Hydrol. Earth Syst. Sci. Discuss., 4, S252–S261, 2007
www.hydrol-earth-syst-sci-discuss.net/4/S252/2007/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Sensitivity analysis and parameter estimation for the distributed modeling of infiltration excess overland flow" *by* W. Castaings et al.

W. Castaings et al.

Received and published: 7 May 2007

The authors would like to thank the referees for their interest in the work carried out and for the valuable comments which should improve the quality of the paper.

General comments

Although some efforts are required for its practical implementation (as emphasized by referee ♯1), the adjoint state method, and more particularly its implementation using the reverse mode of automatic differentiation, provides derivatives accurate to machine precision for a computational cost independent from the dimension of the input space. For this reason, the approach is particularly suited for the analysis and control of spatially distributed systems like meteorological and oceanic models.

The authors agree with referee $\sharp 2$ stating that more references related to the use of deterministic sensitivity analysis in hydrological modeling should be included in the paper. The same argument probably applies to the use of local methods for parameter estimation. Parameter estimation and sensitivity analysis are intimately related. As mentioned in page 379 and also emphasized by Kavetski et al. (2006), "the classical and very efficient gradient-based parameter estimation methods have been abandoned". Therefore, the same happened for the use of sensitivity analysis based on derivative information. Deterministic sensitivity analysis or gradient-based optimization techniques have been used in the past and the contributions of McCuen (1973a,b) and Gupta and Sorooshian (1985), in trying to obtain analytic derivatives rather than classical finite difference approximation are directly in line with the current research endeavor.

McCuen (1973a) states that "since most complex simulation models are expressed in finite difference from, it would be impossible to derive a closed form of the sensitivity function". The model is splitted into components and the author argues that the "concept of component analysis eliminated one of the primary mathematical difficulties affecting the use of sensitivity analysis as a design tool for complex systems". Similarly, Gupta and Sorooshian (1985) derive analytically all the "subcomponents operations" carried out between the occurrence of switches introduced by thresholds in the model formulation (very common in *bucket models*).

This is exactly the strategy adopted by algorithmic differentiation developed in parallel in other scientific fields. While the forward mode of automatic differentiation is simply a more convenient and efficient way to apply the previously mentioned approach for complex models, by making the computational cost independent from the dimension of the input space the reverse mode (discrete equivalent of the adjoint state

HESSD

4, S252–S261, 2007

method) represents a significant improvement.

Sensitivity Analysis

When compared to other local sensitivity analysis techniques, the adjoint method is the only one able to cope with high-dimensional systems for a reasonable computational cost. Of course the approach is only local and suffers from the drawbacks of all local methods for non-linear systems. The referee $\sharp 2$ is absolutely right in pointing that there is no reason why the same conclusions can be achieved regarding parameter importance in different locations of the parameter space. Although this can and should be checked, the fact that the obtained ranking with a local method is stable over the parameter space cannot be transposed to other applications.

For this particular model, page 374 the authors state: "Additional experiments, which are not reported here, show that the wetter the soil at the beginning of the event, the faster the decay of the infiltration rate to the hydraulic conductivity and therefore the greater the relative influence of K compared to the initial soil moisture θ." This sentence refers to a transect in the parameter space ($\theta \in [0, 1]$) for which the stability of the ranking was investigated. Other transects where explored and not reported in the paper. A more systematic approach would imply a comparison with global sensitivity analysis techniques like the method of Sobol' (Sobol', 1993; Homma and Saltelli, 1996; Saltelli, 2002) or factors screening with the Morris method (Morris, 1991) which is probably more suited for computer intensive models like the one analysed in this study.

However, the type of rainfall event also influences the outcomes from the sensitivity analysis. This was also investigated using very simple rainfall scenarii (not reported in the paper) but would require a systematic approach to draw more universal conclusions. Therefore, in order to obtain a robust ranking for this particular model

HESSD

4, S252–S261, 2007

Interactive Comment

EGU

a rigorous exploration of the model behavior would require the coupling between a stochastic rainfall generator (or a large database of rainfall events) with a global sensitivity analysis method for different watershed configurations.

As underlined by McCuen (1973b), sensitivity analysis should play a key role in model formulation, model calibration and model verification. If global sensitivity analysis techniques are really suited for ranking factors importance over the multi-dimensional parameter space, the role of sensitivity analysis in hydrological modeling should not be reduced to this aim. For instance, if one is interested in the behavior of the system for regions in the parameter space leading to satisfactory fit to observations, local sensitivity analysis often prove very informative. The sign and magnitude for the sensitivity of the response to a given input factor greatly facilitates the analysis and interpretation in mechanistic terms. In the ensemble prediction framework, beyond an exploration of the input space using probabilistic methods, a detailed analysis of the behavior of the model for the ensemble mean can be also a very fruitful exercise.

