Methods for improving verity of groundwater modelling

A. Spalvins, J. Slangens, R. Janbickis, I. Lace * E.Gosk†

Abstract

Verity of hydrogeological models (HM) is often limited by a low quality of initial data and by inapt software tools applied. The team of the Environment Modelling Centre (EMC) of the Riga Technical University has developed methods and tools helping to deal with problems occurring in the course of groundwater modelling. Weak points of some wide-known software tools are discussed.

Key words: : hydrogeological model, interpolation of data.

AMS subject classifications: 35Q35.

1 Introduction

To narrow the field to be discussed, only typical errors for semi-3D steady state HM will be reviewed. To consider them, some mathematics is needful. The xyz-grid of HM is built of $(h \star h \star h_z)$-sized blocks (h is the uniform block plane size; h_z is a variable block height). They constitute a rectangular $(2s+1)$-tiered xy-layer system where $s+1$ and s are, accordingly, the number of aquifers and of interjacent aquitards separating these aquifers. Its four vertical sides compose the shell of the HM grid. The relief (ground surface) and the lower side of the model are its geometrical top and bottom, respectively.

The vector φ of the piezometric head is the numerical solution of a boundary field problem, which is approximated, in nodes of the HM grid, by the following sparse algebraic equation system:

$$A\,\varphi = b, \quad A = A_{xy} + A_z - G, \quad b = \beta - G\,\psi, \quad \beta = \beta_{in} + \beta_{bot} + \beta_{sh} + \beta_w, \quad \beta_\psi = G\,(\varphi - \psi) \tag{1}$$

where the matrices $A_{xy},\ A_z,\ G$ represent, correspondingly, horizontal links a_{xy} of aquifers (arranged in xy-planes), vertical ties a_z originated by aquitards, elements connecting nodes of the grid with the piezometric boundary conditions ψ specified on some HM areas (ψ-surfaces); the vector β accounts for boundary flows: β_w is the water production rate in wells; β_{in} , β_{bot} and β_{sh} are the boundary flows, which may be specified on some areas (β-surfaces) of top, bottom, and shell surfaces, accordingly; β_ψ is the computed flow passing through the elements g_{xy}, g_z forming the diagonal matrix G.

Each grid plane contains $n_{xy} = n_x \star n_y$ nodes (n_x and n_y represent, accordingly, the number of nodes on its x and y sides). Thus, the whole grid of (1) contains $n = n_{xy} \star (s+1)$ nodes, only the neighbouring nodes are interconnected. For this reason, the matrix A is sparse (it contains no more than $7n \ll n^2$ nonzero elements). The elements a_{xy}, a_z of A_{xy}, A_z (or g_{xy}, g_z of G) are computed, as follows:

$$a_{xy} = km, \quad a_z = h^2 k/m, \quad k \geq 0, \quad m_i = H_{i-1} - H_i \geq 0, \quad i = 1, 2, \ldots, 2s+1 \tag{2}$$

where H_{i-1} and H_i are the elevation distributions of the top and bottom surfaces of the i-th geological layer; H_0 represents the ground surface map with the hydrographical network included; m, k are, accordingly, elements of the

*Environment Modelling Centre, Riga Technical University, 1 Meza Street, Riga, LV-1048, Latvia, e-mail: emc@egle.cs.rtu.lv

†Geological Survey of Greenland and Denmark Thoravej 8, DK-2400 Copenhagen NV, Denmark, e-mail: eg@geus.dk

digital m, k-maps of computed thickness and permeability distributions of layers. The set of the H_i maps describes the full geometry of HM. It is built in the consecutive way: $H_0 \rightarrow H_1 \rightarrow \ldots \rightarrow H_{2s+1}$, by keeping the thickness of the i-th layer $m_i \geq 0$. If in some area $m_i = 0$ then the i-th layer is discontinuous.
The course of creating and simulation of the system (1) can be conditionally decomposed into the following three sequential phases, during which some specific errors may occur:

- the pre-processing (model formulation) providing both initial constituents: A and b of (1); unfortunately, initial hydrogeological data of practical problems are never complete enough, and calibration processes should be used, which enable to mend insufficient data by employing some veritable calibration targets;
- obtaining of the distribution φ by a special solver programme;
- post-processing of data carried by HM.

