Guidelines for validation and verification of groundwater mathematical models (GULLIVER): A case study

S Troisi[1], S. Straface[1], G. Gambolati[2],
M. Putti[2] & C. Paniconi[3]

[1]*University of Calabria, Italy*
[2]*University of Padua, Italy*
[3]*CRS4, Cagliari, Italy*

Abstract

The goal of this work is to define a set of guidelines for "validating" mathematical models that are to be used in real-world problems of subsurface flow and contaminant transport. Groundwater models today are being used to study increasingly complex problems involving large temporal and spatial scales and with many uncertainties inherent in the data and in the models themselves. It becomes important therefore to establish procedures and benchmark tests with which to assess a given model's adequacy for simulating a specific groundwater problem. In the GULLIVER project we propose to create a public library of phenomenological schemes (PS), each one satisfactorily solved by an established model. A user with a specific application and model at hand can select the PS that most closely corresponds to his application, and corroborate his model against the established solution. The user's model is considered validated when the two numerical solutions match according to specified criteria. In this paper we describe this methodology and present its application to a 2D saturated flow and transport test case based on data from a field study site in Calabria, Italy.

1 Introduction

In addressing groundwater protection and remediation problems, a deep dichotomy exists between the research world, where groundwater problems are approached from a phenomenological point of view, and real-world practitioners or model end-users, who have to confront the problems in all their data, parameter, and boundary condition complexities [*Di Silvio* 1992]. This imbalance has hindered the full application of advanced numerical techniques to real cases. There is a wide literature that demonstrates a close link between between mathematical model and data; if the aim of a mathematical model is to efficiently and reliably describe the behaviour of a physical system under different stresses and forcings, the aim of field study is to obtain accurate measurements (directly, or, more commonly in the case of groundwater, indirectly) of the characteristic parameters of a system.

Reliability of a model solution depends mainly on the model being "valid", in the sense of adequately representing the underlying physical processes of interest and being based on accurate and consistent numerical algorithms, but it also depends on data precision, in the same way that data reliability depends to some extent on model requirements. The phenomenological model through which the characteristic parameters are determined must be the same as the phenomenological model used in simulation, and the model is sensitive to those parameters directly involved in the equations governing the phenomenon [*Troisi* 1995]. The issue of the degree of reliability required of a mathematical model that is used to solve a groundwater problem raises long-standing questions concerning whether groundwater models can actually be validated [*Konikow and Bredehoeft* 1992; *McCombie and McKinley* 1993]. It is broadly accepted that a mathematical model can be a useful too in addressing groundwater flow and pollution problems. A similar consensus is lacking however in attempts to establish procedures and rules for the validation of such models. Indeed there seems to be little agreement both on what validation is and how a theory, and thus a model, can be validated.

A way out of this dilemma is to focus not on verification of a model's validity (valid model), but to assess instead the usefulness of a model in its specific context (validated model): "...a model is developed for real problems and applied to specific situations, it is entirely instrumental. We have not to question if it is valid, only if it is useful - we validate, not verify" [*MISSING-REF-Raitt* 1979]. In other words validation should "guarantee" the ability of a model to make reliable predictions. It should thus be viewed as a verification process, whereby we assess whether the model has an adequate predictive accuracy within its domain of applicability and over characteristic time and space scales.

2 The GULLIVER project

Within this context of "validation", the goal of this work is to define guidelines for assessing and evaluating mathematical models used in hydrogeology for simulation of groundwater flow and transport processes. The GULLIVER project proposes to create a public library of "phenomenological schemes" (PS) which have been solved satisfactorily by established models used in academic and research centres. A user can validate a groundwater model solving the PS that most closely corresponds to the specific problem at hand (e.g., unsaturated flow, 2D groundwater flow with injection/extraction points, density-dependent flow and transport, etc). The user's model is considered validated when the two numerical solutions match according to specified criteria.

Each PS includes a PS proposal set and a PS solution set. The PS proposal set contains:

- A metadata description of the problem;
- An accurate graphical and numerical description of the problem (flow and transport domain, boundary conditions, parameters, load conditions, pollutant types, etc).

The PS solution set includes:

- One or more analytical, experimental, or numerical solutions;
- Data validation and prediction intervals with given probabilities;
- Detailed description of the simulation results;
- A bigliography for the mathematical models used;
- A listing of the commercial codes that have been validated on the specific PS.

The proposed validation procedure will benefit the large community of users needing to adopt simulation models to address groundwater management, protection, and remediation problems, providing them with better indications of the reliability of various models. We limit ourselves in this paper to presenting a case study of 2D saturated flow and transport based on data from a field site in Calabria, Italy. A numerical model based on the method of cells [*MISSING-REF-Troisietal* 1999] is validated against a finite element reference model [*Gambolati et al.* 1993].

