Hydrol. Earth Syst. Sci. Discuss., 3, S1533–S1542, 2006
www.hydrol-earth-syst-sci-discuss.net/3/S1533/2006/

Hydrology and Earth System Sciences Discussions

Interactive comment on "A parameter identifiability study of two chalk tracer tests" *by* S. A. Mathias et al.

S. A. Mathias et al.

Received and published: 22 November 2006

Thank you to Serge Brouyère, Jishan Liu and Jasper Vrugt for their useful and insightful comments about our paper.

A) Response to comments from Serge Brouyère (14/09/2006)

1) Model complexity and parameter identifiability

Brouyère suggests that we consider the "best model" to be the "one that allows fitting as accurately as possible" to "field results... using the most simple model" structure. He then argues that "when modelling, the objective is to translate, as faithfully

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

as possible, the complexity of physical system" of concern. We wholeheartedly agree with the latter and regret that our paper implies the former.

Our main contention is that there is no point in trying to characterise a process (be it the reflection of solutes at the centre of matrix blocks or hydrodynamic dispersion) on the basis of field data that is unaffected by it. The paper proposes that the significance of a process in a given context (e.g. a radially convergent tracer test) can be assessed by comparing results from models that ignore it.

Brouyère goes on to argue that "modelling a given "event" with great accuracy does not mean to have a good model. If the essential processes are not captured in that "event", it is likely that the model will rapidly fail when conditions change." This is absolutely correct and our paper never advocated extrapolating these simple models to alternative situations.

2) Well bore mixing and capturing

Brouyère argues that greater justification is needed for the neglecting of well-bore mixing other than the fact that this introduces further parameters into the calibration procedure.

For the South Farm test, tracer concentrations in the injection well indicated that 80% of the tracer had entered the aquifer after just 22 hours and 99.5% after 46 hours (see Atkinson et al. 2000). Given that the first arrival time was after just 32 hours it is likely that well-bore mixing was important. On this basis Atkinson et al. (2001) argued that the injection well could be better represented by the boundary condition

$$C_f + \frac{1}{P}\frac{\partial C_f}{\partial r_D} = \frac{1}{t_0}\exp\left(-\frac{t}{t_0}\right), \quad r_D = 1, \quad t > 0 \tag{1}$$

with $t_0 = 15.19$ hours.

HESSD

3, S1533–S1542, 2006

Note that the limit as $t_0 \rightarrow 0$ leads to equation (5) in hessd-2006-0092.

Figure 12 (http://www3.imperial.ac.uk/portal/pls/portallive/docs/1/7293590.PDF) shows a comparison of SFM models with instantaneous and exponential tracer injections. Indeed, it can be seen that consideration of well bore mixing does lead to significant perturbation of the tracer breakthrough curve (compare thick solid and thin dashed lines). The thin solid line shows an optimised SFM that takes into account well-bore mixing. Accounting for well-bore mixing leads to comparable model performance (compare thick and thin solid lines). However, markedly different optimal parameters are obtained depending on the assumed value of t0 (see legend in Figure 12).

The issue is that t_0 is unknown and the exponential decay could be a gross-simplification of the tracer injection function. This is the subject of a more recent paper currently under review with Water Resources Research. A detailed poster on this subject:

Mathias, S. A., Butler, A. P., Peach, D. W., Williams, A. T. (2005) Recovering tracer test input functions from fluid electrical conductivity logging in fractured porous rocks, AGU Fall Meeting, Poster H41F-0474, can be downloaded at http://www3.imperial.ac.uk/people/simon.mathias/research.

Our paper has clearly shown that it is not possible to delineate a unique optimal set of t_a, t_{cf} and P from tracer test breakthrough data such as these. The above results suggest that consideration of well-bore mixing leads to even further uncertainty, and hence identifiability issues. These have now been included as a short appendix in the final revised manuscript.

3) Analytical solutions and the true complexity of chalk

Brouyère points out that "one can have some doubts about the real applicability

HESSD

3, S1533–S1542, 2006

of analytical solutions, considering parallel fractures connecting injection and pumping wells ideally etc." However, the reality is that such models can be extremely good at replicating data from radially convergent tracer tests. Tracer tests are practically the only way of observing the in situ response of a subsurface transport system. Given that this paper has demonstrated it is virtually impossible to condition simple analytical solutions (e.g. the PFDM) to such data sets, what kind of data would be needed to appropriately condition a more complicated (although arguably more realistic) model?

