DOI 10.7343/as-2016-231

The role of numerical models in environmental decision-making

Il ruolo dei modelli numerici nel processo decisionale ambientale

Francesca Lotti - Kataclima S.r.l., Vetralla (VT), Italy - f.lotti@kataclima.com
John Doherty - Watermark Numerical Computing Pty Ltd., Brisbane, Australia

Keywords: *environmental modelling, calibration, uncertainty analysis, decision-making.*

Parole chiave: modellazione ambientale, calibrazione, analisi dell'incertezza, processo decisionale.

Modelling is undertaken for a variety of reasons. The present short discussion focuses on everyday model usage as undertaken by government and consulting companies to assist in environmental management. In this context, a model's purpose is to predict the behaviour of a system under a management regime which is not yet in place or has just started. The term *management regime* may refer to excavation of a mine, extraction of water, contaminant remediation, allocation of water for irrigation, etc. A common aspect of all these practices is evaluation of the risk of *bad things* happening. The decision-making process should include exploration of the costs and benefits associated with alternative management practices, and quantification of risks accompanying different management decisions.

Numerical models, being able to make predictions, play the latter role. It is important to understand that any model prediction cannot be made with certainty, even if a model has been *well calibrated*. It therefore follows that a model *cannot* predict what will happen in the future. However, *following proper uncertainty analysis, it may be able to predict what will NOT happen in the future*. This is an important statement with important implications.

Environmental management is undertaken to prevent the outcomes of human activity from having unwanted consequences (*bad things*). Given that model predictions are necessarily probabilistic in nature, how can they support environmental management?

To use models properly in the decision-making context, a change in the philosophy of model usage away from the current *crystal ball* mentality is required. Environmental managers must refrain from asking *What will happen if we do this or that?* Instead they should ask *If we implement this particular management strategy, can it be demonstrated that this costly bad thing will NOT happen?* and *if it MAY happen, how can we get early warning, and what measures can we take to prevent it?* The latter course of action is often referred to as *adaptive management*.

How can the possibility of *bad things* happening be explored with a model? The unlikelihood of occurrence of a *bad thing* can be demonstrated by showing that its occurrence is incompatible with expert knowledge and/or with the historical behaviour of the system. A model is unique in its ability to do this. However, it cannot do it alone; to achieve this, a model must be used in conjunction with software [such as PEST (Doherty, 2015), and geostatistical packages], which, through inversion and uncertainty analysis, can demonstrate such incompatibilities.

It is still commonplace for a calibrated model to be used for predictive purposes without uncertainty analysis. *Once we calibrate the model, it will be fine. We provide the model with correct boundary conditions, we calibrate the parameters on the basis of historical data, and then we obtain a perfect support for decisions.* This is, of course, fantasy.

A model requires thousands, maybe millions, of parameters. The earth and processes associated are heterogeneous. It is simply impossible to supply the correct values for the thousands of parameters that a *perfect* simulator of environmental processes would require. So why is *calibration* such an established part of modelling culture? Part of the answer lies in wishful thinking – the same thinking that has underpinned all attempts by all societies over all ages to prophesize the future. Part of it lies in the fact that calibration, and the quest for parameter uniqueness, is an easier problem to solve numerically than quantification of post-history-matching probability distributions.

Unfortunately, reality rarely conforms to the dictates of wishful thinking; the inability to uniquely parameterize it does not make complexity go away. A modeller can neglect parameterization complexity only if he/she wishes to weaken a model's ability to quantify predictive uncertainty. But given the role of risk assessment in decision-making, why would he/she want to do that?

Acceptance of the notion that *calibration* cannot provide a model with prophetic abilities induces us to abandon parameter parsimony as a parameterization philosophy, and turn to the use of underdetermined methods. It is contended that parsimony and highly parameterized methods are not ends in themselves; they are strategies that a modeller implements as part of his/her desire to achieve a greater end. The validity of either strategy in any particular context must therefore be judged according to its capacity to achieve that end.

Highly-parameterized methods and characterization of history-matching as an ill-posed inverse problem define a working environment where the comforting notion of parameter uniqueness is abandoned, and where we must learn to live in a world of mingled parameter relationships. But still, amidst this chaos, we can begin to answer some important questions, and gain important insights into what a model can and cannot achieve, as we strive to understand and manage the environment (Doherty and Simmons, 2013).

But even if we reject the traditional approach to calibration based on parameter parsimony, we are still faced with a decision of how much complexity we should include in a model, for our computing resources will always be limited. There will always be compromises. The point is that the level of complexity must be set according to its salience to decision-informative predictions; decision-relevant complexity should not be eschewed simply because it cannot be parameterized uniquely by a limited calibration dataset. If this approach is not adopted, a model's ability to contribute to the decision-making process will be compromised. Specifically, if a prediction of interest is sensitive to a specific system detail, that detail must be represented in the model. Since all parameters required to represent such detail can hardly be estimated uniquely, they must be represented on a stochastic basis. The importance of uncertainty analysis should be obvious.

Highly parameterized, model-based uncertainty analysis can form a powerful basis for open and informed environmental decision-making. That analysis also implies a large subjective component, requiring the exercise of expert judgement at every turn. This is a good thing, for it ensures that the modelling process becomes an instrument for making inquiries, and forms a basis for open debate amongst competing hypotheses, rather than a laboratory instrument expected to give the "right answer". If modelling is done in an open manner, with different stakeholders involved in the inquiry process, it will inevitably lead to more effective resolution of disputes than that which often accompanies model-based decision-making today.