Parameter estimation

Some of the comments of referee $\sharp 3$ and $\sharp 1$ were related to the efficiency of the adopted technique in comparison to other optimization methods but also to the difficult problem of equifinality. The authors do not claim that the presented approach will overtake all available local and global parameter estimation techniques in all cases. A very efficient local method is presented and some concerns are expressed on the abusive and inefficient use of global non-smooth optimization methods in catchment scale hydrology.

For the calibration of model parameters, as Ebel and Loague (2006) stress in a recent invited commentary, "all the issues associated with ill-posed problems in hydrologic-response simulation have been lumped together and termed equifinality".

HESSD

4, S252–S261, 2007

Interactive Comment

While referee $\sharp 1$ states that "equifinality is one of the major reasons why the adjoint methodology has not found applicability in hydrological modelling despite its success in order fields", we argue that ill-posed inverse problems are encountered in many scientific disciplines and do not prevent the use of variational methods. Would referee $\sharp 1$ contest the indisputable fact that equifinality does occur in the estimation of a distributed hydraulic conductivity in groundwater hydrology, in the estimation of a distributed initial state for the atmosphere in meteorology ? For the previously mentioned fields, given the dimension of the control space and the computational cost of the models, very often a regularized inverse problem (Tikonov regularization using *a priori* information) is solved using variational methods rather than a global search for all behavioral sets. Independently from the uncertainty associated to observations and from the representativeness of the model w.r.t to the complex hydrological reality, multi-modality due to a limited structural identifiability is very often the principal reason for equifinality.

When the response surface of the model contains a single region of attraction, gradient based techniques are indisputably the most efficient optimization strategy. As underlined by Kavetski et al. (2006) "convergence safeguards such as line-searches and trust-regions in modern gradient based algorithms improve significantly the reliability of the estimates". Uni-modality or at least limited multi-modality can be achieved by providing enough information content from the observations (quantity, quality and heterogeneity) and a parsimonious parametrization; even if this is done using empirical and somehow arbitrary dimension reduction for distributed models. For the experiments with parametrization of reduced dimensionality, the authors are not convinced that a comparison with a widely accepted but very expensive techniques like the Shuffle Complex Evolution from Duan et al. (1992) would strengthen the argument of the paper. Figure 10 shows the number of direct and adjoint models runs (including the model evaluations for the line search) required for the retrieval of 5 parameters ($\sim$ 50 iterations) and this is already an argument when compared to the

HESSD

4, S252–S261, 2007

EGU

number of model runs characterizing evolutionary algorithms. Kavetski et al. (2006) emphasize that "5 minutes of Newton computing often replaces 24 hours of SCE search; moreover Newton methods yield useful additional information."

If limited muti-modality can be explored with gradient-based optimization (global smooth strategy also suggested by Kavetski et al. (2006)), the calibration of distributed models should be tackled in a different paradigm instead of simply transferring the experience gained on global models using scalar multipliers adjusting an *a priori* spatial distribution for the parameters with global non-smooth methods. The authors argue that more efficient techniques should be used given the mathematical foundations of most distributed models. Moreover, given that more sizable contrains are usually available for the parameters of a distributed model, the calibration paradigm should be perceived as an improvement of the *a priori* estimates rather than a blind search over the multi-dimensional parameter space. The paper open prospects for a dimension reduction procedure which can be seen as a regularization strategy and should be combined classical regularization techniques (Tonkin and Doherty, 2005; Doherty and Skahill, 2006).

Applicability to other hydrological models

Referee $\sharp 3$ states that the generalization of the variational methods to other hydrological models is not very clear. Some additional information is provided below and should complement the final version of the paper. Although models based on partial differential equations form the classical framework for the application of variational methods, most of the time their algorithmic representation does not significantly differ from the one characterising models based on less sophisticated mathematical formulations. The advent of algorithmic differentiation open new trends for the application to other type of models. Representative examples are provided by Lauvernet (2005) or Wu (2005).

HESSD

4, S252–S261, 2007

Interactive Comment

If the application to other type of models raise new issues, the main limitation of the approach is related to the differentiability, or more precisely piecewise differentiability. Even if it is sometimes overlooked, this limitation also applies for models based on partial differential equations. Since most models are only piecewise differentiable the gradient does not exist but subgradients can be exhibited. As emphasized in page 380 of the paper, the direct implication is that the derivatives are valid only in the control flow trough the conditional statements defining the model behavior. The flow is defined by the nominal values for the input factors and the derivatives hold only for a neighborhood of this point in the parameter space. While the accuracy of the calculated derivatives is a critical element of the efficiency of descent methods, for sensitivity analysis, threshold processes and strong non-linearity inevitably tend to reduce the vicinity around which the sensitivity outcomes are rigorously valid. Moreover, even for a well defined optimal parameter set, this also constitutes a serious limitation for the propagation of uncertainty using derivatives (local uncertainty analysis).