3 Numerical simulation models

3.1 Governing equations

Using indicial notation where the indices i and j denote summation over the three coordinate dimensions $(i, j = 1, 2, 3)$, the equation for groundwater flow can be written as

$$\frac{\partial}{\partial x_i}\left(k_{ij}\frac{\partial h}{\partial x_j}\right) = S_s\frac{\partial h}{\partial t} - q \tag{1}$$

where k_{ij} is the hydraulic conductivity tensor, x_i is the ith Cartesian coordinate, S_s is the specific elastic storage of the medium, t is time, and q represents distributed source or sink terms. Initial and boundary conditions for equation (1) can be expressed as

$$h\,(x_i, 0) \;=\; h_o(x_i) \tag{2}$$

$$h(x_i, t) \;=\; \overline{h}(x_i, t) \qquad \text{on } \Gamma_1 \tag{3}$$

$$v_i n_i \;=\; -q_n(x_i, t) \qquad \text{on } \Gamma_2 \tag{4}$$

where h_o is the potential head at time $t = 0$, $\overline{h}$ is the prescribed head on boundary Γ_1, and v_i is the Darcy velocity

$$v_i = -k_{ij}\frac{\partial h}{\partial x_j} \tag{5}$$

on the complementary boundary Γ_2, with n_i being the direction cosine of the outward normal to the boundary and q_n being the prescribed flux on Γ_2. We use the sign convention of q_n positive for an inward flux and negative for an outward flux. Boundary condition (3) is of Dirichlet type while (4) is of Neumann type.

Using the Darcy velocities v_i given by (5) we can write the conservation equation for the contaminant. With the dispersion part of the contaminant velocity expressed as $-D_{ij}(\partial c/\partial x_j)$, we obtain the equation describing transport of a non-reactive solute

$$\frac{\partial}{\partial x_i}(D_{ij}\frac{\partial c}{\partial x_j}) - \frac{\partial}{\partial x_i}(v_i c) = n\frac{\partial c}{\partial t} - qc_1 - f \tag{6}$$

where $D_{ij} = n\tilde{D}_{ij}$, n is the porosity of the medium, $\tilde{D}_{ij}$ is the dispersion tensor, c is the concentration of the solute, c_1 is the concentration of solute injected (withdrawn) with the fluid source (sink), and f is the distributed flow rate of the solute per unit volume. The dispersion tensor for a two-dimensional porous medium is given by [*Bear* 1979]

$$D_{ij} = \alpha_T\,|v|\,\delta_{ij} \;+\; (\alpha_L - \alpha_T)\frac{v_i v_j}{|v|} \;+\; D_o\,n\,\tau\,\delta_{ij} \tag{7}$$

where α_L is the longitudinal dispersivity, α_T is the transverse dispersivity, δ_{ij} is the Kronecker delta, τ is the tortuosity ($\tau = 1$ usually assumed), D_o is the molecular diffusion coefficient, and $|v| = \sqrt{v_1^2 + v_2^2}$. The initial and boundary conditions for equation (6) can be expressed as

$$c(x_i, 0) = \qquad c_o(x_i) \tag{8}$$

$$c(x_i,t) = \overline{c}(x_i,t) \qquad \text{on } \Gamma_3 \tag{9}$$

$$D_{ij}\frac{\partial c}{\partial x_j}n_i = q^d(x_i,t) \qquad \text{on } \Gamma_4 \tag{10}$$

$$(v_i c - D_{ij}\frac{\partial c}{\partial x_j})n_i = q^c(x_i,t) \qquad \text{on } \Gamma_5 \tag{11}$$

where c_o is the initial concentration, $\overline{c}$ is the prescribed concentration on the Dirichlet boundary Γ_3, q^d is the prescribed dispersive flux normal to the Neumann boundary Γ_4 (positive inward), and q^c is the prescribed flux of solute across the Cauchy or Rubin boundary Γ_5.