Faults of the model formulation are the dominant ones [1], and they are also interdependent with the errors of solving and post-processing. The above simulation cycle should be repeated many times, because the calibration of HM is iterative. The team of EMC has developed some rather effective tools to fight most of the known groundwater simulation faults.

2 Model formulation errors

During the model formulation phase, digital maps of the parametres: H, k, ψ, and β should be prepared, as inputs for modern modelling service systems: Groundwater Vistas (GV) [2], Visual MODFLOW (VM) [3], etc. They provide a common environment for the MODFLOW [4] and MT3D [5] type codes used, correspondingly, as HM and a mass transport model. The GV and VM systems do not use the thickness m-maps, as inputs.
The main sources of the formulation errors are, as follows: mistakes of initial data, wrongly applied or inapt data interpolation methods, use of unfit boundary flows, human errors, incongruity of software used. Indications of erroneous pointwise initial data (technical faults of measurements, careless attitude of personnel, mismatched observations of seasonally strongly dependent head values in wells, etc.) - have been described in [1], and we are not going to discuss them here.
The safety and amount of initial information increase considerably if data lines are applied, as verified generalization of pointwise data. For example, the border of a layer is the isoline of its zero thickness: $m = 0$. A line can also carry any variable function: the long line profile of a river is obtained by interpolating along the river between the points of observed water levels, and this profile is much more informative then the initial point data used separatively; a geological vertical cross section includes a matched set of elevation lines of layer interfaces providing veritable geometrical data.
None of available ordinary interpolation methods is perfect to make input maps for complex HM. The wide-known methods: Kriging, Minimal Curvature and Inverse Distance, which are also included in the graphical system SURFER [6], may provide results unfit for HM [7], because these methods cannot control strictly the interpolated surface via initial data. Two basic reasons cause this limitation: the surface does not hold the maximum/minimum principle towards initial data applied, and it also does not carry a minimum of a potential energy.
To guarantee controllable results, the Geological Data Interpolation (GDI) programme has been developed [7]. It creates an interpolation, as the numerical solution of an arbitrary boundary field problem, which is formulated, on the HM grid for the parameter, to be interpolated. The initial data of this parameter are applied, as boundary conditions (points, lines). The GDI programme possesses the following advantages: any interpolation surface can be exactly ruled, because it always satisfies both above mentioned principles; to create surfaces of utmost complexity, the sequential recurrent step mode can be applied - during the current interpolation step, new information is added, but the results of the previous step serve, as the base, which should be modified [7].
The GDI programme creates the necessary exact surfaces for H_{i-1} and H_i-maps of the discontinuous geological layer interfaces, which should hold the condition - $m_i = 0$, within the areas where the i-th layer does not exist.
The computer GDI-generated permeability k-maps serve, as the main instruments, to calibrate HM. Their impact is so considerable that they enable to account for geometrical changes, which may be caused by dewatering of aquifers. For this reason, there is no need to vary geometry of HM, until the calibration is completed. Moreover, any attempt to act otherwise may lead to a collapse of HM (see the next section).
The quality of H, k and β maps depends on the difference between the initial data and their interpolated values. It is often overlooked that the formal mass balance of an interpolation surface is always disturbed, at locations where initial

data are applied, as the input. For this reason, the ψ-maps cannot be created freely, because the virtue of these maps depends on two residues simultaneously (value and mass balance). For example, a groundwater table map cannot be generated properly by SURFER if only pointwise initial data are applied, because unwonted flows will appear, at these points where observed head values are fixed, as the veritable ones.
The fixed flows β_{in} , β_{bot} and β_{sh} should not be used, as the boundary conditions, because of the following undesirable effects:

- practically no veritable experimental data are available about distributions of these flows;
- the β-surfaces distort greatly shapes of depression cones created by the flow β_w;
- the stability of the solver programme decreases considerably if the β-surfaces are present;
- an increased sensitivity of HM towards possible errors.

