European Geosciences Union (EGU) - General Assembly
Vienna, Austria, 19 - 24 April 2009

Session HS2.1: *Hydroinformatics: computational intelligence and technological developments in water science applications*

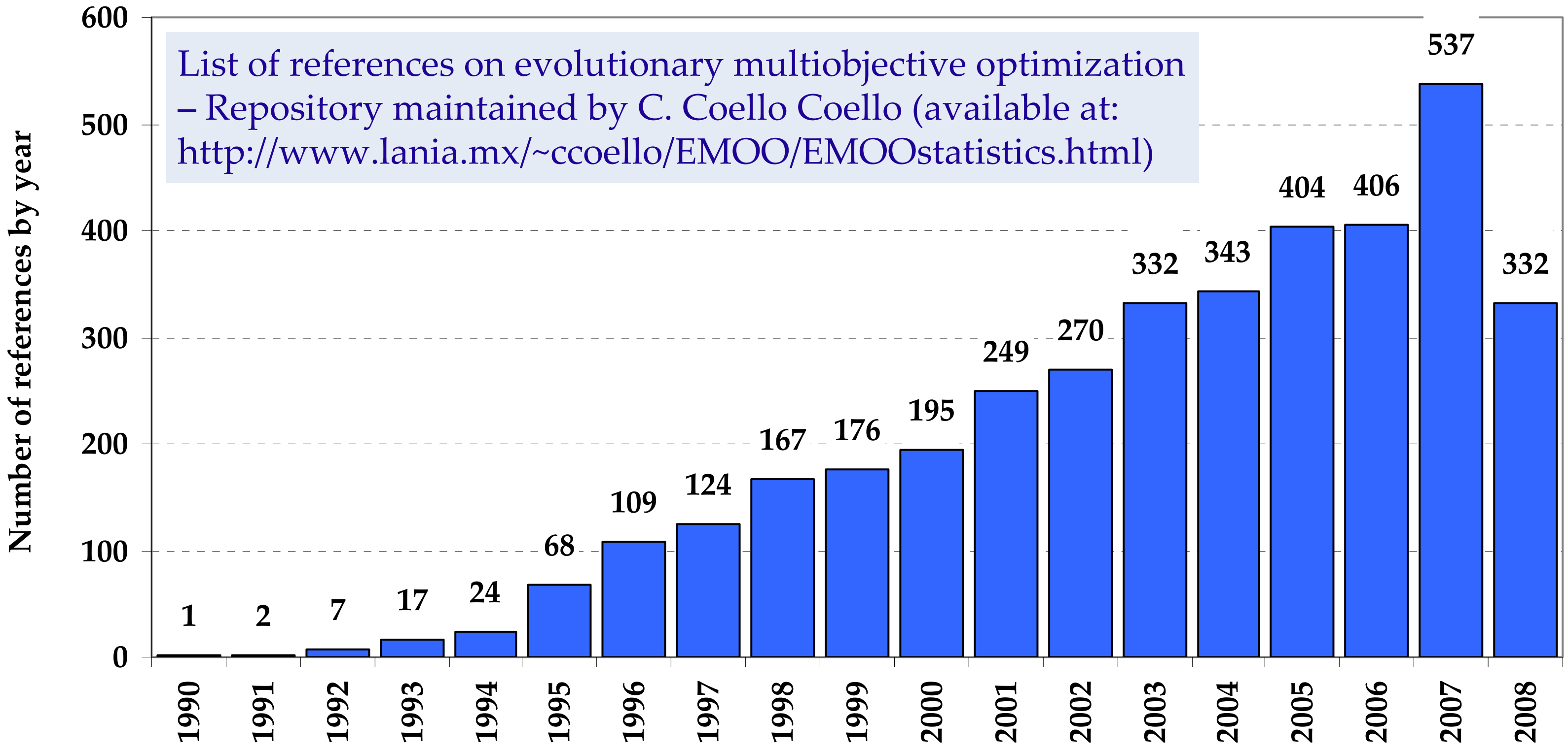
On the practical use of multiobjective optimization in hydrological model calibration

Andreas Efstratiadis and **Demetris Koutsoyiannis**

Department of Water Resources and Environmental Engineering
National Technical University of Athens, Greece

Multiobjective optimization in computer and engineering science: The progress so far

- Expansion of research progress from the early 1990’s, providing already three “generations” of evolutionary multiobjective algorithms.
- Many domains of application, including water resources technology and hydroinformatics.



Important steps in the history of multiobjective hydrological calibration

- **1991**: Early hybrid attempts (Harlin).
- **1998**: Recognition of the value of multiple criteria information in calibration (Gupta *et al*.; Kuczera & Mroczkowski) and first Pareto-based approach, within the presentation of the Multiobjective Complex Evolution method (Yapo *et al*.).
- **2000**: First calibration study involving more than two objectives to optimize (Madsen) and first conjunctive calibration based on criteria accounting for discharge and groundwater level information (Seibert).
- **2002**: The concept of "soft" data (Seibert & McDonnell).
- **2003**: Development of the Multiobjective Shuffled Complex Evolution Metropolis method (Vrugt *et al*.).
- **2005**: First application of the NSGA-II algorithm in hydrological calibration (Khu & Madsen).
- **2006**: First comparative assessment study, involving three well-recognized techniques (NSGA-II, SPEA-II, MOSCEM) (Tang *et al*.).
- **2007**: Parallel implementations of multiobjective algorithms in time-consuming calibration problems (Confesor & Whittaker; Tang *et al*.).

