linking information carriers with the HM grid of (1);
- adjusting the υ-distribution of (2); obtaining the interpolation system (3);
- obtaining the solution of (3) in the form suited for the software of the chosen HM.

As the above GDI structure is practically identical with the one of HM, the following benefits are gained:

- identical algebraic system solvers and data storage arrangements can be applied in GDI and HM;
- because of a close synchronisation of GDI and HM data files it is possible to exclude most of human errors;
- it is relatively easy to build a 3D version of GDI. E. g., such is needed for creating 3D thickness distributions in complex multilayered HM.

4. Practical applications and conclusions

The GDI programme has been developed since 1993 by the EMC team in order to prepare elements of the REgional hydrogeological MOdel (REMO) for the central part of Latvia (SPALVINS A., et al, 1996). During this REMO project, GDI has been fully updated in 1996 for the first time and now it is being applied for the data preprocessing when various kinds of HM and digital maps must be created quickly and reliably.

The GDI programme has been successfully used in various practical applications. Some new ideas and suggestions have been produced in order to develop the next GDI version. It will be fully compatible with most of available software tools which are used with HM.

References

ATRUSKIEVICS J. et al, (1994) Computer Based Hydrogeological Model "Large Riga", Proc. of Latvian-Danish Seminar on "Groundwater and Geothermal Energy" (Boundary Field Problems and Computers, 35-th issue), Riga-Copenhagen, Vol. 2, 203 - 224.

SPALVINS A., SLANGENS J., (1994) Numerical Interpolation of Geological Environment Data., Proc. of Latvian-Danish Seminar on "Groundwater and Geothermal Energy" (Boundary Field Problems and Computers, 35-th issue), Riga-Copenhagen, Vol. 2, 181 - 196.

SPALVINS A., SLANGENS J., (1995) Updating of Geological Data Interpolation Programme, Proc. of International Seminar on "Environment Modelling" (Boundary Field Problems and Computers, 36-th issue), Riga-Copenhagen, Vol. 1, 175 - 192.

SPALVINS A. et al, (1996) Hydrogeological model "Large Riga". Atlas of Maps, Riga-Copenhagen, (Boundary Field Problems and Computers, 37-th issue), (bilingual: Latvian and English), 107.

SURFER (1997) Version 6.04, Users Manual, Golden Software, Inc.

9TH INTERNATIONAL SYMPOSIUM ON

SMC

SYSTEM MODELLING CONTROL

Edited by P. S. Szczepaniak

April 27 – May 1, 1998, Zakopane, Poland

English translation: Ewa Jasińska Graphic design: Krzysztof Tyczkowski
ISBN 83-909161-0-X