3.2 Finite element model

In the finite element discretization of the flow equation an approximate solution $\hat{h}$ for h is defined as

$$h \approx \hat{h} = \sum_{i=1}^{l} \hat{h}_i(t)N_i(x_1,x_2,x_3) \tag{12}$$

where $N_i(x_1,x_2,x_3)$ are linear shape or basis functions for three-dimensional tetrahedral finite elements, $\hat{h}_i$ are the unknown nodal potential heads, and l is the number of nodes. Substituting $\hat{h}$ in equation (1) and prescribing an orthogonality condition between the residual

$$L(\hat{h}) = \frac{\partial}{\partial x_i}\left[k_{ij}\left(\frac{\partial \hat{h}}{\partial x_j}\right)\right] - S_s\frac{\partial \hat{h}}{\partial t} + q \tag{13}$$

and each basis function yields the Galerkin integral:

$$\int_V L(\hat{h})N_i\,dV = 0, \qquad i = 1,\ldots,l \tag{14}$$

where V is the flow region or integration domain. Assuming the coordinate directions x_j to be parallel to the principal directions of hydraulic anisotropy and applying Green's lemma we get

$$\begin{aligned} -\int_V &\left[k_{11}\frac{\partial \hat{h}}{\partial x_1}\frac{\partial N_i}{\partial x_1} + k_{22}\frac{\partial \hat{h}}{\partial x_2}\frac{\partial N_i}{\partial x_2} + k_{33}\left(\frac{\partial \hat{h}}{\partial x_3}\right)\frac{\partial N_i}{\partial x_3}\right]dV \\ +\int_\Gamma &\left[k_{11}\frac{\partial \hat{h}}{\partial x_1}n_1 + k_{22}\frac{\partial \hat{h}}{\partial x_2}n_2 + k_{33}\left(\frac{\partial \hat{h}}{\partial x_3}\right)n_3\right]N_i\,d\Gamma - \\ &\int_V S_s\frac{\partial \hat{h}}{\partial t}N_i\,dV + \int_V qN_i\,dV = 0, \qquad i = 1,\ldots,l \end{aligned} \tag{15}$$

where n_i, $i = 1,2,3$ are the direction cosines of the outer normal to the boundary Γ.

In matrix form this equation can be written as

$$H\hat{h} + P\frac{\partial \hat{h}}{\partial t} + q^* = 0 \tag{16}$$

where H is the stiffness matrix, P is the capacity matrix, and vector q^* accounts for the prescribed boundary flux, the withdrawal or injection rate, and the coupling with the transport equation. Integration in time of (16) by a weighted finite difference scheme yields

$$\begin{aligned} &\left(\nu H_{t+\Delta t} + \frac{\nu P_{t+\Delta t} + (1-\nu)P_t}{\Delta t}\right)\hat{h}_{t+\Delta t} = \\ &\left(\frac{\nu P_{t+\Delta t} + (1-\nu)P_t}{\Delta t} - (1-\nu)H_t\right)\hat{h}_t - \nu q^*_{t+\Delta t} - (1-\nu)q^*_t \end{aligned} \tag{17}$$

where for stability reasons the weighting parameter ν must satisfy the condition $0.5 \leq \nu \leq 1$. Selecting $\nu = 0.5$ leads to the Crank-Nicolson scheme while $\nu = 1$ gives the fully implicit backward difference scheme.

Finite element integration of the transport equation (6) follows the same procedure as the flow equation, and yields the discretized system of equations

$$(A + B + D)\hat{c} + C\frac{\partial \hat{c}}{\partial t} + r^* = 0 \tag{18}$$

where matrices A, B, and D account for the dispersion, convection, and Cauchy boundary condition terms, respectively. The vector r^* accounts for point and distributed contaminant sources and sinks. Integration in time of equation (18) is again performed using a weighted finite difference scheme:

$$\begin{aligned} &\left[\nu(A+B+D)_{t+\Delta t} + \frac{\nu C_{t+\Delta t} + (1-\nu)C_t}{\Delta t}\right]\hat{c}_{t+\Delta t} = \\ &\left[\frac{\nu C_{t+\Delta t} + (1-\nu)C_t}{\Delta t} - (1-\nu)(A+B+D)_t\right]\hat{c}_t - \nu r^*_{t+\Delta t} - (1-\nu)r^*_t \end{aligned} \tag{19}$$

where as before $1/2 \leq \nu \leq 1$. Unlike the flow equation, the transport equation is quite sensitive to the value of the weighting parameter ν. A value close to 1/2 leads to accurate but unstable solutions while values close to 1 yield good stability but with larger numerical dispersion [*Peyret and Taylor* 1983].

3.3 The method of cells

If we consider a cell complex in which the most nonuniform cells are the smallest, we may use as an approximation the same constitutive law used in the differential context. Since a numerical process produces approximate expressions of physical laws, consideration of small reasonably uniform regions enables us to obtain a natural discretized form. The forming of densities and rates and the passage to the limit to form field functions, typical of field and continuum theories, deprives physical variables of their geometrical content. This renders arbitrary any discretization process restored from the differential formulation. From these considerations it follows that to obtain a discrete formulation of the fundamental equation of a physical theory, it is not necessary to go down to the differential form and then to go up again to the discrete form. It is enough to apply the elementary physical laws to small regions in which the uniformity of the field is attained to a sufficient degree, linked to the accuracy of the input data and to the degree of approximation requested of the solution.

