Hydrol. Earth Syst. Sci. Discuss., 7, 8265–8308, 2010
www.hydrol-earth-syst-sci-discuss.net/7/8265/2010/
doi:10.5194/hessd-7-8265-2010

Hydrology and Earth System Sciences Discussions

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Holistic versus monomeric strategies for hydrological modelling of modified hydrosystems

I. Nalbantis[1], A. Efstratiadis[2], E. Rozos[2], M. Kopsiafti[2], and D. Koutsoyiannis[2]

[1]Laboratory of Reclamation Works and Water Resources Management, School of Rural and Surveying Engineering, National Technical University of Athens, Athens, Greece
[2]Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Athens, Greece

Received: 30 September 2010 – Accepted: 5 October 2010 – Published: 20 October 2010

Correspondence to: I. Nalbantis (nalbant@central.ntua.gr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The modelling of modified basins that are inadequately measured constitutes a challenge for hydrological science. Often, models for such systems are detailed and hydraulics-based for only one part of the system while for other parts oversimplified models or rough assumptions are used. This is typically a bottom-up approach, which seeks to exploit knowledge of hydrological processes at the micro-scale at some components of the system. Also, it is a monomeric approach in two ways: first, essential interactions among system components may be poorly represented or even omitted; second, differences in the level of detail of process representation can lead to uncontrolled errors. Additionally, the calibration procedure merely accounts for the reproduction of the observed responses using typical fitting criteria. The paper aims to raise some critical issues, regarding the entire modelling approach for such hydrosystems. For this, two alternative modelling strategies are examined that reflect two modelling approaches or philosophies: a dominant bottom-up approach, which is also monomeric and very often, based on output information and a top-down and holistic approach based on generalized information. Critical options are examined, which codify the differences between the two strategies: the representation of surface, groundwater and water management processes, the schematization and parameterization concepts and the parameter estimation methodology. The first strategy is based on stand-alone models for surface and groundwater processes and for water management, which are employed sequentially. For each model, a different (detailed or coarse) parameterization is used, which is dictated by the hydrosystem schematization. The second strategy involves model integration for all processes, parsimonious parameterization and hybrid manual-automatic parameter optimization based on multiple objectives. A test case is examined in a hydrosystem in Greece with high complexities, such as extended surface-groundwater interactions, ill-defined boundaries, sinks to the sea and anthropogenic intervention with unmeasured abstractions both from surface and groundwater. Criteria for comparison are the physical consistency of parameters, the reproduction

HESSD

7, 8265–8308, 2010

of runoff hydrographs at multiple sites within the basin, the likelihood of uncontrolled model outputs, the required amount of computational effort and the performance within a stochastic simulation setting.

1 Introduction and motivation

Two different general approaches or philosophies are applied in modelling of natural processes at large scales (in the order of at least a few km^2): The first approach, called bottom-up or upward (BU), seeks to exploit knowledge (typically physical laws) at the micro-scale (in the order of a few m^2) and then proceeds to larger scales through spatial aggregation. The second approach, called top-down or downward (TD), examines processes directly at the large scale and then eventually proceeds to making inferences about processes at smaller scales (Klemeš, 1983; Sivapalan et al., 2003b). Apart from this categorization of modelling approaches whose criterion is the base spatial scale of process representation, another categorization arises when the criterion is the level of modelling detail: Very often, some parts of the studied system are modelled in detail (in space-time) while for other parts simplified models are employed; in that way essential interactions among system components may be poorly represented or even omitted; we will call this approach monomeric (M), which originates from the Greek words "μόνος" and "μέρος" respectively denoting "solely" and "part". Conversely, when all parts of the studied system are modelled in the same detail and are linked via feedback mechanisms the approach will be called holistic (H) from the Greek word "όλον", which means "whole". The distinction between the monomeric and the holistic character of the approach can be extended to the type of information about the system that is exploited in modelling; in a monomeric approach often limited information is used, which encompasses a small number of measured system outputs or responses whereas in an holistic approach often one seeks to exploit information that is more general. Apart from observations, any qualitative information about the system

HESSD
7, 8265–8308, 2010

responses also can be taken into account together with the empirical interpretation of the unmeasured variables (whether output or internal ones).

Focusing on hydrological modelling, Savenije (2009) pointed out that "*. . . the dominant paradigm of hydraulics is reductionism, or a bottom-up approach, whereas in hydrology it is (or should be) empiricism and a top-down approach looking for links with fundamental laws of physics.*" The implementation of the BU approach into hydrology has led to modelling of hydrological processes at the small scale (e.g., local, plot or hill-slope), which has been an active research area in recent years (Zhang and Savenije, 2005; Zehe et al., 2006; Bárdossy, 2007). The practical usefulness of such models lies in allowing hydrological predictions at the catchment scale, supposedly without using any information on hydrological responses (Kilsby et al., 1999). Essentially this was the initial central focus of the "Predictions in Ungauged Basins (PUB)" initiative (Sivapalan et al., 2003a).

Critiques on the fundamental limitations of this approach, promising substantial reduction of uncertainty through reduction (i.e., theoretical explanation of small-scale processes) rather than deduction (i.e., explanation based on "lumped" response data) have appeared recently (Koutsoyiannis et al., 2009), but the underlying idea has been criticised from its early steps (Beven, 1989). Savenije (2009) reports examples where the BU approach fails while taking a broader perspective of the system under study through a top-down (TD) approach manages to better explain reality. Applications of the latter approach, which is rather macroscopic and, in this sense, holistic, are few (e.g., Tekleab, 2010).