Data processing and uncertainty analysis in which system complexity is acknowledged and numerically accommodated (supported by packages such as PEST) must include a high degree of user-involvement. The subjective component of uncertainty analysis thus implies that matching the historical behaviour of a system, and predicting its future behaviour, must be a partnership between the software and the expert. Modellers are required to understand the systems which they manage, characterize numerically what they know and what they do not know, and explain to decision makers the risks associated with alternative courses of action as these arise from an incomplete knowledge of the complex system which is being managed.

In conclusion, it seems that environmental modelling is a murky business – a business that will always be subjective and hence a hotbed for arguments. But if models can serve to focus those arguments on issues that matter – issues that science rather than rhetoric may ultimately resolve - then these arguments will have served the decision-making process well.

This is a far more productive outcome of model usage than that which seeks to award one duelling model victory over another, when both in fact make false and undeserving claims to prophetic visions of the environmental future.

RFERENCES

Certes C, de Marsily G (1991) Application of the pilot-points method to the identification of aquifer transmissivities. Advances in Water Resources 14(5):284-300. doi:10.1016/0309-1708(91)90040-U

Doherty J (2003) Groundwater model calibration using Pilot Points and Regularization. Ground Water 41(2):170-177.

Doherty J (2008) Guidelines for groundwater model calibration using regularized inversion. Watermark Numerical Computing, Brisbane, Australia.

Doherty J (2015) Calibration and uncertainty analysis for complex environmental models - PEST: complete theory and what it means for modelling the real world. Watermark Numerical Computing, Brisbane, Australia.

Doherty J, Christensen S (2011) Use of paired simple and complex models in reducing predictive bias and quantifying uncertainty. Water Resources Research. doi:10.1029/2011WR010763

Doherty J, Simmons CT (2013) Groundwater modelling in decision support: reflections on a unified conceptual framework. Hydrogeology Journal 21:1531-1537. doi:10.1007/s10040-013-1027-7

Fienen MN, Muffels CT, Hunt RJ (2009) On constraining pilot point calibration with regularization in PEST. Ground Water 47(6):835-44. doi:10.1111/j.1745-6584.2009.00579.x

LaVenue M, de Marsily G (2001) Three-dimensional interference test interpretation in a fractured aquifer using the pilot-point inverse method. Water Resources Research 37(11):2659-2675. doi:10.1029/2000WR000289

Marsily G de (1978) De l'identification des systems hydrologiques "*Identification of hydrogeological systems*". PhD thesis, University of Paris VI, Paris, France.

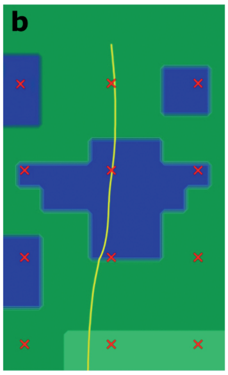
Marsily G de, Lavedan G, Boucher M, Fasanino G (1984) Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, In: Verly G, David M, Journel AG, Marechal A (eds.) Geostatistics for natural resources characterization. Part 2. Springer, Amsterdam, The Netherlands, pp 831-849.

Moore C, Doherty J (2005) The role of the calibration process in reducing model predictive error. Water Resources Research 41(5):W05020. doi:10.1029/2004WR003501

RamaRao BS, LaVenue AM, de Marsily G, Marietta MG (1995) Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. Theory and computational experiments. Water Resources Research 313:475-493.

Tonkin M, Doherty J (2009) Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. Water Resources Research 45:W00B10. doi:10.1029/2007WR006678

Calibrated! But is it that simple?



Some groundwater model calibrations are based on user-defined zones in which hydraulic conductivity is assumed to be uniform. This may be a valid representation of subsurface properties where underground geology is as clearly depictable as surface geology and when each lithotype is in fact homogeneous. In (**a**) the "real" hydraulic conductivity distribution of an aquifer is represented; this aquifer is tapped by 12 observation wells in which the head has been measured (Moore and Doherty, 2005). Zone-based calibration depicted in (**b**) provides good agreement between observed and simulated heads. However its prediction of the time taken for a particle to travel to the southern boundary of the domain is wrong by a large margin. This is not necessarily bad as all models are wrong. What is bad is that when predictive uncertainty is analysed with parameterization based on these zones, the true travel time does not fall within the uncertainty limits.

An alternative parameterization approach can be adopted. The possibility of aquifer heterogeneity can be maintained during the calibration process by using pilot points as a parameterization device (**c**). Pilot points have actually been used for a long time (de Marsily, 1978; Doherty 2003; 2008; Fienen et al. 2009; de Marsily et al. 1984; RamRao et al. 1995; LaVenue et al. 2001; Certes and de Marsily 1991). However they are still not used so widely in everyday modelling, even though their use is growing. A special quality of this parameterization method is its ability to represent the possibility of geological heterogeneity when undertaking calibration and (more importantly) calibration-constrained uncertainty analysis (Tonkin and Doherty, 2009). Calibration still yieds a parameter field that leads to a wrong prediction. But calibration -constrained uncertainty analysis yields e.g.1000 parameter fields, all of which include parameterization detail that is somewhat representative of the real system, and all of which allow the model to fit the 12 heads which comprise the calibration dataset. All of these lead to different particle path and travel time predictions (**d**). A posterior probability distribution of the particle exit time can be constructed from these (**e**). This distribution includes the correct travel time. Decision making can then be based on statements such as "*The travel time of the contaminant is between 15 and 22 years with a probability 0.72*". This is far better than "*The travel time of the contaminant is 20 years, and we can only guess at how much that may be wrong*".