Model calibration: The (diachronic) dilemma between consistency and optimality

- A calibration based on the concept of **consistency**:
  - aims to represent as faithfully as possible the whole aspects of the real system (not only the observed responses);
  - recognises the major and inherent role of uncertainty;
  - requires some manual guidance (difficult to fully automate);
  - reveals the importance of hydrological experience within all modelling stages (conceptualization, schematization, parameterization, calibration).
- A calibration based on the concept of **optimality**:
  - aims to ensure the most accurate fitting on observed data, which is trivial to automate (formulation of a global optimization problem);
  - usually assumes a single numerical criterion as an overall evaluator of the model performance;
  - is proved too weak against errors and uncertainties;
  - is possible to degenerate to a black box procedure, providing over-fitted schemes with limited predictive capacity.

Model calibration: Back to the fundamentals

- General formulation of model fitting to observed data:

$$\text{minimize } \mathbf{e}(\boldsymbol{\theta}) = \{|e_1(\boldsymbol{\theta})|, |e_2(\boldsymbol{\theta})|, \ldots, |e_M(\boldsymbol{\theta})|\}$$

  where $e_i(\boldsymbol{\theta})$ model residuals (departures of observed responses from the computed ones), M number of observations, and $\boldsymbol{\theta}$ vector of parameters.

- The above formulation:
  - is inherently multiobjective (since models are imperfect simulators of highly complex systems);
  - is impractical to interpret (the number of residuals is too large, thus the Pareto front tends to cover the entire M-dimensional objective space);
  - is unnecessarily not parsimonious in multiobjective terms (the residuals are highly correlated).

- "Reduced" formulation of the calibration problem:

$$\text{maximize } \mathbf{g}[\mathbf{e}(\boldsymbol{\theta})] = \{g_1[\mathbf{e}(\boldsymbol{\theta})], g_2[\mathbf{e}(\boldsymbol{\theta})], \ldots, g_m[\mathbf{e}(\boldsymbol{\theta})]\}$$

  where $g_j(\boldsymbol{\theta})$ scalar fitting criteria that account for representative aspects of the model performance (should be approximately uncorrelated), and m the reduced dimension, with $m << M$.

Pareto optimality: An alternative to equifinality?

- Both concepts reject the concept of a unique optimal parameter set, thus dividing the feasible parameter space into two sub-areas, containing the acceptable (called Pareto-optimal and behavioral, respectively) and non-acceptable solutions (the sub-areas are *not identical*).
- The **equifinality** hypothesis:
  - is more general, since it is also applicable to model structures;
  - requires a *scalar* likelihood function (LF) and an *arbitrary* cut-off threshold to distinguish between behavioral and non-behavioral solutions;
  - is widely used within Bayesian inference methods, which evaluate model uncertainty around the LF, employing Monte-Carlo approaches.
- The **Pareto optimality** concept :
  - requires at least two fitting criteria to make sense;
  - is based on a strict mathematical notion, i.e. the principle of dominance, to distinguish between optimal and non-optimal solutions;
  - allows to effectively handle non-commensurable criteria;
  - is computationally much simpler, if compared to Bayesian methods;
  - is more practical and easier to understand.

Why multiobjective calibration?
(a) A view based on the principle of parsimony

- Parsimony is a key principle in mathematical and statistical modelling, where model parameters are estimated through data-fitting, which favours models having the simplest possible structure.
- In lumped modelling, it is accepted that only 5-6 parameters can be identified from external variables (e.g. runoff), while a more detailed structure, in the absence of supplementary control data, may result to poorly identified parameters; therefore **non-parsimony is a major source of uncertainty**.
- Given that semi- and fully-distributed modelling involves a large number of free parameters to identify, a multiobjective approach is essential to include more information in calibration.
- Models should allow for more **flexible parameterizations**, adapted to the available data.