The problems of the bottom-up approach become apparent when hydrological models are called to support engineering and management decisions, i.e., meet their major role (Efstratiadis and Mamassis, 2009). Supporting of decisions often requires modelling hydrosystems that involve extended surface-groundwater interactions and extended anthropogenic interventions in the hydrological cycle, such as abstractions from surface water bodies, pumping, and returns through artificial drainage systems. Theoretically, applying the BU approach for such modified hydrosystems would necessitate

HESSD
7, 8265–8308, 2010

putting together all physical processes and process interactions. Obviously, this would require a tremendous amount of information, which is absent in every real-world application. What is very frequently encountered within the BU approach is the monomeric character of modelling as this is defined earlier in this section. For example, a very detailed model is often formulated for one part of the system (or subsystem), while using oversimplified models for other parts or even ignoring dynamic links between subsystems. More often than not, the focus is on the detailed hydraulic model of a specific subsystem, such as an aquifer. According to the two categorization criteria the approach will be characterized as bottom-up/monomeric (BU-M). Although one may say in advance that this is simply a bad modelling practice, which merits no further study, the use of such approach is still so widespread that analysis of its implications is, to our view, justified.

We will concentrate our effort to modelling of complex, human-modified hydrosystems, which is a practical problem of high interest. To represent the BU-M approach we will consider a particular modelling strategy, called here strategy A. This focuses on hydraulic modelling of one natural subsystem only, which is the basin aquifer. To cope with the system complexities, a multi-stage modelling process is used that involves five stages: (1) splitting the hydrosystem into a number of natural sub-systems (sub-basins and the aquifer) and one man-made sub-system; (2) modelling natural sub-systems individually; (3) transferring predictions from the natural sub-systems to the man-made sub-system; and (4) optimizing the operation of the latter sub-system to represent as close as possible the observed conditions of the past (calibration). This is typically the strategy followed in engineering studies with the aid of popular commercial computer packages for water resources management. It presupposes that: (a) pure natural sub-systems can be effectively found, and (b) sufficient information is available for each sub-system modelled.

Strategy A may lead to erroneous predictions in complex basins where no simple natural sub-systems can be identified due to complex water exchanges. Moreover, inadequate information on some sub-systems may prohibit successful modelling. Data

HESSD
7, 8265–8308, 2010

inadequacy involves, among others, ill-defined system boundaries due to unknown leakages, sinks to the sea and anthropogenic intervention with unmeasured abstractions both from surface and groundwater. In the last years some approaches have appeared that cope with some of the above problems, although they do not cover the case of modified basins (e.g., Panday and Huyacorn, 2004; Bari and Smettem, 2004; Singh and Bhallamudi, 1998; Gauthier et al., 2009) neither do they treat the case with unknown abstractions (e.g., Schoups et al., 2005).

An alternative modelling strategy, called here strategy B, will be used to represent a top-down/holistic approach. The hydrosystem is viewed as a whole, having the input and the required information as guides to formulate spatial modelling units and process models. Ultimately, this approach leads to model integration, parsimonious parameterization and simultaneous optimization of all model parameters. All these provide flexibility to strategy B, which may be critical in cases with modified but poorly measured hydrosystems.

The motivation for this work is to test the applicability of modelling strategy A when the latter is used for modified hydrosystems. The target is precisely to examine the every-day modelling strategies in a critical spirit. It is the effects of these strategies that are investigated and not the value of the models used therein. In this respect, our work differs from the few comparative studies reported in the literature, such as the distributed model intercomparison project (Smith et al., 2004). The potential benefit when the problems of strategy A are faced is evaluated through applying strategy B. To achieve this, we extended the typical split-sample procedure for model building (i.e., calibration/validation based on historical data) to examining also the system response under hypothetical future conditions, using synthetically-generated forcing (stochastic simulation). This offers advantages over the typical model validation procedure: (1) it helps testing a modelling strategy within a framework that is similar to an operational one; (2) it can help examining future water management scenarios that are different from the historical ones; (3) it provides an opportunity to check for unreasonable long-term statistical trends or jumps of any model variable, which tests model credibility

HESSD
7, 8265–8308, 2010

(e.g., the model should normally generate stationary outputs, if fed with stationary forcing); (4) it can provide estimates of the uncertainty of model predictions against synthetic forcing data (e.g., precipitation, temperature).

To implement the two strategies, two modelling frameworks were chosen. The choice merely reflects the authors' experience on specific models. Modelling framework A implements strategy A and is based on the well-known groundwater package MODFLOW, coupled with a simple infiltration scheme. Modelling framework B, which is chosen to implement strategy B, uses a recently proposed framework that integrates a semi-distributed rainfall-runoff model, a coarse groundwater model and a network-type water allocation model (Efstratiadis et al., 2008).

A challenging operational case study was chosen involving the Boeoticos Kephisos river basin, Greece. This comprises all complexities described above and has been studied by the authors in the past (Nalbantis and Rozos, 2000; Nalbantis et al., 2002; Rozos et al., 2004; Efstratiadis et al., 2008). All above works present sequentially improved modelling strategies, from relatively simple to more detailed ones, which are consistent with the top-down/holistic type of approach. Taking advantage of this effort, we detected and investigated five key modelling options within the selected modelling strategies, which are discussed hereinafter.

2 Key modelling options in hydrological modelling strategies

When formulating a modelling strategy, critical decisions are made in regard to selecting, formulating and fitting hydrological models. These decisions lead to defining key modelling options that constitute the "ingredients" of the formulated strategy; these options are described next.