Brouyère argues that such analytical solutions "rely on the strong assumption that mass transfer processes occurring between mobile and immobile water are related to molecular diffusion. Recently, there have been several works showing that a contrasted heterogeneity of hydraulic conductivity is likely to produce non-equilibrium solute transport with same "visible" consequences such as concentration attenuation and tailing." The issue is that matrix diffusion certainly occurs in chalk aquifers. Incorporating alternative mechanisms in addition to matrix diffusion leads to more parameters and more uncertainty.

At this stage it is useful to reemphasise, we are not saying the models we have used are the right ones. They are gross simplifications of reality. The point this paper demonstrates is that even these overly simplified analytical solutions, when assessed using these types of tracer test data, are over parameterised and cannot be uniquely conditioned.

4) Dealing with tracer mass recovery

Brouyère states he does "not agree too much with using tricks to constrain calculated and observed mass recovery in models." The method we used to estimate mass recovery is transparently described in section 4 of the paper. The argument is that we can not always be sure how much tracer will ultimately arrive at the abstraction well because it is not possible to observe the effluent concentrations indefinitely. How-

HESSD

3, S1533–S1542, 2006

ever, the quantity of tracer contained in the abstraction water during the monitoring period can be reasonably estimated. The methodology used in section 4 ensures that all models (including all model structures and parameter sets) deliver the same amount of tracer mass during the monitoring period.

Brouyère argues that "if the recovered factor is $<$ 100%, the model should be able to explain it as much as possible". For the South Farm test, a possible model could be that an alternative pathway diverted the 44% of unaccounted tracer mass away from the abstraction well. It is not possible to describe this pathway any further as the tracer test data only gives information about the pathway from the injection well to the abstraction well.

Brouyère suggests that we should consider fitting to the evolution of mass recovery rather than that of concentration data so as to better fit the breakthrough tail. A problem with this method is that the calculation of cumulative mass recovery involves an arbitrary interpolation between data points, which leads to serious loss of accuracy with the early time data.

He further highlights the fact that for the Horseheath test, "the calculated tail decreases more rapidly than the experimental one". However, this discrepancy is comparable with the noise signal. Furthermore, it is unlikely to have been fixed by changing the objective function as this was already biased towards the tail due to the high density of data points at that end.

5) Mechanical dispersion and diffusion in the matrix

Brouyère claims that he is unsurprised "that there is an indetermination between mechanical dispersion and matrix diffusion". Whilst the result in this paper is in some way common sense it is a useful demonstration.

B) Response to comments from Jishan Liu (28/09/2006)

HESSD

3, S1533–S1542, 2006

1) Liu claims that “conceptually, because of the high mass recovery fraction and negligible dispersion for the Horseheath test, SFM could be the best model to use; because of the low mass recovery fraction and high dispersion for the South Farm test, PFDM could be the best model to use. The authors may have to address the discrepancy between the conceptual models and their findings.”

It is not clear what is meant here. One of the main findings of the paper was that it is not possible to determine whether dispersion is negligible or not.

2) Liu argues that “More justifications are required for the conclusion: “Overall, this study emphasises the importance of adequate temporal sampling of breakthrough curve data prior to peak concentrations, to ensure adequate characterisation of mechanical dispersion processes, and continued monitoring afterwards, to ensure adequate characterisation of fracture spacing (where possible), when parameterising dual-porosity solute transport models”.”

It is still our opinion that the justification is perfectly adequate. The “importance of adequate temporal sampling of breakthrough curve data prior to peak concentrations, to ensure adequate characterisation of mechanical dispersion processes” is clearly emphasised in Figures 10 and 11. The importance of ”continued monitoring afterwards, to ensure adequate characterisation of fracture spacing (where possible), when parameterising dual-porosity solute transport models” is clearly emphasised by Figure 6 and Table 4.

C) Response to comments from Jasper Vrugt (29/09/2006)

1) Vrugt suggests our comment (Page 2440, Line 1 - 2) “Generally modellers have not addressed this question because they have not used a formal methodology

HESSD

3, S1533–S1542, 2006

to investigate parameter identifiability" is "incorrect, and misleading". This comment is indeed misleading when taken out of context in this way. In the context of radially convergent tracer tests (RCTT) it is pertinent and not misleading as demonstrated by the list of relevant references that accompany it. However, to avoid further confusion, this statement has been changed to "Generally modellers of RCTTs have..." in the final revised manuscript.

2) Vrugt suggests that our "identifiability analysis is extremely simple and inadequate". He believes that we "should consider stochastic optimization using appropriate likelihood functions to draw statistically correct conclusions". He claims that our "current method is very weak, and lacks any statistical basis".

