Creating reliable program for preparing line data of hydrogeological models

J. Slangens & A. Spalvins
Environment Modelling Centre, Riga Technical University, Latvia

Keywords: hydrogeological models, line data interpolation

ABSTRACT: Line data are the major part of information needed to create hydrogeological model (HM) by the team of the Environment Modelling Centre (EMC), the Riga Technical University. The CPR program prepares these data, as the input for interpolation. This paper describes the updated CRP version of improved reliability.

1 INTRODUCTION

Various types of data lines (isolines, geological borders and sections, long lines profiles of rivers, etc) can be applied for creating HM. The data line L_σ has its carrier L and the data profile σ. The line L is the broken one passing through the master points $j = 1, 2, \ldots, J$. The set of $J - 1$ directed straight line segments represents the following vectored form of L:

$$L = \{ c_j \},\;\; j = 1, 2, \ldots, J;\;\; l_{1,J} = \sum_{j=1}^{J-1} l_{j,j+1},\;\; l_{j,j+1} = \sqrt{(x_j - x_{j+1})^2 + (y_j - y_{j+1})^2} \quad (1)$$

where $\{ c_j \} = \{ x_j, y_j \}$ is the coordinate set, on the continuos xy-plane, of the master points serving, as the dots where L turns; $l_{1,J}$ and $l_{j,j+1}$ are the lengths, accordingly, of L and the elementary vector linking the adjacent points j and $j + 1$.

The physical nature of the σ function may be different (elevation z of some surface, thickness h_z, or permeability k of a geological stratum, etc). The simplest graph of σ is a horizontale used for detailing isolines. Generally, the shape of σ may be complex, for example, such as elevation profiles of an eroded ground surface. If the additional ponts $k = 1, 2, \ldots, K$ are introduced, to aproximate some curvilinear graph σ then the data line L_σ includes $N = J + K$ points (located on L), and L_σ is specified, as follows:

$$L_\sigma = \{ c_i, \sigma_i \},\;\; i = 1, 2, \ldots, N,\;\; \{ c_i = c_j \wedge c_k \},\;\; \{ \sigma_i = \sigma_j \wedge \sigma_k \}. \quad (2)$$

The duo sets of the xy-coordinates and the σ-values $\{ c_i, \sigma_i \}$,$\{ c_j, \sigma_j \}$, and $\{ c_k, \sigma_k \}$ represent, accordingly, the current i-th data points, the master wells j of L, and the additional points k dividing the elementary segments $l_{j,j+1}$ of L into collinear pieces.

It was necessary to develop the special CRP program (Atruskievics et al., 1995), because of the following main reasons:

- in practice, rarely any line L_σ is available, in the arranged form of (2); some data processing is needful to obtain this form;
- special calculations must be performed, to extract from L_σ (2) the duo set $\{ c_s, \sigma_s \}$ where c_s and σ_s are the xy-coordinates and the σ-values, accordingly, at the intersections of L_σ with the HM grid; this set represents the "c-data" applied by the GDI program for incorporating the lines into HM (Spalvins and Slangens, 1994); the set $\{ c_s, \sigma_s \}$ carries the line L_s, which approximates L_σ; closeness of these lines depends on the plane step h used for the grid (fine grid → close L_s and L_σ).

Let us consider, for example, how the CPR program creates data lines for the road shown in Fig.1. The electronic image b is scanned from the hard copy a. Digitizing of the road provides table c (*road.dig*:) where some part of elevations $\sigma_i = z_i$ is unknown (marked by ?). By applying the linear interpolation along the road line $\{ c_i \}$, $i = 1, 2, \ldots, 17$, the CPR program creates table f (*road.int*:) which represents the arranged form (2) of the L_σ. From table *f*, CPR obtains tables *d* (*road.cpr*: - the set { c_s, $\sigma_s = z_s$ } of L_s applied by GDI) and *e* (*road.prf*: - applicable for visualization of the *z*-profile). The third column of table *d* is used for service marks (1; 0 and 9) of GDI (1 - { σ_s } are applied at { c_s }, as numerical values; 0 - { c_s } are cutpoints of the HM grid, no { σ_s } values are used; 9 - { c_s, σ_s } represent sources for generating some masked GDI areas). More information about the regimes of GDI is given in (Spalvins and Slangens, 1995). The pictures *g* and *h* of Fig. 1 demonstrate good closeness of L_σ and L_s lines.

Ample practice of applying the CPR program discovered the following drawbacks of this tool:

- wrong values c_s, σ_s may occur, especially, when the segment $l_{j,\, j+1}$ is directed horizontally or vertically;
- the coordinate c_s often coincide with a node of the HM grid; this should not be so, because such an occurrence is hardly probable;
- the formal line L_s based on the set { c_s, σ_s } gets occasionally interrupted; this failure may worsen, even ruin the GDI results; usually, the reason of such an interruption is some error, in the sequence of the points *i* of (2).