Fortunately, these fixed flows can be always substituted for computed flows β_ψ , which involve the condition ψ [1]. The EMC team has developed the following effective method for substituting the most influential infiltration flow β_{in}. The ground surface map H_0 is applied, as a fixed condition ψ_{rel}; the unsaturated aeration zone belonging to the underlying, conditionally confined q aquifer is presented by a formal *aer* aquitard that accounts for the variable leakance k_{aer}/m_{aer}. Its value alters from the small one of an aquitard (a dewatered soil of recharge areas) to a multifold larger ratio of an aquifer (in water-saturated discharge areas, beds of rivers and lakes included). The infiltration rate is automodelled, as $\beta_\psi = G$ (ψ_{rel} - φ_q), in the form of a vertical flow passing through the *aer* aquitard, where φ_q is the computed head of the q aquifer. To simplify the infiltration control (only k_{aer} variable!) and to avoid any real influence of the formal thickness m_{aer} on the HM geometry, m_{aer} is chosen small and constant, everywhere ($m_{aer} =$ 2 cm may be used in most cases). When the computed infiltration distribution is provided by HM, as the concorded part of (1), the value ψ_{rel} - φ_q (real thickness of the aeration zone) should be used, as a calibration target. Due to the above innovatory approach, good simulation results are guaranteed for the q aquifer and whole HM.
The following annoying errors may occur, regarding modelling of the production rate β_w [8]:

- often co-ordinates and rates of imaginary "large" production wells are given, instead of the individual ones; such an approach should be definitely avoided, because it may result in crude simulation errors, especially, when the HM grid must be refined, in order to create detailed local models:
- when a production well extracts water from several aquifers simultaneously, the real production rates for each aquifer are uncertain, even if the total rate of the well is exactly known.

The REMO code has been developed and used by the EMC team for creating and calibrating of HM, because of the following advantages [8]:

- the impact of human errors is considerably reduced, because all maps of HM are obtained digitally from initial data by the GDI programme, which is synchronised with REMO;
- the ground surface map is always used for controlling the infiltration flow, on the HM top;
- no special software is needed to account for the hydrographical network, because flows caused by it are components of the computed infiltration flow;
- data from irregularly located production and observation wells can be interpolated, to nodes of an HM grid, or backwards, if necessary [9];
- the shell of HM acts, as an interpolation device for creating ψ_{sh}-maps;
- vertical cross section pictures of high quality can be created, along any line chosen;
- REMO is handy for a hydrogeologist, to assist the calibration process of HM.

The above mentioned ability of REMO to interpolate elements of the flow β_w, considerably reduces simulation errors, especially, if $h = 0.5$ - 4.0 km (regional HM!). The GV and VM codes roughly move a production rate of an irregularly located well, to the nearest node (REMO interpolates each rate among four nearest nodes!).
In view on pre-processing, the EMC team has estimated the GV and VM systems. The following results have been obtained [1]: the GV and VM systems are under notable influence of principles, which have been applied, in the era of slow mainframe computers (a zone principle used to save memory and to prepare input maps manually, evident

expectations that HM should be simple, needlessly strong human involvement). These outdated principles not only stimulate formulation errors, but also cause some incongruity with other related novel software tools (REMO, GDI, SURFER, etc.), which should be used for preparing the input maps of HM. For example, due to the zone principle, the beforehand-prepared maps may be seriously distorted if imported, into the GV and VM environments. Moreover, the VM system accepts maps, which are prepared in MODFLOW codes only! Hence the VM system applies only the pure 3D approximation scheme, the number n of nodes increases roughly twofold (compared with GV) - each aquitard of HM also contains an unnecessary (for most cases) set of n_{xy} nodes.
It follows from the above materials that fighting against model formulation errors is a very hard task. For this reason, any complex HM should be prepared by a team including specialists from overlapping fields (data interpolation, hydrogeology, computer software, etc.). A lot of related software tools should be applied and developed, to create HM. No single system is able to cope with such a problem.