HESSD
7, 8265–8308, 2010

2.1 Key modelling option SW-GW: link between models for surface and groundwater processes

In strategy A, different stand-alone models are used for surface and groundwater hydrology, which precludes accounting for feedback interaction between the corresponding processes. Very often, the two models differ in the spatial and temporal scale used. They may also differ in their modelling approach or philosophy. For, the surface processes are usually represented via conceptual (hydrological) approaches; even the fully-distributed physically-based schemes are considered conceptual at the grid scale (Beven, 1989). Yet, groundwater modelling typically follows a hydraulic rather than a hydrologic rationale. All these aspects affect the parameter estimation procedure, which requires either to provide unrealistic simplifications (e.g., assume that the entire runoff is derived from the surface system, thus omitting the contribution of groundwater runoff) or (rarely) make successive approximations, i.e., calibrate the one component after the other, which is computationally inefficient. In strategy B, the main hydrological interactions are explicitly represented, and thus model parameters can be simultaneously optimized, taking advantage of the available measurements across both components (e.g., flow and piezometric data).

2.2 Key modelling option SW-GW-WM: link between models for hydrological processes and water management

In the staged modelling procedure of strategy A, hydrological models are constructed exclusively for undisturbed parts of the system (e.g., sub-basins) and the outputs thereof (e.g. river flows) are transferred as inputs to the water management model of the man-made sub-system (usually implemented within a Decision Support System or DSS). This serial operation, apart from being computationally inefficient, suffers from a number of drawbacks: (a) it is unrealistic when decision-related interactions between hydrological and man-made sub-systems are significant; (b) it is infeasible when real abstractions are unknown, since it precludes the assessment of boundary

HESSD
7, 8265–8308, 2010
Holistic versus monomeric strategies
I. Nalbantis et al.

conditions for hydrological and hydrogeological models; (c) it may impose serious limitations to the representation of the water allocation problem (e.g., a "first-come, first-served" management policy may be mandatory); (d) it requires data transfer between models, which puts stringent requirements regarding space-time scale compatibility of hydrological model outputs and inputs to the water management model; (e) it precludes automatic calibration of models, which, in presence of the above problems, is at least questionable (Efstratiadis et al., 2010). Attempts to cope with the above problem are rare in literature (Fredericks et al., 1998; Dai and Labadie, 2001). Strategy B adopts model integration, which copes with the problem.

2.3 Key modelling option SCALE-PARAM: link of spatial scale and model parameterization

Since parameterization is designed to represent factors that influence the spatial variability of hydrological processes, it is naturally linked to the spatial scale (Klemeš, 1983). The large heterogeneity of mechanisms and properties makes it difficult to achieve compatibility between measurements made at the local scale and model predictions. Quite often, in strategy A, very detailed models are chosen in the hope to achieve scale compatibility between data and predicted variables. Yet, the resulting high dimensionality leads to extremely time-consuming schemes, which is a major restricting factor affecting not only calibration but also the operational applicability of models; the latter arises because models have to co-operate with DSS that run in forecast mode, using synthetic forcings for long time horizons (Nalbantis et al., 2002).

2.4 Key modelling option SCHEM-PARAM: link between hydrosystem schematization and parameterization

Inevitably, in strategy A, the hydrosystem schematization (i.e., the simplification of the process representation in space) dictates parameterization. Parameters are assigned to individual spatial elements (e.g., sub-basins, grid cells), thus having limited physical meaning. Moreover, due to the detailed spatial scale adopted, the principle of parsimony is broken, which results to poorly identified parameters (Kuczera and Mroczkowski, 1998) and increased predictive uncertainty (Efstratiadis and Koutsoyiannis, 2010; Savenije, 2010). Attempts to reduce the number of control variables of the optimization problem require hybridized strategies, such as detecting only the most important parameters while estimating the rest of them on the basis of field data (Refsgaard, 1997; Eckhardt and Arnold, 2001) or using zonation approaches (i.e. spatial grouping of parameters). Contrary to the above, in strategy B schematization and parameterization are disconnected, thus ensuring that models are by construction parsimonious. In this approach, the schematization is adapted to the engineering objectives (i.e., which processes should be simulated and where), while the parameterization is only linked to the available information (cf. Dehotin and Braud, 2008).

2.5 Key modelling option OPT: appropriate use of optimization in calibration

In theory, physically-based approaches enable their free variables to be derived from field measurements. Yet, in practice, their applicability is significantly restrained not only by the heterogeneity of processes and the unknown scale dependencies of parameters (Beven, 2001; Wagener et al., 2001; Rosberg and Madsen, 2005), but also by the high computational effort and the subsequent inability to co-operate with DSSs (Nalbantis et al., 2002). Hence, optimization is always required for at least some of model parameters (Refsgaard, 1997; Eckhardt and Arnold, 2001). However, within strategy A, non-expert users frequently adopt a "black-box" approach, seeking for a "global optimal" parameter set (typically against a single, quantitative performance criterion), through an automatic algorithmic procedure. Very often, this leads to: (a) parameter values that are inconsistent with their physical interpretation; (b) poor predictive capacity against independent control data (validation); (c) unreasonable values of uncontrolled responses (e.g., evapotranspiration, underground losses) and internal model variables (e.g., soil and groundwater storage) (Rozos et al., 2004; Efstratiadis et al., 2008). On the other hand, strategy B emphasizes less on the optimization task and

HESSD
7, 8265–8308, 2010

more on the comprehensive understanding of the problem components (real system, model and data), to ensure reliable results. Particularly in models of complex parameterization, it aims to increase the information that is exploited in calibration as much as possible, by means of both multi-response measurements and empirical metrics ("soft" data), which account for the hydrological expertise (Efstratiadis and Koutsoyiannis, 2010).