Traditional approach

1. A priori specified parameterization
2. Control data → objective function
3. Calibration

Splitting schematization and parameterization

Alternative approach

1. Control data → multiple objectives
2. Parameterization adapted to data
3. Multiobjective calibration

Why multiobjective calibration?
(b) Controlling multiple system components

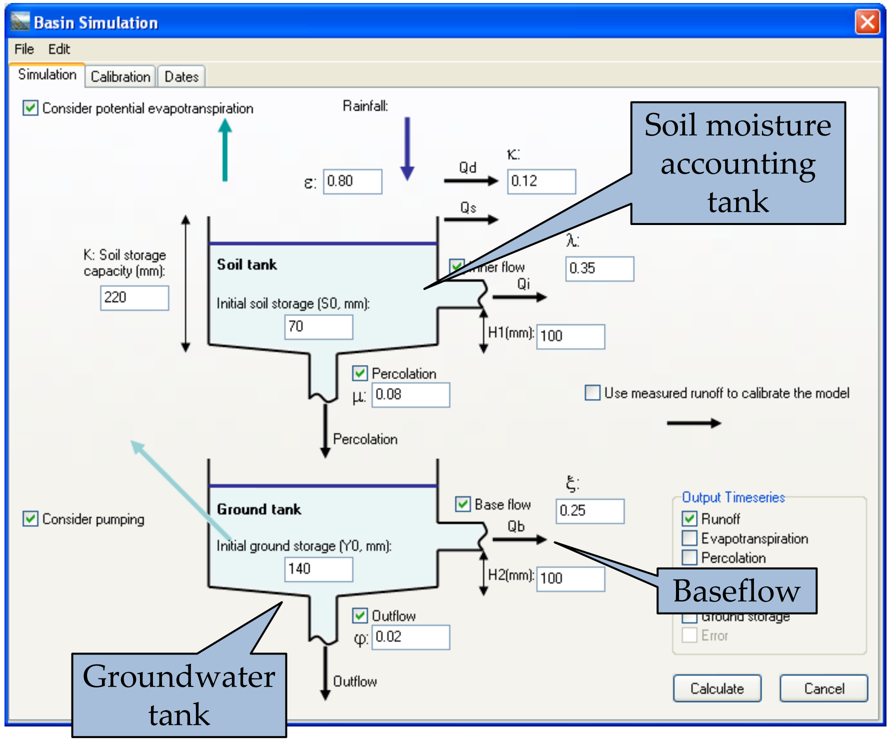
- Typical multiobjective formulation of calibration problems, based on "hard" data, i.e. systematic observations (terminology taken by Madsen, 2003):
  - multi-variable data (fitting criteria for different types of fluxes);
  - multi-site data (fitting criteria for a flux measured at different sites);
  - multi-response models (different aspects of a single flux).
- In complex models, the number of observations is usually incompatible with the principle of parsimony in parameterization; this makes it necessary to also take advantage of "soft" types of information, including:
  - sparse and non-systematic measurements;
  - average water balance statistics;
  - long-term fluctuation of internal fluxes (e.g. storage variables);
  - any qualitative information about the system behaviour.
- An effective combination of "hard" and "soft" data within a multiobjective framework allows for augmenting the information contained in calibration and taking advantage of the hydrological experience → **an enhanced view of manual calibration practices**.

Why multiobjective calibration? (c) Optimization of conflicting criteria—made easy

- Should fitting criteria be conflicting?
  - In theory, the augmentation of information embedded in calibration should systematically improve the quality of model fitting against all controlled aspects of the real system.
  - In practice, due to errors and uncertainties in both model structure and data, it is not always easy to recognize a priori whether two criteria are conflicting or not; moreover, their behaviours may differentiate across the search space.
- What is the practical value of Pareto-based approaches?
  - **Objectivity**, since the user avoids to employ arbitrary aggregating approaches, thus risking to hide significant competitions among criteria.
  - **Efficiency**, since state-of-the-art algorithms provide representative and well-dispersed solutions in a single run and with reasonable computational effort (much less if compared to step-by-step approximations of the Pareto front).
  - **Feasibility**, since the searching procedure is not easy to be trapped due to geometrical peculiarities of the Pareto front (e.g. non-convexities).
  - **Comprehensibility**, since the irregularities in the shape of the Pareto front may help to explain (and possibly remedy) model weaknesses, while the bounds of the Pareto set may be associated with parameter uncertainty ranges.

Example A: Lumped conceptual modeling of the Boeoticos Kephisos river basin

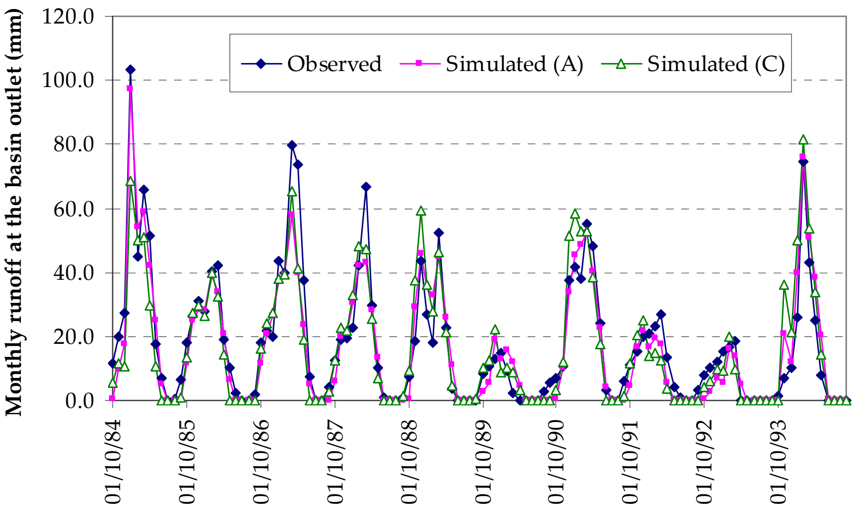
- Input data for monthly hydrological simulation (1984-1994):
  - Precipitation, potential evapotranspiration, pumping, river abstractions.
  - Observed runoff at the basin outlet.
  - Monthly average and standard deviation of baseflow series (estimated by aggregating the hydrographs of the six major springs).
- Hydrological characteristics:
  - Significant contribution of baseflow, with small variability.
  - Mediterranean climate, substantial losses due to evapotranspiration.
- Simulation via the **Zygos** model:
  - Soil and aquifer processes represented by two tanks.
  - Runoff made up by four components.
  - Up to 8 parameters to estimate.