3 Faults of solver programmes

A modeller should be aware that mistakes may occur, during the phase of solving the system (1). There are two basic sources of errors: a solver programme is not able to produce the distribution φ of required accuracy; the system (1) gets ruined by the solver programme.
The solver programme may fail, because of the following reasons: an obsolete numerical solution method used (one should avoid to use old-fashioned methods: SOR, SIP, etc., which are unwisely offered by GV, VM and other systems); the number of grid nodes n is too large; a harmful impact of the β-surfaces; the system (1) presents more then one practically independent numerical problem to be solved (such a situation may arise when discontinuous geological layers are involved).
The system (1) may be ruined if the solver programme is allowed to change the geometry of HM by accounting for the influence of unconfined aquifers. For example, the MODFLOW programme annihilates dewatered grid cells of such aquifers, during the iterative solution process [4]. This move destroys the vertical continuance of HM. For discontinuous aquifers such an algorithm inevitably leads to a collapse of HM. For this reason, a modeller should avoid to use any regimes allowing presence of unconfined aquifers. It is much safer to account indirectly for possible changes of the HM geometry by applying the permeability maps. For example, the real thickness of the unsaturated aeration zone can be always applied after the calibrated leakance distribution of the formal *aer* aquitard is found.

4 Mistakes of data post-processing

Not only the distribution φ, but also all other main digital maps forming HM (H, m, k, β, ψ), represent results of modelling. By applying them, complex synthetic maps may be obtained (vertical cross sections with various lines shown, isometric diagrams, computed vertical flows through aquifers, vulnerability maps, etc.). All these maps should be available, as printed documents, with isolines of involved parametres shown. Various maps may be used, as backgrounds (situation or basemaps are often applied here) of the documents. Although, the GV, VM, and SURFER systems are able to produce isolines, the task is not so simple, as it may appear, because serious mistakes may spring up here:

- a location of an isoline may be computed wrongly;
- a form of the isoline may get misshaped by tools of line smoothing applied.

To draw an isoline, one should find its geometrical place by applying gridded data of the parameter to be visualized and information regarding A and b of (1). For example, the back-interpolation $\varphi \rightarrow \varphi_{xy}$ module of REMO [9] can find the value φ_{xy}, at any irregularly located observation well. This value is applied, during the calibration of HM (GV and VM systems do not have such an improvement). The module accounts for the following items: the solution φ; the location of the observation well; the impact of A_{xy} (space heterogeneity); the possible influence of the local depression cones, caused by irregularly located production wells (components of β_w). Although, this tool is not used for drawing isolines, it discovers the weak points of all ordinary isoline-creating programmes - none of them considers for the influence of A_{xy} and β_w! When isolines are drawn, additional assumptions about their smoothness can be introduced. If the map, to be visualised, represents a complex, not smooth surface (for example, the H_0-map) then the SURFER system may distort the isoline picture awkwardly, unless any line smoothing is blocked: not smoothed straight lines may turn into wavelike "smoothed" ones, unpredictable ups and downs on the smoothed surface may

appear. For this reason, one should always compare the pictures of smoothed and not smoothed isolines, in order to avoid serious mistakes of data visualisation.
Unfortunately, the GV and VM systems produce misleading φ-pictures of the vertical cross sections [10]: no vertical orientation of isolines, within aquifers; the isolines have no breaking points (as it should be), on interfaces of aquifers and aquitards; not all interfaces of geological layers are shown. These pictures can be drawn along rows and columns of the HM grid only (in REMO, the cross-section of a high quality can be drawn along any line chosen!). This serious fault of post-processing is caused by the superficial assumption that the φ-pictures of the vertical cross sections and of the xy-planes behave similarly, and the same interpolation method can be used in both cases. It is not true, because values of vertical hydraulic gradients existing, on aquifers and aquitards, differ drastically! It follows from the above material that interpolation programmes used for data post-processing may provide incorrect graphical pictures. The consequences may be even more serious if inapt principles of these programmes are applied to control a movement of tracer particles for the mass transport models [5].