Statistical likelihood functions assume a correct model structure and a known error distribution (even when using state updating) and therefore are not always useful (see Beven and Freer, 2001). We agree that our identifiability analysis is simple. In the context of the paper, we regard this as an advantage. The benefits of a simple model screening analysis, which has explicit and clear limitations, are well-recognised (Beck, 1997, Hornberger and Spear 1980; Saltelli et al. 2000). The SODA method and other similarly complex methods which Vrugt refers to also have limitations - but these are more difficult to make clear or explicit (see response 3). However, it is evident that the 'statistical basis' of Vrugt's method is also flawed, as it assumes a model structure that is not "correct" (Beven and Freer, 2001).

As a model structure is never "correct" a rigorous statistical method for assessing parameter uncertainty is not possible. Therefore a subjective statistical approximation must be made. Consequently, we have opted for the most transparent approximation possible. That is, to construct multivariate plots of parameter values against objective function values and make binary decisions about parameter identifiability (identifiable or non-identifiable) on the basis of the visual appearance of these plots. Previously

HESSD
3, S1533–S1542, 2006

identifiability has been quantified in a very similar way but using statistical descriptors of the calculated objective function values (e.g. Wagner et al., 2001; Beven and Freer, 2001; Sincock et al., 2003; Smith and Wheater, 2004). However, due to the subjective choice of objective function (and other decisions that have to be made), this process is arguably as arbitrary as making a personal decision based on a visual inspection, as we have done.

3) Vrugt requests that we consider his paper (Vrugt et al., 2005, Geophys. Res. Lett., 32, L18408). In this paper Vrugt et al. reanalyse a weak-dipole tracer test carried out in the fractured volcanic tuff at Yucca Mountain using a Simultaneous Optimization and Data Assimilation method (SODA). SODA uses an ensemble Karman filter to recursively update model states while estimating time-invariant values for the model parameters using the Shuffled Complex Evolution Metropolis stochastic-ensemble optimization approach (SCEM-UA). This method is capable of estimating more reliable confidence intervals than those obtained from methods that do not consider data assimilation. The reason is that the recursive updating of model states allows the discarding of noise which can help towards preventing a biasing of parameter estimation to accommodate measurement error.

However, if the wrong model and/or parameter set are used (which is almost always the case), this 'measurement error' includes other sources such as conceptual errors, process uncertainty and process misrepresentation. Furthermore (as discussed above in response to comment 2) SODA does not have a rigorous statistical basis. Several arbitrary assumptions have to be made, for example: the Kalman Gain function (see Vrugt's equation 7) and the Bayesian Density Criterion (see Vrugt's equation 8). Alternative assumptions (which may be as justifiable) will lead to alternative confidence limits.

While we concede that the approach of Vrugt et al. (2005) is much more advanced

HESSD
3, S1533–S1542, 2006

than our own, their work does not negate or make obsolete the usefulness of our paper. Although they present SODA in the context of a weak-dipole tracer test (which is very similar to a radially convergent tracer test) they do not concern themselves with the identification of parameters that describe the transport of conservative solutes (which is the main focus of our paper).

Specifically Vrugt et al. (2005) look at Bromide and Lithium breakthrough curves (BTC). Both the Bromide and Lithium were originally mixed together before being injected into the aquifer. A residence time distribution (RTD) is estimated from the conservative Bromide BTC a priori. This is then fixed and assumed also to apply to the Lithium BTC leaving two Freundlich Isotherm parameters to be defined. It is only these two parameters that are studied in the context of parameter identifiability. With the exception of the conceptual basis contained in Robinson and Viswanathan (2003), Vrugt et al. (2005) give no information as to how the RTD is obtained.

References not contained in the HESSD paper

Robinson, B. A., and H. S. Viswanathan (2003), Application of the theory of micromixing to groundwater reactive transport models, Water Resour. Res., 39(11), 1313, doi:10.1029/2003WR002368.

Beven, K. J., (2004), Does an interagency meeting in Washington imply uncertainty? Hydrol. Process. 18, 1747-1750.

Beven, K. J. and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11-29

Vrugt, J. A., B. A. Robinson, and V. V. Vesselinov (2005), Improved inverse modeling for flow and transport in subsurface media: Combined parameter and state estimation, Geophys. Res. Lett., 32, L18408, doi:10.1029/2005GL023940.

HESSD

3, S1533–S1542, 2006

Beck, M.B. 1987. Uncertainty in water quality models. Water Resources Research 23(8), 1393.

Hornberger, G.M. and Spear, R.C. 1980. Eutrophication in Peel Inlet, 1. Problem-defining behaviour and a mathematical model for the phosphorus scenario. Water Research 14, 29-42.

Saltelli, A., Chan, K., Scott, E.M. (2000) Sensitivity Analysis, Probability and Statistics Series. New York: John Wiley & Sons.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 2437, 2006.