3 Overview of alternative modelling frameworks

3.1 Modelling framework A

Within this framework an off-line coupling of MODFLOW is made with a simple rainfall-runoff model, as illustrated in Fig. 1, left. MODFLOW is a classical tool for 3-D simulation of groundwater flow, where the flow field is discretized into a number of rectangular cells and all quantities are referred to cell centres. The 3-D continuity equation and the Darcy's law written in finite differences form provide one final flow equation for each cell as a function of the unknown hydraulic head h and other known variables. The latter are either parameters of the aquifer (conductances for the three axis directions and the specific yield) or external stresses of the cell (percolation from soil or the river bed, pumped water, water outflow to the sea). After defining initial and boundary conditions, the final system of equations on h is solved via the Preconditioned Conjugate Gradient method.

For model implementation, we used the computer package by Waterloo Hydrogeologic Inc. (1999). This includes an optimization module, employing a deterministic local-search type method for the automatic calibration of model parameters.

The rainfall-runoff process is represented through two modelling components; the first calculates the effective rainfall on the basis of precipitation and potential evapotranspiration, whereas the second divides the effective rainfall into runoff and percolation, assuming a constant ratio between these two variables, which stands as the single

HESSD
7, 8265–8308, 2010

model parameter. The above scheme runs independently, to provide external stresses to MODFLOW cells (due to percolation).

3.2 Modelling framework B

In modelling framework B the computer package HYDROGEIOS 2.0 is used. This is a GIS-based model suite, which allows for flexible representation of hydrological processes, as summarized below and more analytically described by Efstratiadis et al. (2008). A synoptic description of the modelling framework is illustrated in Fig. 1, right. More precisely, surface flow is considered within the hydrographic network, which is extracted from a digital terrain model through adjusting the flow accumulation parameter and adding control points that correspond to flow measurement stations or diversion nodes. The network conveys flow of sub-basins, which are subject to different hydrological stresses (precipitation and potential evapotranspiration). Surface processes are considered homogeneous within partitioned areas of the basin termed as the Hydrological Response Units (HRUs). This idea has found a limited number of applications to distributed modelling also (Flügel, 1995; Srinivasan et al., 2000; Gurtz et al., 2003). The HRUs represent soil and land use types within portions (patches) of the studied area, which are not necessarily contiguous. They are defined as the product of partitions of the basin on the basis of different properties, such as soil permeability, land cover and topography (i.e. terrain slope). Through classification of the above properties, one can adjust the number of HRUs and, consequently, the number of the parameters of the surface hydrological model; the latter comprises six parameters per HRU and transforms precipitation into real evapotranspiration, percolation and surface runoff (direct, overland, lateral) for each sub-basin and HRU combination.

Groundwater flow is modelled through a multi-cell approach involving discretization of the aquifer into non-rectangular cells (Rozos and Koutsoyiannis, 2006, 2010). This allows the description of complex geometries on the basis of the physical characteristics of the aquifer (e.g., geology), through parsimonious structures. Two parameters are assigned to each cell (conductivity and specific yield). Springs and underground

HESSD
7, 8265–8308, 2010

losses are implemented as virtual cells of very large base, which allows keeping almost constant hydraulic head regardless storage variations. Model stresses are: (a) areal inflows due to percolation through each sub-basin and HRU combination; (b) inflows due to infiltration underneath each river segment; (c) outflows due to pumping from each borehole.

Regulated flow through man-made structures and portions of the hydrographic network is modelled with the aid of a water management network, which is an extension of the scheme described by Efstratiadis et al. (2004). This has as nodal inflows the surface and the groundwater runoff, as nodal outflows the withdrawals for water uses, and as distributed fluxes the water losses due to infiltration and river discharge. Major hydraulic works are also represented along with the corresponding water uses and constraints as well as their interactions with the natural system. Model properties are discharge and pumping capacities, target priorities, demand time series and unit transportation costs of water. The allocation of flows is based on a linear programming approach where virtual unit costs, positive or negative, are assigned either to prohibit undesirable fluxes or to force the model fulfil the hydrosystem targets (Efstratiadis et al., 2010).

HYDROGEIOS embeds an advanced calibration module that provides a number of statistical and empirical criteria for model fitting on multiple responses (river and spring discharge, hydraulic heads) and various options regarding the delineation of the feasible search space. Optimization is carried out through the evolutionary annealing-simplex method (Efstratiadis and Koutsoyiannis, 2002; Rozos et al., 2004).

4 Case study

4.1 The study area

The Boeoticos Kephisos river basin lies on the Eastern Sterea Hellas, to the north of Athens, and drains a closed area of 1930 km^2 (Fig. 2). The catchment is formed on heavily karstified limestone covered with alluvial deposits in plain areas. Considerable groundwater amount (more than half of the annual catchment runoff) is discharged through large springs in the upper and middle part of the basin, whereas an unknown amount is leaking to the sea.

The surface runoff reaching the basin outlet is conducted to the neighbouring Lake Hylike, which is one of the major reservoirs of Athens water supply system. Groundwater of the middle part is also considered as an emergency resource. Finally, a significant part of the surface and groundwater resources of the basin is used in agriculture; more specifically, river abstractions are mainly implemented in the lower part of the basin, practically eliminating flow availability during the summer months. For a more detailed description of the basin features, the reader may refer to previous publications (Rozos et al., 2004; Efstratiadis et al., 2008).