Let T be a triangulation of the domain into N cells. For example, a 2D domain may be discretized via a Delaunay triangulation. In this case the cells are the Voronoi polygons, whose union is the dual of the Delaunay triangulation. Note that by definition the boundary of a Voronoi cell is orthogonal to the edge of a Delaunay triangle. For each cell $\bar{v}_h$ the discrete equation enforcing conservation of the fluid mass m_h in the time interval $\bar{\Delta}t = \bar{t}_{n+1} - \bar{t}_n$ is

$$\bar{\Delta}_t m_h + \sum_\alpha \bar{d}_{h\alpha} \Phi_\alpha = 0 \tag{20}$$

where $\bar{\Delta}_t m_h = (m_h(\bar{t}_{n+1}) - m_h(\bar{t}_))/\bar{\Delta}t$ and Φ_α represent the mass flux across boundary α of $\bar{(v)}_h$. The discrete form of Darcy's law (5) over boundary α can be written as

$$q_\alpha = K \frac{H_\alpha}{\ell_\alpha} \tag{21}$$

where H_α, called the head coboundary, is the head difference discretizing the opposite of the gradient normal to boundary α and ℓ_α is the size of the boundary of the Delaunay simplex (the triangle edge). The mass flux Φ_α is then given by

$$\Phi_\alpha = \rho q_\alpha s_\alpha \tag{22}$$

where s_α is the size of the boundary α of the Voronoi cell. If this direction does not coincide with the principal axis of anisotropy, a rotation of the local reference frame is performed to maintain the diagonal form of the conductivity tensor.

For a confined aquifer, we can write:

$$\rho S_s \bar{v}_h \bar{\Delta}_t h_h + \sum_\alpha \Phi_\alpha = 0 \tag{23}$$

where for simplicity we have written $\bar{v}_h$ to indicate the volume of cell $\bar{v}_h$. This equation can be written for every cell of the domain discretization, thus yielding a system of equations with h_h as unknowns

$$\bar{H} h_h + \bar{P} \bar{\Delta}_t h_h = \bar{q} \tag{24}$$

where $\bar{H}$ and $\bar{P}$ are the stiffness and capacity matrices of the method of cells. This system can be solved using an approach entirely similar to the Crank-Nicolson scheme used for the finite element equation (17) with $\nu = 0.5$.

The discrete form of the solute mass conservation equation may be written as

$$\bar{\Delta}_t m_{c_h} + \sum_{\alpha} \bar{d}_{h\alpha} \Phi_{c_\alpha} = \Phi^*_{ch} \tag{25}$$

The solute mass variation in time is

$$\bar{\Delta}_t m_{c_h} = n\, \rho_s\, \bar{v}_h\, \frac{C(t_{n+1}) - C(t_n)}{t_{n+1} - t_n} \tag{26}$$

where ρ_s is the density of the solute and C_h is the concentration within cell $\bar{v}_h$. The total flux Φ_{c_α} is the sum of the dispersive flux $\Phi_{c_\alpha}^{\text{disp}}$ and the convective flux $\Phi_{c_\alpha}^{\text{conv}}$ The latter is given by

$$\Phi_{c_\alpha}^{\text{conv}} = \frac{\Phi_\alpha}{\rho}\, C_h \tag{27}$$

while the dispersive flux is

$$\Phi_{c_\alpha}^{\text{disp}} = D_\alpha \bar{s}_\alpha \frac{C_\alpha}{l_\alpha}\, \rho_s \tag{28}$$

where C_h is the concentration coboundary and D_α is the projection of the dispersion tensor (7) along the α-direction. If this direction does not coincide with the principal axis of anisotropy, a rotation of the local referance frame is performed to maintain the diagonal form of the dispersion tensor.

Again, assembling the equations for all the cells in the discretized domain yields a system of equations with C_h as unknowns

$$(\bar{A} + \bar{B} + \bar{D})C_h + \bar{C}\bar{\Delta}_t C_h = \bar{f} \tag{29}$$

for which all the considerations reported in the finite element case still apply.