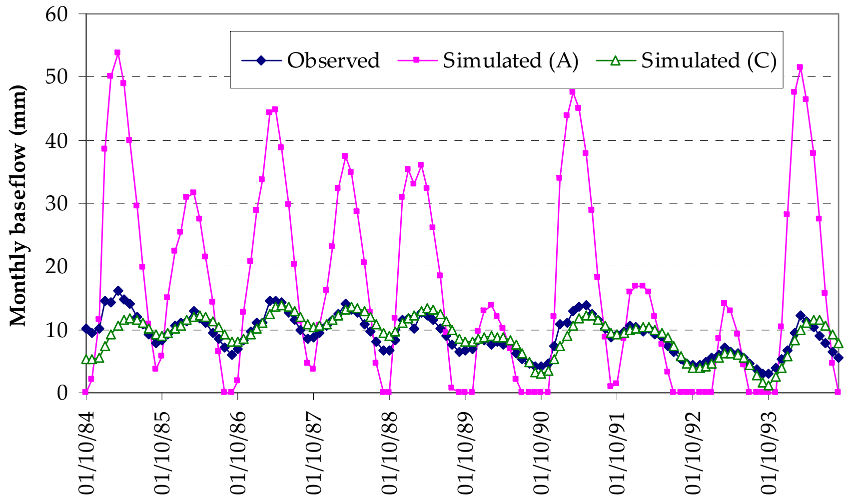


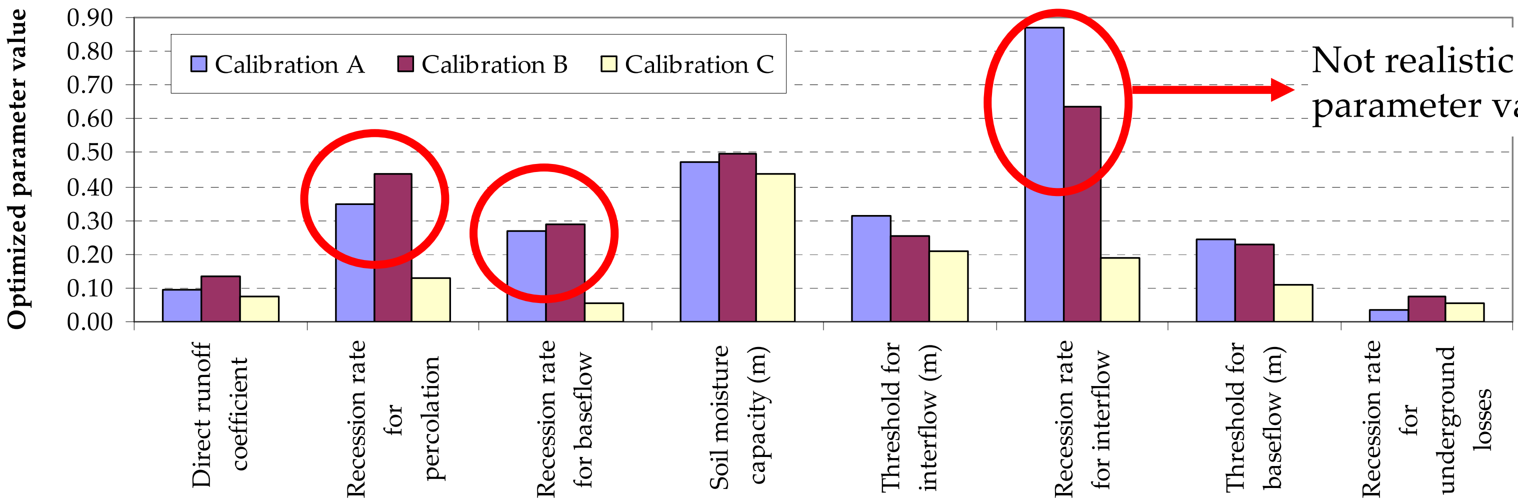
Example A: Multiobjective calibration attempts with augmenting information

- **Calibration A**: Single criterion, efficiency of total basin runoff (“hard” data);
  - Very satisfactory reproduction of runoff; efficiency = 87.8%.
  - Unrealistic fluctuation of spring runoff; ~2 times greater mean value, 4 times greater standard deviation.
- **Calibration B**: A + Penalty for departure from mean and standard deviation of monthly baseflow series (statistics provided by “hard” data)
  - Decrease of runoff efficiency from 87.8% to 78.4%.
  - Faithful fluctuation of baseflow, very close to the “observed” one.
  - Unrealistic overall water balance (mean annual evapotranspiration ≈ 1/3 of mean annual rainfall, which is unexpectedly low).
- **Calibration C**: B + Empirical criterion to maximize actual evapotranspiration (combination of “hard” and “soft” information)
  - Satisfactory high efficiency (80.7%) and faithful fluctuation of baseflow.
  - Realistic increase of mean annual evapotranspiration contribution up to half of the mean annual precipitation.
  - Consistent parameter values.

