Incorporating Initial Conditions in the Model Calibration Process

John Doherty
Watermark Numerical Computing, Brisbane, Australia.

ABSTRACT
Means by which estimation of initial heads and concentrations over the entirety of a model domain is incorporated into the model calibration process are discussed. This brings the benefits of reduced potential error in parameters employed by the model, and of predictions made by it. It also allows better quantification of the possible extent of these errors. A two step process is briefly described. First the location of a notional freshwater-saltwater interface is parametrically defined. Second, these parameters are estimated simultaneously with aquifer hydraulic properties using regularised inversion techniques which can accommodate the estimation of large numbers of parameters. A maximum likelihood solution to the combined inverse problem is sought through use of appropriate Tikhonov regularisation constraints which collectively attempt to enforce a set of initial conditions which are as hydraulically realistic as possible.

INTRODUCTION
One of the most difficult tasks involved in construction of a saltwater intrusion model is the assignment of initial conditions. Realistic heads and concentrations must be assigned to every cell in the model domain. To make matters more difficult, these are often the outcome of unknown stresses over unknown amounts of time. Furthermore, spatial sampling of these quantities is often very poor.

If initial conditions are wrongly assigned, the model will move water and salt in order to correct this misassignment, at the same time as it moves them in response to stresses represented in the model. The problem is exacerbated when a model is run for calibration purposes. Parameter estimates may be compromised if initial conditions are wrongly assigned, and if hydraulic property estimates are the only quantities that are allowed to vary through the parameter estimation process in order to minimise misfit between model outputs and historical field measurements. This problem is compounded by the fact that spatial variation of initial heads and concentrations is often strongly dependent on spatial variation of subsurface hydraulic properties. Assignment of the former independently of a calibration process which seeks to estimate the latter is unlikely to result in initial head and concentration fields that are physically plausible on the one hand, and acceptable to the model on the other hand.

If initial conditions are uncertain, then (like any other uncertain model input) their estimation should be included in the model calibration process. Reasons for this include the following.

1. Field data that is directly or indirectly informative of the initial disposition of groundwater heads and salt concentrations can be included in the calibration dataset, and the information content of this data with respect to these dispositions thereby “tapped”; this will hopefully result in better assignment of initial conditions than would otherwise be the case.
2. With better assignment of initial conditions, parameter estimates will be better, because their need to play a compensatory role for misassignment of initial conditions is thereby mitigated.
3. With uncertainty of initial conditions recognised by their inclusion in the calibration process, this uncertainty (together with that of correlated hydraulic properties) can be quantified as a by-product of the parameter estimation process.

4. Quantification of parameter uncertainty then leads to quantification of predictive uncertainty. Removal of initial conditions from the purview of the calibration and predictive uncertainty analysis processes may result in serious underestimate of potential predictive error as it depends directly on initial conditions, and/or on parameters whose estimates may be compromised through misassignment of initial conditions.

Two problems must be addressed, however, if initial conditions are to be included in the parameter estimation process. The first is that of how to represent these conditions parametrically so that the parameter estimation process is provided with "handles" through which it can alter them. The second is that of accommodating the large number of parameters that parametric representation of initial conditions may entail.

PARAMETRIC REPRESENTATION OF INITIAL CONDITIONS

Where a simulated groundwater system is comprised of one or a number of aquifers separated from each other by aquitards of lower hydraulic conductivity, the disposition of salt and fresh water in each aquifer is often described, qualitatively at least, using the freshwater-saltwater interface concept. In truth, fresh and salt water are not separated by a sharp interface, as there is a significant mixing zone between the two which is the site of complex flow processes. Furthermore, even to the extent that an "interface" can be identified, its landward extension can vary along the coast in response to local pumping and aquifer heterogeneity.

Nevertheless, because it is easier to describe a surface parametrically than a three-dimensional body, the interface concept is useful. The diffuse nature of salt concentrations in the vicinity of the "interface" can then be described using parameters which govern the degree of "concentration spread" about it. Above the interface, salt concentrations can be described as asymptotically approaching some (possibly location-dependent) freshwater background concentration with increasing elevation, while below the interface concentrations asymptotically approach a (possibly location-dependent) background saltwater concentration with decreasing elevation. The width of the interface (this being determined by the rate of approach to asymptotic concentrations both above and below it) then becomes a parameter that is adjustable through the parameter estimation process. This too can be location-dependent if desired.

In numerical experiments conducted to date, pilot points have been employed for parametric description of the freshwater-saltwater interface. Pilot points have often been used as parametric descriptors of two-dimensional hydraulic property variability. They were introduced by Certes and de Marsily (1991) and have been employed by many others since then. Doherty (2003) first combined their use with regularised inversion (see below), allowing the estimation of many more pilot point parameters than had hitherto been possible. Tonkin and Doherty (2005) illustrate the use of hundreds of pilot points in a complex multi-layer setting.

In work undertaken so far, pilot points used for parametric description of the freshwater-saltwater interface have been placed in two rows within each aquifer represented in the model – one row being in front of the toe of the freshwater-saltwater interface and the other being behind its head. The latter position is often well defined in a model as it coincides with the aquifer's maximum seaward extent. The location of the interface toe is often only poorly known. However this does not matter, for it is part of the role of the parameter estimation process to assess where this may be; it is only necessary that pilot points be placed somewhere near this toe, preferably on its landward side.