5 Conclusions

Model formulation, solving, and pre-processing errors can be subdued if special methodology and software are focused on this purpose. The following main measures are of importance:

- creating and testing of initial data bases, which include information carried by points and lines;
- development and use of computer based reliable interpolation methods, to prepare necessary digital maps;
- withholding involvement of people in routine processes, in order to minimize the impact of human errors;
- substitution of uncertain fixed boundary flows for more reliable computed ones;
- preventing use of inappropriate algebraic system solvers and avoiding such regimes of solvers, which may result in a collapse of a model;
- cautious use of rather precarious software used for post-processing purposes;
- orientation on a teamwork involving specialists from overlapping fields (data interpolation, hydrogeology, software, etc.), in order to create complex groundwater models;
- applying different, highly effective specialised software tools for creating of a complex model, because no single system can cope with such a task.

We hope that results reported, in this paper, will be of some interest for modellers involved, in creating of complex hydrogeological models.

References

[1] A. Spalvins, J. Slangens, R. Janbickis, I. Lace, *Reducing of Model Formulation Errors as an Effective Remedy for Improving Simulation Results*, in Proceedings of International Conference on "Calibration and Reliability in Groundwater Modelling, ModelCARE'99". - September 20-23, 1999, Zurich, Switzerland, 1999. - pp. 161 - 166.

[2] *Groundwater Vistas. Guide to Using*, Environmental Simulations, Inc., 1997.

[3] *Visual MODFLOW. User's Manual*, Waterloo Hydrogeologic, 1998.

[4] M. McDonald and A. Harbaugh, *A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model*, U.S. Geological Survey. Open File Report 83-875. Washington, 1988.

[5] *MT3D'96. Documentation and Input Instructions*, S. Papadopulos and Associates, Inc., 1996.

[6] *SURFER 6.0 for Windows. User's Manual*, Golden Software. Inc., 1997.

[7] A. Spalvins and J. Slangens, *Numerical interpolation of geological environment data*, in Proceedings of Latvian - Danish Seminar on "Groundwater and Geothermal Energy", Riga - Copenhagen, Vol. 2, 1994, pp. 181 - 196, (Boundary Field Problems and Computers; 35-th issue).

[8] A. Spalvins , R. Janbickis, J. Slangens, E. Gosk , I. Lace , J. Atruskievics, Z. Viksne, N. Levina , I. Tolstovs, *Hydrogeological Model "Large Riga". Atlas of Maps*, Riga - Copenhagen, 1996, 101 p. (Boundary Field Problems and Computers; 37-th issue; biling.; Latvian and English).

[9] I. Lace, A. Spalvins, J. Slangens, *Algorithms for Accounting Groundwater Discharge in the Regional Hydrogeological Model and Interpolating of Simulation Results at Observation Wells*, in Proceedings of International Seminar on "Environment Modelling". - Riga - Copenhagen, Vol. 1, 1995, pp. 201 - 216, (Boundary Field Problems and Computers, 36-th issue).

[10] A. Spalvins, I. Semjonovs, E. Gosk, I. Gobins, O. Aleksans, *Development of a mathematical model for contamination migration in the area of the sulphur-tar sludge waste pools in Incukalns, Latvia*, in Proceedings of XXIX International Association of Hydrogeologists Congress on "Hydrogeology and Land use Management", Bratislava, 1999, pp. 253 - 258.