For the study area, a major question is about the impacts of abstractions though the Vasilika-Parori boreholes to the overall hydrological regime of the basin. These boreholes were drilled within the frame of emergency measures taken during a severe drought in the period from 1989 to 1994, at the end of which almost all surface resources dried out. Yet, due to the considerable reduction of precipitation and the intense pumping for providing drinking water to Athens, the discharge of the neighbouring Mavroneri springs was interrupted twice during 1990 and 1993, thus resulting to various social and environmental problems (these springs account for 15% of the total basin runoff). In the last 15 years, several engineering studies were carried out to investigate the complex dynamics of the hydrosystem and provide reliable estimations regarding the consequences of water supply abstractions to the system responses. Modelling framework A originates from an earlier engineering study (Nalbantis and Rozos, 2000) while framework B capitalizes the experience gained after continuous research attempts (Nalbantis et al., 2002; Rozos et al., 2004; Efstratiadis et al., 2008).

HESSD
7, 8265–8308, 2010

4.2 Model schematization and parameterization

Within strategy A, the karst aquifer is considered as a single layer with free surface flow and cell size varying from $800 \times 800\,\mathrm{m}^2$ near the boundaries to $150 \times 150\,\mathrm{m}^2$ in central areas (near the springs); this resulted to a total number of cells equal to 3631. To reduce the number of control variables, the flow field was divided into zones (Fig. 3), which resulted to totally 18 parameters to calibrate. The alluvial aquifer was not simulated since its water yield is low and of limited interest. In the whole perimeter of the karst aquifer no-flow boundaries were assigned, apart from two areas: a small area in SW, where lateral recharge was considered using dummy recharge wells, and the NE boundary, where small-scale outflows to the sea were simulated through a cluster of pumping wells. Unsteady flow was assumed whereas external stresses due to percolation were modelled in a simple way as explained in Sect. 3.1. Precipitation was considered homogenous in the interior of three sub-basins, which represent different hydrogeologic units (low, middle and upper course). Infiltration depth was taken as a constant fraction of the effective precipitation, which varied from one type of surface geological formation to another; three types were assumed, corresponding to limestone, alluvial and flysch (impermeable) formations. Stream-aquifer interactions were represented using the River module of MODFLOW assuming zero infiltration during the April–August period. No water management model was explicitly considered; to estimate unmeasured withdrawals from groundwater; the irrigated area that is supplied by each well component and the related water demand per unit area were used by assuming (wrongly) that the entire demand is fulfilled via pumping (Nalbantis and Rozos, 2000; Nalbantis et al., 2002).

Strategy B followed a semi-distributed schematization to 15 sub-basins, as shown in Fig. 2. The spatial average of precipitation of each sub-basin was calculated through the Thiessen method, whereas potential evapotranspiration series was estimated via the Penman-Monteith method, by adjusting air temperatures to the mean elevation of each sub-basin. For the definition of the HRUs, three categories of permeability (low, medium, high) and two categories of terrain slope were taken. The groundwater flow field was discretized into 40 non-rectangular cells, four of which represent underground leakages to neighbouring basins and the sea. In addition, six dummy cells were used to model surface outflows through the major karstic springs (Fig. 4). This spatial discretization is two orders of magnitude coarser than the one implemented within strategy A, thus making the computational effort for groundwater simulation almost negligible, in comparison to that of strategy A. For the parameterization of the groundwater flow field, we used three categories of permeability and porosity, which reflect topography and geology. Moreover, we used particular permeability values for the rest of cells representing springs and underground losses (ten parameters in total). In combination with the parameterization of the surface processes via the six HRUs, the total number of model parameters was 52 (36 for the surface model and 16 for the groundwater model). We note that both the schematization and parameterization differ from those reported by Efstratiadis et al. (2008), in an attempt to significantly reduce the number of groundwater parameters; this ensures a more parsimonious modelling approach (Kopsiafti, 2009).

The water management network includes consumption nodes for agricultural areas, abstraction nodes (river diversion works and borehole groups) and aqueducts. Water management policy was implemented through assigning water supply targets to a number of consumption nodes (seven for irrigation and one for water supply) and virtual costs to the system aqueducts with the purpose to represent a realistic and close to the actual abstraction priority for water uses (e.g., in case of combined abstractions, priority was given to river abstractions instead of pumping, which is the historical practice). This is a critical difference between the two strategies, since strategy A requires known abstractions whereas strategy B is much more flexible, since it is based on theoretical demands, costs and priorities, thus allowing to choose among different abstraction policies to fulfil demands. For instance, pumped water through Vasilika-Parori boreholes is either conducted downstream, for the irrigation of Kopais plain, or diverted

HESSD
7, 8265–8308, 2010

for the water supply of Athens. Withdrawals from other resources, both surface and groundwater, are also implemented to serve local agricultural demands (Fig. 5).

4.3 Calibration strategies and data for parameter estimation

Both strategies were tested against the observed hydrological responses for a 10-year control horizon (October 1984–September 1994), employing a monthly time step. For this period, discharge series at seven locations are available, precisely at the basin outlet (Karditsa tunnel) and downstream of the six karst springs, illustrated in Fig. 2. With regard to groundwater, several level gauges were available, mostly located in the vicinity of the main river branch.

Regarding the rainfall-runoff component of the framework A, no attempt was made for parameter optimization, since infiltration fractions were estimated empirically, on the basis of the main geological formations of the basin. The MODFLOW parameters were manually optimized on the basis of 18 observed level series and through visual inspection of the closeness of observed and simulated spring hydrographs.