Example A: Graphical comparison of multiobjective calibration results

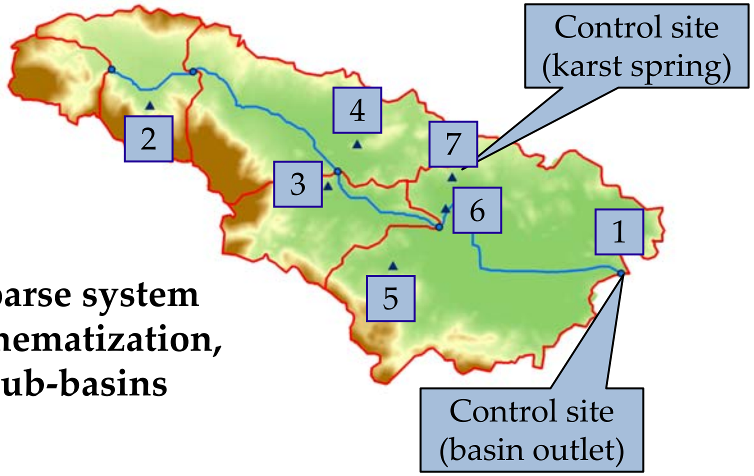




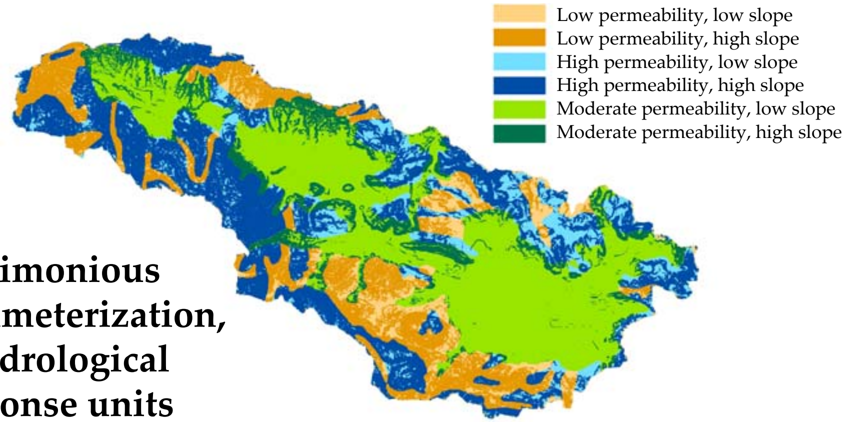


Not realistic parameter values

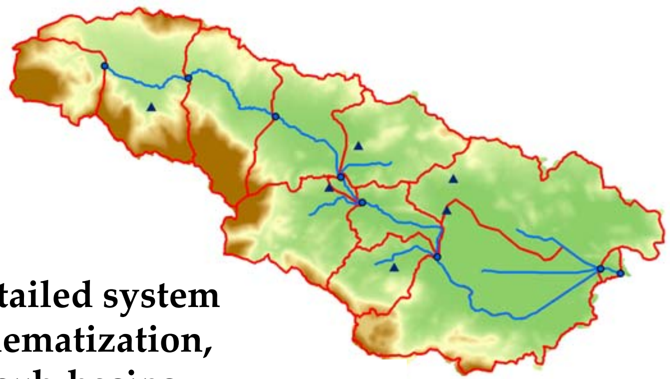
Example B: Semi-distributed modeling of Boeoticos Kephisos basin through the Hydrogeios model (*)



Coarse system schematization, 5 sub-basins



Parsimonious parameterization, 6 hydrological response units



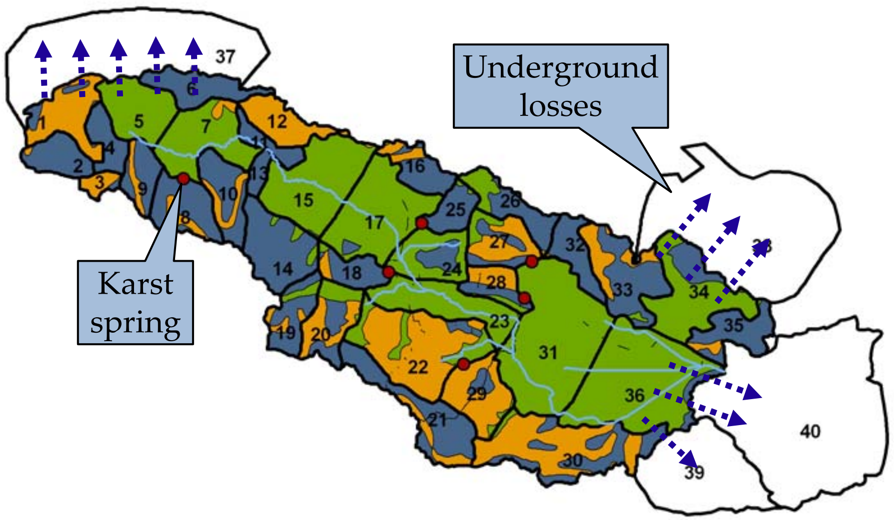
Detailed system schematization, 13 sub-basins

(*) Model and case study description (coarse schematization): Efstratiadis *et al.* (2008)