To subdue the above faults, the CPR program has been updated.

2 UPDATED CPR VERSION

The above mentioned faults of the old CRP version were mainly due to mishaps of the search methods for c_c of crosspoints. The methods were based on the following conventional equation for $l_{j,j+1}$:

$$y - y_j = d_j\,(x - x_j), \quad d_j = (y_{j+1} - y_j) / (x_{j+1} - x_j) = \tan \delta_j. \tag{3}$$

The gradient d_j has infinite limits ($\infty > d_j > -\infty$) charesteristic for the trigonometric tangent (*tan*) function; δ_j is the angle between $l_{j,j+1}$ and the axis *x*. Computations based on (3) fail if $\delta_j = \pi / 2$; $-\pi / 2$, or 0 (vertical or horizontal $l_{j,j+1}$). We guess that this feature of (3) partially caused the occurrence of wrong CPR results.

In the updated CRP program, other classic equation for $l_{j,j+1}$ is applied:

$$(x - x_j) / (x_{j+1} - x_j) = (y - y_j) / (y_{j+1} - y_j) = D_{jx}\,, \quad 1 \geq D_{jx} \geq 0 \tag{4}$$

where the ratio D_{jx} is always positive and finite. It is used for computations, as follows:

$$x = x_j + D_{jx}\,A_j\,, \quad y = y_j + D_{jx}\,B_j\,, \quad A_j = x_{j+1} - x_j\,, \quad B_j = y_{j+1} - y_j\,,$$
$$x = x_j \;\text{ if }\; A_j = 0\,, \qquad y = y_j \;\text{ if }\; B_j = 0. \tag{5}$$

In the new CPR version, the computations of (5) are performed with the double precision, and no mishaps caused by the form (3) occur.

The most difficult task for CPR is to find the coordinates c_c of crosspoints and the new algorithm for this search has been applied. On the elementary segment $l_{j,j+1}$, it contains the following main steps:

- to find the current *h*-block where the *j*-th point is located:

$$x_j = x_0 + (i_x - 1)\,h + \eta_j\,, \quad h > \eta_j \geq 0\,, \quad i_x = 1, 2, \ldots, n_x\,,$$
$$y_j = y_0 + (i_y - 1)\,h + \lambda_j\,, \quad h > \lambda_j \geq 0\,, \quad i_y = 1, 2, \ldots, n_y \tag{6}$$

where x_0, y_0 are the initial coordinates for the uniform plane grid of HM; n_x, i_x and n_y, i_y are total and current numbers of the grid lines for the *x* and *y* directions, accordingly; η_j, λ_j are the local coordinates of the *j*-th point with respect to the *h*-block; the further search for the nearest crosspoint η_c, λ_c is performed, in this local η, λ - coordinate system;

- for the *j*-th point, to detect which of the four possible variants of its location, is dialed with:

$$1.\ \eta_j > 0\,, \lambda_j > 0\,; \qquad 2.\ \eta_j > 0\,, \lambda_j = 0\,;$$
$$3.\ \eta_j = 0\,, \lambda_j > 0\,; \qquad 4.\ \eta_j = 0\,, \lambda_j = 0\,; \tag{7}$$

- for the *j+1*-th point, by checking its position to find out for each of the variants, if any crossspoint is located on $l_{j,j+1}$:

$$1.\ 0 < \eta_{j+1} < h,\quad 0 < \lambda_{j+1} < h;\qquad 2.\ 0 < \eta_{j+1} < h,\quad -h < \lambda_{j+1} < h;$$
$$3.\ -h < \eta_{j+1} < h,\quad 0 < \lambda_{j+1} < h;\qquad 4.\ -h < \eta_{j+1} < h,\quad -h < \lambda_{j+1} < h. \tag{8}$$

The search for crosspoints must be switched to the next segment $l_{j+1,\ i+2}$ if the condition (8) is satisfied; if not then the computations of the coordinates η_c, λ_c must be started.

These computations are organized, as a trial and error type process containing the following two stages:

- in the formula (5), apply the values:

$$\eta_c = h \text{ if } A_j > 0,\ \eta_c = -h \text{ if } A_j < 0;$$
$$\lambda_c = h \text{ if } B_j > 0,\ \lambda_c = -h \text{ if } B_j < 0, \tag{9}$$

as the try for computing of the coordinates λ_c and η_c, accordingly;

- for the nearest crosspoint, the right computed values λ_c and η_c must satisfy the following conditions:

$$h \ge \lambda_c \ge 0 \text{ if } \lambda_j > 0;\qquad h \ge \lambda_c \ge -h \text{ if } \lambda_j = 0;$$
$$h \ge \eta_c \ge 0 \text{ if } \eta_j > 0;\qquad h \ge \eta_c \ge -h \text{ if } \eta_c = 0. \tag{10}$$

The condition (10) sorts out the false crosspoints external to the h-block considered. The new, valid crosspoint (η_c, λ_c) is declared, as the formal start for the remaining part of the line segment analysed, and the algorithm proceeds, until all $J-1$ segments of (1) are processed. All necessary co-ordinate transformations "global ↔ local" are based on the representation (6), and their performance are not detailed in this paper.