In strategy B, as thoroughly explained by Efstratiadis et al. (2008), the parameters of the surface and the groundwater models were simultaneously calibrated. Since the number of parameters was large, to cope with the resulted uncertainty which is directly related to equifinality (Freer et al., 1996), multiple criteria were taken into account, including "soft" data (Seibert and McDonnell, 2002; Efstratiadis and Koutsoyiannis, 2010). Thus, a weighted objective function was formulated comprising the following statistical and empirical measures: (a) efficiency and bias of the monthly hydrographs at the seven locations mentioned above; (b) penalties for not reproducing flow intermittency; and (c) penalties for generation of unrealistic trends regarding the groundwater levels. The first group of criteria accounts for the so-called "hard data" (measurements); this is essential for reproducing the global water balance and the spring mechanisms, but not sufficient for representing the regime of the groundwater fluxes; these criteria will be later referred to as performance indices. The second group simply accounts for the information "zero" or "non-zero discharge", which is easily observable, reliable and is of major interest in water management. Finally, the third group of criteria is a kind of "soft" data, ensuring reasonable fluctuation of the non-observable internal variables of the model. Optimization was carried out through a hybrid strategy, which combines human experience and automatic tools (Boyle et al., 2000; Rozos et al., 2004). In that manner, search was guided towards a realistic, best-compromise parameter set, ensuring satisfactory predictive capacity for all model responses.

4.4 Operational use of models through stochastic simulation

As mentioned in the Introduction, calibration is essentially an intermediate step in the modelling procedure, which allows for optimising the predictive capacity of the model on the basis of observed data. Yet, the reproduction of the past responses has limited interest, if not accompanied by further analyses with "projected" inputs, thus providing support to decisions for future. In this respect, the two strategies are evaluated within a stochastic (Monte Carlo) simulation framework, aiming to examine the system response under different stress conditions, comprising both natural (precipitation, potential evapotranspiration) and anthropogenic (abstractions from surface and groundwater recourses) forcing.

For the representation of rainfall, a multivariate stochastic scheme was used to generate point series of 1000-year length, which preserve the essential statistical characteristics of the observed samples of 12 rain gauges across the basin, at the annual and monthly time scales (Koutsoyiannis et al., 2003). Next, the point series were aggregated to the appropriate spatial scale, thus providing areal rainfall series for the 3 and 15 sub-basins, which correspond to modelling frameworks A and B respectively.

The synthetic rainfall records were divided into clusters of ten-year length, to formulate a hundred of statistically equivalent forcing scenarios. Each modelling scheme ran 100 times, each one under different stochastic forcing, whereas for all runs the same initial conditions were supplied. For instance, for modelling framework B we used the soil moisture depths and groundwater levels of the beginning of calibration (October 1984). The model outputs within this configuration represent statistically consistent trajectories of the system responses, for a ten-year horizon. Such type of Monte Carlo analysis is typically used in operational applications (e.g., forecasts), and is also known as terminating simulation (Koutsoyiannis et al., 2003; Koutsoyiannis, 2005).

Regarding the potential evapotranspiration throughout each sub-basin, we assigned the same areal values with the control period 1984–1994, since this is a forcing variable of very low interannual variability. Finally, for the anthropogenic forcing, we examined two alternative water management scenarios, to evaluate the impacts of water supply abstractions to the system responses (especially in the middle course of the basin). For both policies, we assigned the actual agricultural demands across the basin (223 hm^3/yr) and assumed either zero or extensive (46 hm^3/yr, equal to that of the water year 1993–1994) demand through the Vassilika-Parori boreholes, for providing drinking water to Athens (Fig. 5). We remind that in modelling strategy A, the water requirements are by definition fulfilled via pumping (which is an erroneous yet obligatory assumption), while in strategy B the demands can be satisfied through multiple sources, following the cost optimization approach.

5 Results

5.1 The testing framework

Ideally, testing the effect of adopting modelling strategy A for modified hydrosystems, would require a complex computer experiment based on a series of alternative frameworks implementing strategy A, each one differing from framework B in only one key modelling option. Careful examination of the commonly available models shows that it is almost impossible to build frameworks that perfectly fulfil the above requirements. So, we simplified our experiment by considering the two frameworks presented in Sect. 3. Thus, the combined effect of all key modelling options is tested. Table 1 summarises the a priori knowledge about the relation of each key modelling option to each modelling framework. Application to a common data set of hydrological variables of the aforementioned test basin allowed an objective evaluation of the effect of adopting strategy A. For this, both numerical and empirical criteria are used for the comparison of the two approaches, in calibration and simulation mode.

We emphasize on the monthly flows at control points (i.e., the main karstic springs), whose hydrographs are also depicted for visual comparison. Comparisons on the river flow at the basin outlet were not feasible since framework A did not provide such output. Regarding the groundwater levels, we only examined their long-term behaviour rather than their actual values, since a direct comparison of aquifer water levels was meaningless, due to enormous differences of scale between the two alternative groundwater models used.

5.2 Comparison of model performance in calibration

The model performance during the calibration and the validation periods is evaluated on the basis of two criteria, efficiency values and bias in the mean. Due to the different assumptions regarding the system delineation (e.g., in strategy A, the river network and the alluvial areas of the aquifer were not simulated), comparisons were possible only at three observation sites, namely downstream of Mavroneri, Melas and Polygyra springs. In Tables 2 and 3, values of the corresponding performance indices are provided, which show a clear improvement of model performance in spring flow predictions when passing from strategy A to strategy B for both the calibration and the validation periods. This is confirmed by hydrograph comparisons on Fig. 6 for the two most important springs of the basin (Mavroneri and Melas). We note that the first approach, although concentrated on the detailed representation of the aquifer dynamics, fails to reproduce the key characteristics of the observed flows, namely the monthly variability, which is overestimated for Melas and underestimated for Mavroneri springs; it also fails to reproduce the interruption of the discharge of the latter, during 1990 and 1993.