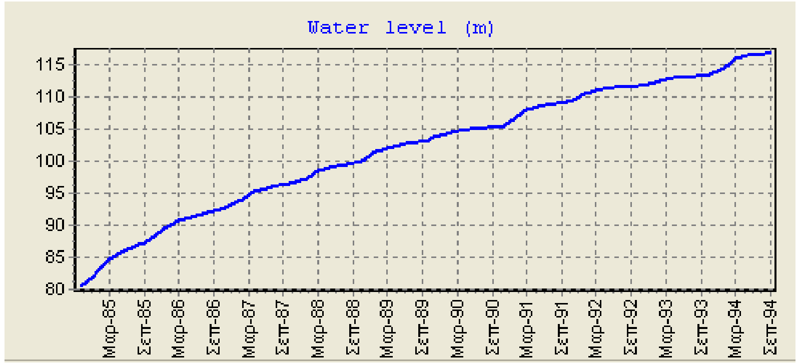
- Observed variables = discharge series at the basin outlet and downstream of six major karst springs (7 control responses).
- Parameterization through 6 Hydrological Response Units (HRUs; product of three types of permeability and two types of terrain slope).
- The processes of each HRU are represented by a conceptual model of 6 parameters.
- The total number of parameters remains 36, independently of the number of sub-basins.

Example B: Groundwater modelling issues

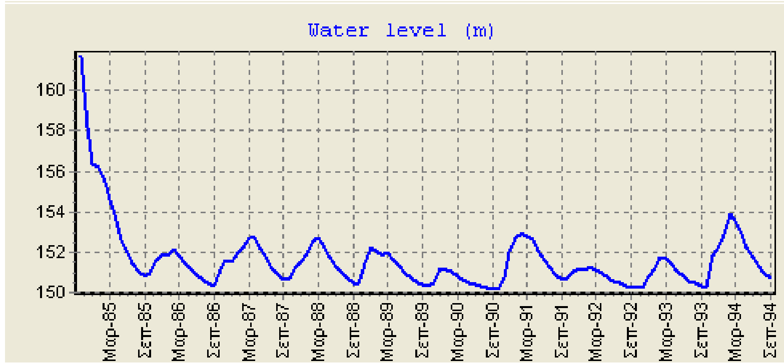
- Groundwater modelling through a multi-cell approach
  - Conceptualization: Darcian representation of flow field.
  - Schematization: 36 cells (conceptual tanks stressed by percolation, infiltration and pumping), 4 dummy cells representing underground losses, 6 dummy cells representing springs; delineation based on topography and permeability.
  - Parameterization: 3 categories of permeability and porosity, particular permeability values for cells representing springs and underground losses.
- Criteria used in conjunctive calibration:
  - Efficiency of observed hydrographs (“hard data”, sufficient for reproducing the water balance, not sufficient for representing of the entire groundwater regime).
  - Additional criteria for reproducing spring flow intermittency (**easily observable information** of major interest in water management).
  - Penalty functions to prohibit unrealistic water level “trends”, indicating systematic evacuation or filling of tanks (“soft” data ensuring **reasonable fluctuation** of the non-observable groundwater variables).



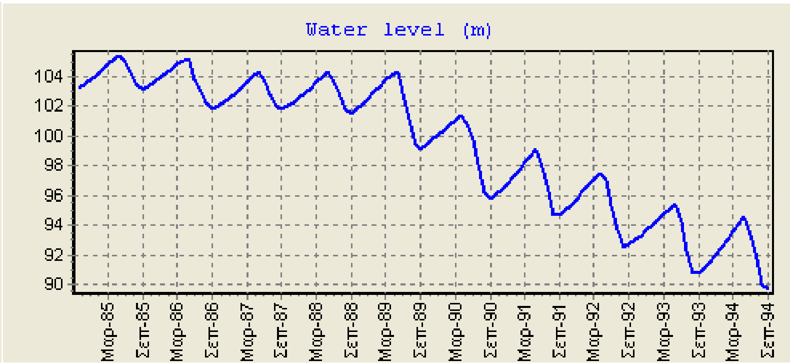
Example B: Why use “trend penalties” for the simulated groundwater level series?



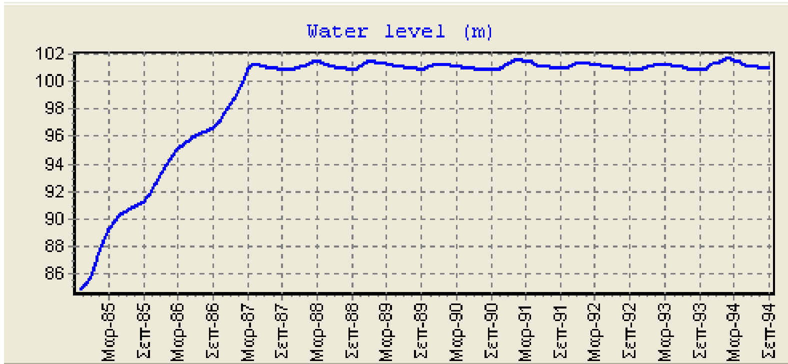
Systematic increase of cell level, without seasonal fluctuations (almost perfect linear trend).



Sharp decrease of cell level at the beginning of simulation, next followed by reasonable seasonal fluctuations.



Systematic decrease of cell level, although seasonal fluctuations are represented.