In the past simplified versions of the direct model were used for the derivation of the adjoint model in meteorology. Although the indisputable benefits from the inclusion of the full physics of the original model for sensitivity analysis and parameter estimation (or data assimilation) were demonstrated (Li et al., 2000), the way these threshold processes should be treated for the development of an adjoint is still subject of controversy. While some authors stress the need for simplified regularized physics (Janiskova et al., 1999) or propose generalized adjoint formulations (Xu, 1996; Mu and Wang, 2003), others like Zhang et al. (2000) state that the very efficient quasi-newton algorithms may still work with the subgradient calculated from the discrete direct model. They also underline the fact that "local smoothing introduced to remove discontinuities may lead to more problems than solutions due to the insertion of artificial stationary points."

HESSD

4, S252–S261, 2007

Interactive Comment

In order to complete this picture, it is important to mention that although the approach discretize and derive generally provides the most reliable and accurate derivatives, for some cases it might lead to instability in the solution of the adjoint model (Sirkes and Tziperman, 1997). Therefore, as commented by Sandu et al. (2003), "there is no general rule to decide which approach (discrete vs. continuous) should be used to implement the adjoint model. The performance of the adjoint code is problem dependent and a selection can be made only after an extensive analysis and testing for the particular problem to be solved have been performed."

All the previously mentioned difficulties are mainly the reason why a very simple and common model structure, on a small watershed using synthetic observations was adopted for this prospective study. Investigations are in progress concerning the application of this technique to other model structures (based on TOPMODEL) and will be reported in due course.

References

Doherty, J. and Skahill, B. E.: An advanced regularization methodology for use in watershed model calibration, Journal of Hydrology, 327, 564-577, 2006.

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and Efficient Global Optimisation for Conceptual Rainfall-Runoff Models, Water Resources Research, 28, 1015-1031, 1992.

Ebel, B. A. and Loague, K.: Physics-based hydrologic-response simulation: Seeing through the fog of equifinality, Hydrological Processes, 20, 2887-2900, doi:DOI:10.1002/hyp.6388, invited Commentary, 2006.

Gupta, V. and Sorooshian, H.: The automatic calibration of conceptual catchment models using derivative based optimization algorithms, Water Ressources Research,

HESSD

4, S252–S261, 2007

Interactive Comment

21, 473-485, 1985.
Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering and System Safety, 52, 1-17, 1996.
Janiskova, M., Thpaut, J.-N., and Geleyn, J.-F.: Simplified and Regular Physical Parameterizations for Incremental Four-Dimensional Variational Assimilation, Monthly Weather Review, 127(1), 26-45, 1999.
Kavetski, D., Kuczera, G., and Franks, S. W.: Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis, Journal of Hydrology, 320, 187-201, 2006.
Lauvernet, C.: Assimilation variationnelle des observations de télédetection dans les modèles de fonctionnement de la végétation : utilisation du modèle adjoint et prise en compte des contraintes spatiales., Ph.D. thesis, Université Joseph Fourier (Grenoble I), 2005.
Li, Z., Navon, I. M., and Zhu, Y.: Performance of 4D-Var with Different Strategies for the Use of Adjoint Physics with the FSU Global Spectral Model, Monthly Weather Review, 128(3), 668-688, 2000.
McCuen, R.: Component sensitivity: a tool for the analysis of complex water resources systems, Water Resources Research, 9 (1), 243-247, 1973a.
McCuen, R.: The role of sensitivity analysis in hydrologic modelling., Journal of Hydrology, 18, 37-53, 1973b.
Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33, 161-174, 1991.
Mu, M. and Wang, J.: A Method for Adjoint Variational Data Assimilation with Physical OnOff Processes, Journal of the Atmospheric Sciences, 60(16), 2010-2018, 2003.
Saltelli, A.: Making best use of model valuations to compute sensitivity indices, Computer Physics Communications, 145, 280-297, 2002.
Sandu, A., Daescu, D., and Carmichael, G.: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I - Theory and software tools, Atmospheric Environment, 37, 5083- 5096, 2003.

Sirkes, Z. and Tziperman, E.: Finite difference of adjoint or adjoint of finite difference?, Monthly weather review, 49, 5-40, 1997.

Sobol', I.: Sensitivity analysis for non-linear mathematical models, Mathematical Modelling and Computational Experiment, 1, 407-414, 1993.

Tonkin, M. and Doherty, J.: A hybrid regularized inversion methodology for highly parametrized environmental models, Water Ressources Research, 41, 2005.

Wu, L.: Variational methods applied to plant functional-structural dynamics: Parameter Identification, Control and Data Assimilation, Ph.D. thesis, Université Joseph Fourier (Grenoble I), 2005.

Xu, Q.: Generalized Adjoint for Physical Processes with Parameterized Discontinuities. Part I: Basic Issues and Heuristic Examples, Journal of the Atmospheric Sciences, 53(8), 1123-1142, 1996.

Zhang, S., Zou, X. amd Ahlquist, J., Navon, I., and Sela, J.: Use of Differentiable and Non-Differentiable Optimization Algorithms for Variational Data Assimilation with Discontinuous Cost Functions, Monthly Weather Review, 128, 4031-4044, 2000.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 4, 363, 2007.