On the other hand, the advantages of model integration as offered by strategy B are not restricted to flow predictions only. Other improvements are equally significant, which are commented in Table 4 and provide explanation of the superior performance

HESSD
7, 8265–8308, 2010

indices regarding observed groundwater variables, despite the fact that framework B follows a much simpler modelling approach, which focuses on the surface processes and the water management practices.

5.3 Comparison of model performance in simulation

Obviously, when comparing two modelling approaches in a stochastic simulation setting, it is impossible to use quantitative criteria (e.g., goodness-of-fit measures), as in calibration. Therefore, the evaluations are based on the grounds of common sense, i.e. testing whether the model provides the right answers for the right reasons (cf. Kirchner, 2006), taking advantage of the hydrological experience.

The implementation of the two frameworks under synthetically generated inputs further reveals the drawbacks of strategy A. In Fig. 7 we plot the projected discharge at Mavroneri springs for a ten-year horizon (mean value of 100 flow scenarios and 80% prediction limits), under zero and intensive pumping through the neighbouring boreholes at Vasilika-Parori (respectively referred to as “actual abstraction policy” and “intensive abstraction policy”). For both management policies, there is a sharp decrease of discharge, which is inconsistent with the experience so far. Although one could expect that under an extensive pumping a systematic negative trend could be possible, it is unlikely that such trend is encountered under the actual abstraction policy. The differences with the respective results obtained through the modelling approach B (Fig. 8) are substantial; here, the mean projection for the spring outflows (which corresponds to average rainfall conditions) follows a stationary pattern under the actual abstraction policy, while there is a progressive decrease of the spring resources under the intensive abstraction policy (Fig. 8). This indicates that, in a long-term perspective, the intensive use of the Vasilika-Parori boreholes for the water supply of Athens is not a sustainable option.

Regarding the response of the Melas springs according to modelling approach A, there are negligible differences between the two abstraction policies (Fig. 9). This is also an unexpected result, since the entire karst aquifer should be affected by the water supply abstractions, also because of the fact that the irrigation needs in the downstream part of the basin should be fulfilled though increased pumping in the vicinity of the aforementioned springs. Yet, this presupposes a proper description of the water management alternatives, which is only feasible thorough framework B. Following strategy B, it is projected that Melas springs are obviously affected by the upstream abstractions, and thus the slightly negative trend of the average flow trajectory is reasonable (Fig. 10).

Attempting to investigate the reasons for the unrealistic performance of modelling framework A, we concluded that we should revisit the calibration procedure. In contrast to strategy B, where we accounted for the internal variables of the model by assigning trend penalties on groundwater levels, in strategy A, we just attempted to fit the model responses (spring flows) on the observations. In that manner, we allowed systematic drains and fillings of the cells lying near the boundaries of the aquifer where spring flow observations are not available. Actually, we left calibration to assign unrealistic conductivity values in order to maximize the model performance against the efficiency indices. On the other hand, within framework B soft criteria are also employed but this required a more time-consuming and hard to automate calibration strategy. In Fig. 11 we compare the synthetically generated levels obtained through the two approaches, which correspond to the most upstream part of the karst system, assuming the actual abstraction policy. We observe a questionable behaviour (i.e., negative trend) of the projected level when applying modelling framework A, whereas for framework B the groundwater system exhibits stationary behaviour, which is more reasonable.

6 Conclusions

Our investigations have shown that in watersheds that are modified by human interventions the classical modelling strategy based on the monomeric bottom-up approach is, in general, inefficient. It makes use of a detailed hydraulic model for only a part of the studied system and of separate models for surface hydrological processes,

HESSD
7, 8265–8308, 2010

groundwater flow processes and water management processes. Such serial use of models prohibits the modelling of the process interactions and suffers from increased computational burden. The monomeric character of this approach reflects other modelling aspects, including the assignment of parameters (coincidence of spatial scales of schematization and parameterization). The calibration procedure is based only on measured system outputs; moreover, fully automatic search is employed and internal variable dynamics is ignored. All the above misuse practices are reflected to the predictive capacity of the model, which proved disappointedly poor for such an exhaustive effort (in calibration). This behaviour was identified by employing the model in stochastic simulation mode for operational use; the obtained projections are far from being realistic.

Conversely, a holistic top-down approach allows for model schematization and parameterization that respects the principle of parsimony and ensures computational efficiency by means of both simulation and optimization/calibration. This precludes taking advantage of the power of fully distributed models while, at the same time, favouring the use of the semi-distributed approach. An effective way to reduce the size of the parameter set is to decouple parameterization and model schematization through using Hydrological Response Units (HRUs). Through the HRU concept, the model structure depends on a limited number of landscape classes, whose parameters retain some physical consistency thus allowing for a better identification of their prior uncertainty (cf. Savenije, 2010). Last, model integration allows simultaneous calibration of all models through exploiting all kinds of information and not only information about some basic output variables. A hybrid process of manual and automatic optimization proved very effective in finding a best compromise solution while, at the same time, respecting some physical interpretation of parameters. In this approach all available pieces of information, including hydrological experience, are exploited in model calibration within a multiobjective framework.