4 Validation exercise

4.1 Description of the test case

The study area (Figure 1) is located near the town of Montalto Uffugo in Calabria, Italy. It is a valley at the confluence of the Settimo River on the south, the Mavigliano River on the north and the Crati river on the east. The study area is underlain by layers of sand (0–7 m), clay (7–11 m), and silt (11 to about 40 m) [*Troisi et al.* 1995]. A basal clay underlies the silty sand. A local perched water table is in the alluvium above the clay layer, and the silty sand layer constitutes a confined aquifer above the basal clay. Measurements of the hydraulic conductivity and its distribution

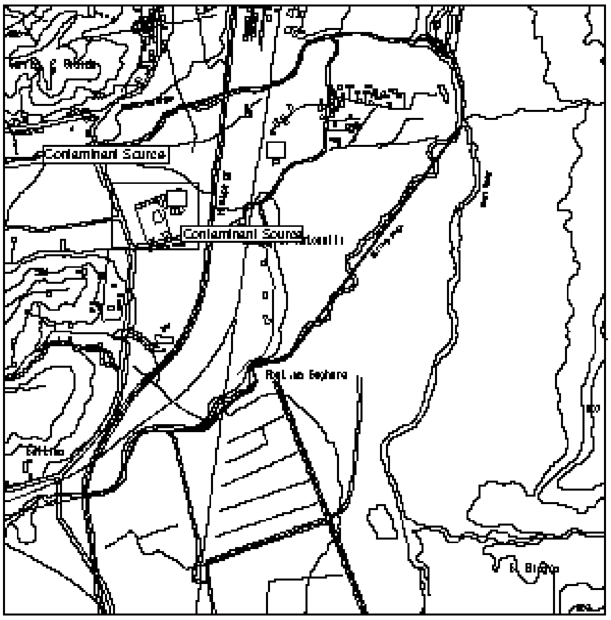

Figure 1: Location of the study area used for the validation exercise.

have been made over the past few years using a kriging with external drift methodology applied to electrical-resistivity data [*Troisi et al.* 2000].

The domain for the flow and transport model application was defined as the area enclosed by the surface water system as shown in Figure XXXX. In the flow model a heterogeneous isotropic porous medium is assumed. Dirichlet boundary conditions were are according to the water levels in the surface water system. Several pumping tests were carried out to estimate aquifer hydraulic conductivity under steady and unsteady flow conditions. A well of depth 40 m within the central monitoring unit of the groundwater monitoring system located in the study area was pumped at a rate of 0.002 m^3/s. The actual steady state drawdown caused by this pumping was measured to be about 6 m at the central well of the monitoring system.

4.2 Comparison of model results

TO COMPLETE

5 Conclusions

TO COMPLETE

References

Bear, J., *Hydraulics of Groundwater.* McGraw-Hill, New York, NY, 1979.

Di Silvio, G., Schemi concettuali e previsioni quantitative in idraulica fluviale. In: *Atti XXIII Convegno di Idraulica e Costruzioni Idrauliche, Vol. 5*, Firenze, pp 97–104, 1992.

Gambolati, G., C. Paniconi and M. Putti, Numerical modeling of contaminant transport in groundwater. In: Petruzzelli, D. and F. G. Helfferich (eds.) *Migration and Fate of Pollutants in Soils and Subsoils.* Springer-Verlag, Berlin. Volume 32 of *NATO ASI Series G: Ecological Sciences*, pp 381–410, 1993.

Konikow, L. F. and J. D. Bredehoeft, Ground-water models cannot be validated, *Adv. Water Resour.* 15, 75–83, 1992.

McCombie, C. and I. McKinley, Validation – another perspective, *Ground Water* 31(4), 530–531, 1993.

MISSING-REF-Raitt, X., ???, *???* ???(???), ???–???, 1979.

MISSING-REF-Troisietal, X., ???, *???* ???(???), ???–???, 1999.

Peyret, R. and T. D. Taylor, *Computational Methods for Fluid Flow.* Springer-Verlag, New York, 1983.

Troisi, S., Problems of mathematical model validation in groundwater hydrology, *Excerpta* 9, 1995.

Troisi, S., C. Fallico, M. Maiolo and R. Coscarelli, La caratterizzazione dell'acquifero interessato da un campo prove per lo studio sperimentale di fenomeni idrodispersivi in mezzi porosi. In: *Quaderni di Geologia Applicata, Vol. 1.* Pitagora Editrice, Bologna, pp 41–53, 1995.

Troisi, S., C. Fallico, S. Straface and E. Migliari, Application of kriging with external drift to estimate hydraulic conductivity from electrical-resistivity data in unconsolidated deposits near Montalto Uffugo, Italy, *Hydrogeology Journal of the U.S. Geological Survey* 8(3), 2000.