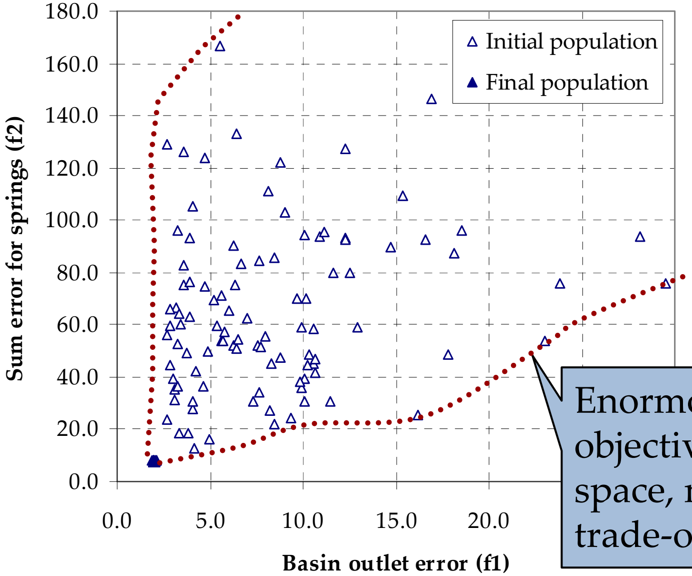


Systematic increase of cell level during the first years, next stabilized.

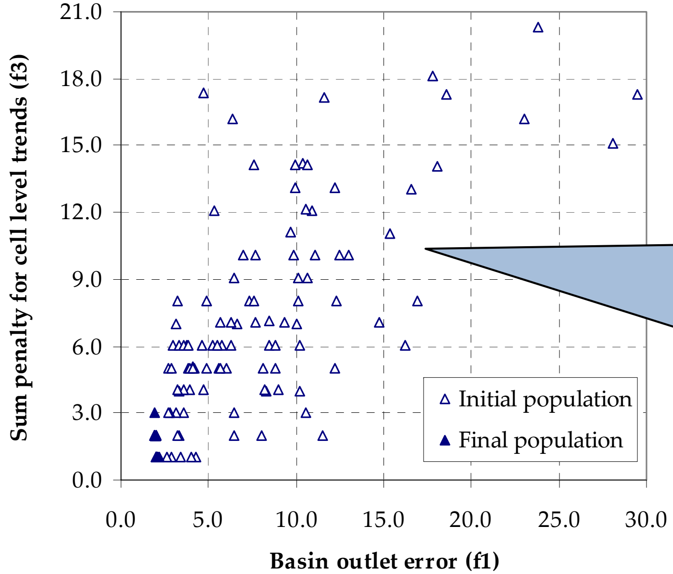
Example B: Calibration approaches

- **Hybrid calibration (weighted objective function)**
  - Step-by-step optimization of relatively small groups of parameters.
  - Manual rejection of solutions performing poorly against even one criterion (either in calibration or in validation) or providing unreasonable parameter values.
  - Very effective while particularly time-consuming strategy, primarily driven by the hydrological experience.
- **Pareto-based calibration**
  - Estimation of 30 out of 54 parameters (the rest obtained from the hybrid approach).
  - "Decomposition" of the performance measure into three components: (i) efficiency and intermittency penalty for outlet runoff; (ii) sum of efficiency values and intermittency penalties for spring runoffs; (iii) sum of penalty functions (trend criteria) for ensuring reasonable fluctuation of all groundwater levels.
  - Optimization employed through the *multiobjective evolutionary annealing-simplex* method (Efstratiadis & Koutsoyiannis, 2008), allowing at most 3000 function evaluations (fewer, if compared to the hybrid calibration approach).
  - Constrained approach, i.e. search for promising compromises across the Pareto front, by imposing feasibility limits to the three objectives.
  - Choice of the most promising compromises, by comparing the overall model performance in calibration and validation.