Elevations are then assigned to all pilot points. The elevation of the freshwater-saltwater interface in every pertinent model cell is then computed through spatial interpolation between the rows of pilot points. If interpolation leads to an interface elevation that is locally above or below the top or bottom of the aquifer this does not matter; fresh and saltwater are dispersed about the interface in the usual manner, but concentrations are only assigned to cells that lie within the hydrostratigraphic unit to which the pilot points (and hence the interface) are assigned. Note that this unit can be comprised of many model layers.

Elevations assigned to pilot points are adjusted through the calibration process. So too are the “spreading parameters” that govern the sharpness (or otherwise) of the freshwater-saltwater interface.

If using the SEAWAT (Langevin et al, 2003) model, computation of a head field that is complimentary to an interface-derived concentration field can be accomplished through running the model for a single steady-state stress period in which an approximation to historical time-averaged stresses is supplied. Computation time for this stress period is normally minimal; if necessary, the MODFLOW component of SEAWAT can be de-coupled from its MT3DMS component for this stress period to lower the burden of initial head computation even further.

PARAMETER ESTIMATION

Hunt et al (2007) describe the advantages of highly parameterised inversion as a vehicle for automatic parameter estimation, and briefly discuss the means through which this can be achieved in a manner that results in estimated parameter fields that are as realistic as possible. In contrast to classical approaches to parameter estimation which rely on the “principle of parsimony”, the regularised inversion approach allows inclusion in the parameter estimation process of as many parameters as are needed to extract all information worth extracting from a calibration dataset. Those parameters, or parameter combinations, for which information is lacking, are simply left unaltered, and/or are assigned values that are enlightened by “plausibility constraints” built into the parameter estimation process.

The hybrid inversion methodology described by Tonkin and Doherty (2005) and available through PEST (Doherty, 2007) implements highly parameterised inversion through combining two different, but complimentary, regularisation techniques, namely the subspace and Tikhonov methods. The former allows subdivision of parameter space into orthogonal estimable and inestimable components. The latter allows re-formulation of the inverse problem as one of minimisation of misfit between model outputs and field measurements subject to constraints of maximum parameter plausibility. Through use of “super parameters” which collectively span estimable parameter space, the number of model runs required per iteration of the parameter estimation process is then reduced to a number which is far smaller than the number of parameters that are actually defined (which can number in the hundreds, or even thousands). Use of this methodology in the saltwater intrusion modelling context allows simultaneous estimation of interface parameters, together with (pilot-point-based) aquifer/aquitard hydraulic property parameters to take place with a numerical burden that is little greater than that of classical parameter estimation, but is far more stable, and infinitely more flexible, thus allowing the benefits outlined above of including initial conditions in the parameter estimation process to be realised.

Tikhonov constraints on initial condition parameters that have been employed successfully to date include the following:

1. maximum adherence of initial concentration fields to "training fields" computed using model runs undertaken on the basis of likely hydraulic property distributions;
2. (where appropriate) minimal changes in concentrations within all active cells of the model domain over an initial model stress period in which model inputs approximate pre-run, long-term stresses.

UNCERTAINTY ANALYSIS

The matter of uncertainty analysis was cited above as one of the benefits of including initial conditions in the calibration process. In contrast to traditional parameter estimation and uncertainty analysis, inestimability of certain parameters (or parameter combinations) does not require that such parameters (or parameter combinations) be omitted from this process. Nor does it compromise computation of values for estimable parameters. In fact if inestimable (and hence uncertain) parameters are *not* represented in these processes, achievement of maximum likelihood solutions to the inverse problem of model calibration becomes less likely, while uncertainty analysis becomes impossible. Once included however it then becomes a simple matter to quantify (with integrity) parameter and predictive uncertainty, the contribution to this uncertainty by different parameter types (which now include initial conditions), and the potential benefits of different data acquisition strategies in reducing predictive uncertainty.

REFERENCES

Certes, C., and de Marsily, G., 1991. Application of the pilot point method to the identification of aquifer transmissivities, *Advances in Water Resources*, 14(5), 284-300.

Doherty, J., 2003. Ground water model calibration using pilot points and regularization, *Ground Water*, 41(2), 170-177.

Doherty, J., 2007. PEST: Model Independent Parameter Estimation. Watermark Numerical Computing, Brisbane, Australia.

Hunt, R.J., Doherty, J, and Tonkin, M.J., 2007. Are models too simple? Arguments for increased parameterisation. *Ground Water* 45 (3), 254–262

Langevin, C.D., Shoemaker, W.B., and Guo, Weixing, 2003, MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model–Documentation of the SEAWAT-2000 Version with the Variable-Density Flow Process (VDF) and the Integrated MT3DMS Transport Process (IMT): U.S. Geological Survey Open-File Report 03-426, 43 p.

Tonkin, M. and Doherty, J., 2005. A hybrid regularised inversion methodology for highly parameterised models. *Water Resources Research*. Vol. 41, W10412, doi:10.1029/2005WR003995.

Contact Information: John Doherty, Watermark Numerical Computing, 336 Cliveden Avenue, Corinda, 4075, Australia. Phone: +61 7 3379 1664. Email: johndoherty@ozemail.com.au