The new algorithm for the crosspoints produces correct values c_c for any lines tried. When the full set $\{ c_c \}$; $c = 1, 2, \ldots, C$ is obtained then the corresponding $\{ \sigma_c \}$ values are computed, as follows:

- the elementary segment $l_{i,i+1}$ of (2) must be found where the current crosspoint c_c is located;
- the current value σ_c is computed:

$$\sigma_c = \sigma_i + D_{ic}\,(\sigma_{i+1} - \sigma_i),\qquad D_{ic} = l_{ic} / l_{i,\ i+1},\qquad c = 1, 2, \ldots, C \tag{11}$$

where σ_i , σ_{i+1} are the values of σ at the adjacent points i and $i + 1$, respectively; l_{ic}, $l_{i,\ i+1}$, are the distances, correspondingly, between the points i and c, i and $i + 1$. The formula (11) reflects the rule of linear interpolation.

In the updated CPR version, two simple rules are introduced, to detect possible errors, in the sequence of the master j-points of L:

$$1.\ l_{j,j+1} \le \varepsilon_l;\qquad 2.\ (l_{j,j+1} + l_{j+1,j+2}) / l_{j,j+2} \ge q_l \tag{12}$$

where ε_l, q_l are empirical variable constants applied, correspondingly, for sorting out very closely located adjacent points (rule 1) and for detecting the questionable situation, when the $j+1$-th point is located far from the line $l_{j,j+2}$ (rule 2).

The updated version of CPR was specially developed by the EMC team, to prepare complex HM of the Noginsk District, Russia (Spalvins et al., 2000). This version performed much better than the old one of CPR.

3 CONCLUSIONS

To prepare line data for interpolation purposes, these data must be preprocessed by the special CPR program.

Because the old CPR version possessed some minor drawbacks bothering creating of complex HM, the new program was developed by the EMC team.

The new CRP version provides reliable results. It can also detect some errors of initial data.

REFERENCES

Atruskievics, J., Viksne, Z., Spalvins, A. & Lace, I. 1995. Lakes and rivers algorithms in the hydrogeological model "Large Riga". *Proc. of International Seminar on "Environment Modelling",* (1), Riga – Copenhagen: 193-200. (*Boundary Field Problems and Computers, 36th issue*).

Slangens, J. & Spalvins, A. 1995. Updating of geological data interpolation programme. *Proc. of International Seminar on "Environment Modelling"* (1). Riga - Copenhagen: 175 - 192. (*Boundary Field Problems and Computers, 36th issue*).

Spalvins, A. & Slangens, J. 1994. Numerical interpolation of geological environment data. *Proc. of Latvian - Danish Seminar on "Groundwater and Geothermal Energy",* (2), Riga – Copenhagen: 181-196. (*Boundary Field Problems and Computers, 35th issue*).

Report. 1996. Modelling of the Groundwater Flow Dynamics and the Contaminant Movement for the Rumbula Airbase Place by Spill CAD, ARMOS and BioTRANS. Contract No 6153/96 (BC-001). Riga: Riga Technical University, Environment Modelling Centre.

Spalvins, A., Slangens, J., Janbickis, R. Lace, I., Gosk, E. & L. Loukiantchikova L. 2000. New simulation technologies applied for creating hydrogeological model of Noginsk district, Russia (*in this book*).

Janis Slangens, Dr.sc.ing.
Aivars Spalvins, Dr.sc.ing.
Riga Technical University, Environment Modelling Centre
Address: 1/4 Meza Str., Riga, LV-1048, Latvia
Phone: +371 7089518; E-mail: emc@egle.cs.rtu.lv

Šlangens J., Spalviņš A. Drošas programmas izveidošana līniju datu gatavošanai hidroģeoloģiskajos modeļos.
Līniju dati ir galvenā daļa informācijai, kuru hidroģeoloģisko modeļu (HM) radīšanai izmanto Rīgas Tehniskās Univesitātes Vides Modelēšanas centrs. CRP programma sagatavo šos datus interpolācijas izpildīšanai. Šis raksts apskata modernizētu CRP versiju, kurai piemīt uzlabota drošība.

Шланген Я., Спалвинь А. Создание надежной программы для подготовки данных, представленных линиями, для гидрогеологических моделей.
Данные, представленные линиями составляют главную часть информации используемой для создания гидрогеологических моделей (ГМ) Центром моделирования окружающей среды Рижского Технического университета. Программа CRP подгатавливает эти данные для последующей интерполяции. В этой статье рассматривается усовершенственная версия CRP обладающая повышенной надежностью.