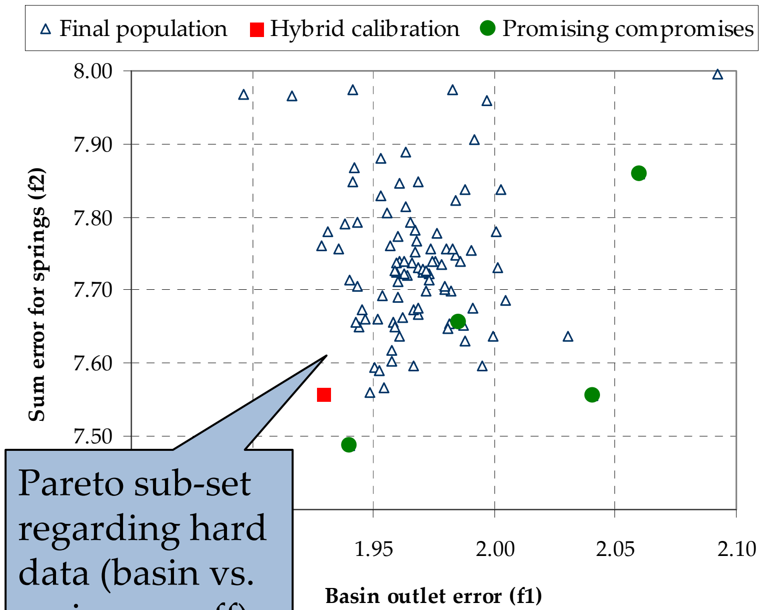
Example B: Characteristic trade-offs



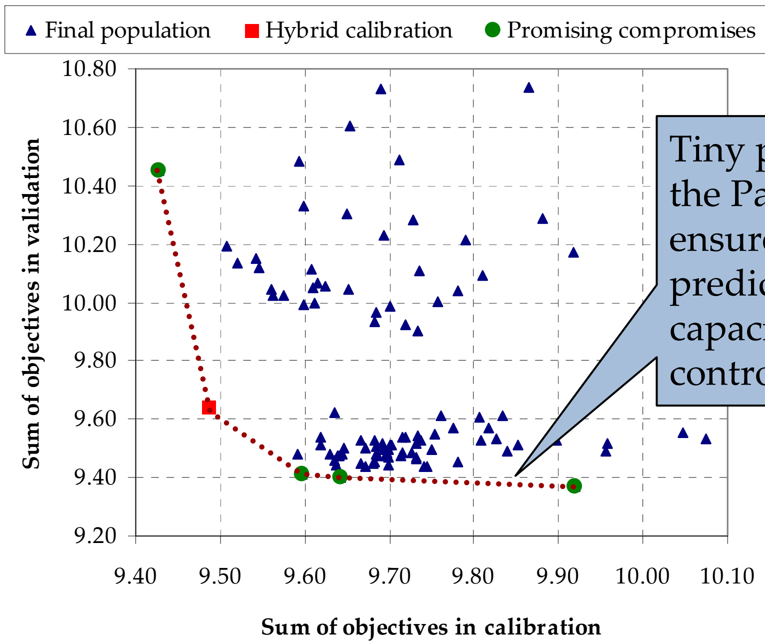
Enormous objective space, major trade-offs



Large areas of the objective space are far from reality regarding groundwater dynamics



Pareto sub-set regarding hard data (basin vs. spring runoff)



Tiny portion of the Pareto set ensures good predictive capacity for both control periods

Conclusions and perspectives

- Key issues of the actual multiobjective calibration research …
  - Impressively rapid development of novel computational tools, provided by many disciplines → quest for more comparative studies and wider dissemination of them in the every-day hydrological practice.
  - Adaptation of the principle of parsimony to distributed and conjunctive hydrological models → quest for more flexibility on model schematization and parameterization, based on data availability.
  - Recognition of the value of multi-criteria information in calibration → quest for more "hard" data **(the foundation of hydrology!**), quest for formulating criteria accounting for "soft" data and the engineering experience.
  - Effective optimization of conflicting criteria → quest for "filtering" principles and related procedures, to detect the most promising solutions through the Pareto-optimal set (e.g. calibration across "space", validation across time).
- … and some questions to be answered:
  - How close (or far) is a unified approach to model calibration and uncertainty assessment?
  - How can the multiobjective paradigm be effectively combined with Bayesian approaches of parameter uncertainty?
  - Can the hydroinformatics community help hydrologist of practice via developing tools allowing interactive multiobjective calibration?

References

- Confesor, R.B. & Whittaker, G.W. (2007) Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and Pareto optimization. *J. Amer. Wat. Res. Assoc.* 43(4), 981-989.
- Efstratiadis, A. & Koutsoyiannis, D. (2008) Fitting hydrological models on multiple responses using the multiobjective evolutionary annealing-simplex approach. *Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications* (ed. by Abrahart, R. J., See, L. M. & Solomatine, D. P.), Springer DE, Vol. 68.
- Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E. & Koutsoyiannis, D. (2008) HYDROGEIOS: A semi-distributed GIS-based hydrological model for modified river basins. *Hydrol. Earth System Sci.* 12, 989-1006.
- Gupta, H.V., Sorooshian, S. & Yapo, P.O. (1998) Toward improved calibration of hydrologic models: multiple and non-commensurable measures of information. *Wat. Resour. Res.* 34(4), 751-763.
- Harlin, J. (1991) Development of a process oriented calibration scheme for the HBV hydrological model. *Nordic Hydr.* 22, 15-26.
- Khu, S.T. & Madsen, H. (2005) Multiobjective calibration with Pareto preference ordering: an application to rainfall-runoff model calibration. *Wat. Resour. Res.* 41, W03004, doi:10.1029/2004WR003041.
- Kuczera, G. & Mroczkowski, M. (1998) Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Wat. Resour. Res.* 34(6), 1481-1489.
- Madsen, H. (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.* 235, 276-288.
- Madsen, H. (2003) Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Adv. Water Resour*. 26, 205-216.
- Seibert, J. & McDonnell, J.J. (2002) On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration. *Wat. Resour. Res.* 38(11), 1241, doi:10.1029/2001WR000978.
- Seibert, J. (2000) Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrol. Earth System Sci.* 4(2), 215-224.
- Tang, Y., Reed, P. & Kollat, J. (2007) Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications. *Adv. Water Resour.* 30(3), 335-353.
- Tang, Y., Reed, P. & Wagener, T. (2006) How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrol. Earth System Sci.* 10(2), 289-307.
- Vrugt, J. A., Gupta, H.V., Bastidas, L.A., Bouten, W. & Sorooshian, S. (2003) Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Wat. Resour. Res.* 39(8), 1214, doi:10.1029/2002WR001746.
- Yapo, P.O., Gupta, H.V. & Sorooshian, S. (1998) Multi-objective global optimization for hydrologic models. *J. Hydrol.* 204, 83-97.

This presentation is available on-line at:

http://www.itia.ntua.gr/en/docinfo/901/

Contact info:

andreas@itia.nua.gr

dk@itia.ntua.gr