Our tests proved that running models in stochastic simulation mode can be a useful tool for their testing and validation since this augments information supplied by typical calibration and validation procedures, while, at the same time, addressing some of the intrinsic restrictions of the available data (Kirchner, 2006). In fact, the stochastic framework offers new perspectives towards multiple paths. First, it allows for assessing the performance of hydrological models in situations that are consistent to those in which it is supposed to be used in practice (Klemeš, 1986). In addition, it provides a "crash-test" for evaluating the model transposability in time, which is a necessary condition for their operational adequacy (Andréassian et al., 2009). Finally, through a proper representation of the varying character of climate and the related processes in stochastic terms, it ensures reliable estimation of the uncertainty and the long-term risk in hydrological studies and water resources management (Koutsoyiannis et al., 2007; Koutsoyiannis, 2010). This can also include the evaluation of extremes, which are not represented in the (usually limited) calibration data (Seibert, 2003).

We believe that the research presented in this paper can contribute in (1) formulating specifications for model packages applicable to modified basins, and (2) opening new research routes regarding different types of approach followed in hydrological modelling.

Acknowledgements. The second version of the HYDROGEIOS package was carried out within the research project "Development of Database and software applications in a web platform for the National Databank for Hydrological and Meteorological Information". The software is freely available at http://www.hydroscope.gr/software/hydrogeios.htm.

References

Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M.-H., and Valéry, A.: *HESS Opinions* "Crash tests for a standardized evaluation of hydrological models", Hydrol. Earth Syst. Sci., 13, 1757–1764, doi:10.5194/hess-13-1757-2009, 2009.

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, doi:10.5194/hess-11-703-2007, 2007.

Bari, M. and Smettem, K. R. J.: Modelling monthly runoff generation processes following land use changes: groundwatersurface runoff interactions, Hydrol. Earth Syst. Sci., 8, 903–922, doi:10.5194/hess-8-903-2004, 2004.
Beven, K. J. : Changing ideas in hydrology – The case of physically-based models, J. Hydrol., 105(1–2), 157–172, 1989.
Beven, K.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, doi:10.5194/hess-5-1-2001, 2001.
Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, Water Resour. Res., 36(12), 3663–3674, 2000.
Dai, T. and Labadie, J. W.: River basin network model for integrated water quantity/quality management, ASCE J. Water Res. Plan. Manag., 127(5), 295–305, 2001.
Dehotin, J. and Braud, I.: Which spatial discretization for distributed hydrological models? Proposition of a methodology and illustration for medium to large-scale catchments, Hydrol. Earth Syst. Sci., 12, 769–796, doi:10.5194/hess-12-769-2008, 2008.
Eckhardt, K. and Arnold, J. G.: Automatic calibration of a distributed catchment model, J. Hydrol. 251, 103–109, 2001.
Efstratiadis, A., and Mamassis, N.: Evaluating models or evaluating modelling practices? – Interactive comment on HESS Opinions "Crash tests for a standardized evaluation of hydrological models", Hydrol. Earth Syst. Sci. Disc., 6, C1404–C1409, 2009.
Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E., and Koutsoyiannis, D.: HYDROGEIOS: a semi-distributed GIS-based hydrological model for modified river basins, Hydrol. Earth Syst. Sci., 12, 989–1006, doi:10.5194/hess-12-989-2008, 2008.
Efstratiadis, A. and Koutsoyiannis, D.: An evolutionary annealing-simplex algorithm for global optimisation of water resource systems, Proc. Fifth Intern. Conf. on Hydroinformatics, IWA, Cardiff, UK, 1423–1428, 2002.
Efstratiadis, A. and Koutsoyiannis, D.: One decade of multiobjective calibration approaches in hydrological modelling: a review, Hydrolog. Sci. J., 55(1), 58–78, 2010.
Efstratiadis, A., Koutsoyiannis, D., and Xenos, D.: Minimising water cost in the water resource management of Athens, Urban Water J., 1(1), 3–15, 2004.
Efstratiadis, A., Nalbantis, I., Rozos, E., and Koutsoyiannis, D.: Accounting for water management issues within hydrological simulation: Alternative modelling options and a network optimization approach, EGU General Assembly 2010, Vienna, Austria, 2–7 May 2010, Geophys.

Res. Abstr., 12, EGU2010-10085, 2010.
Flügel, W. A.: Delineating Hydrological Response Units by Geographical Information System analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the river Bröl in Germany, Hydrol. Process., 9(3–4), 423–436, 1995.
Fredericks, J. W., Labadie, J. W., and Altenhofen, J. M.: Decision support system for stream-aquifer management, J. Water Res. Pl.-ASCE, 124(2), 69–78, 1998.
Freer, J., Beven, K. J., and Ambroise, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach, Water Resour. Res. 32(7), 2161–2173, 1996.
Gauthier, M. J., Camporese, M., Rivard, C., Paniconi, C., and Larocque, M.: A modeling study of heterogeneity and surface water-groundwater interactions in the Thomas Brook catchment, Annapolis Valley (Nova Scotia, Canada), Hydrol. Earth Syst. Sci., 13, 1583–1596, doi:10.5194/hess-13-1583-2009, 2009.
Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., and Vitvar, T.: A comparative study in modelling runoff and its components in two mountainous catchments, Hydrol. Process., 17, 297–311, 2003.
Kilsby, C. G., Ewen, J., Sloan, W. T., Burton, A., Fallows, C. S., and O'Connell, P. E.: The UP modelling system for large scale hydrology: simulation of the Arkansas-Red River basin, Hydrol. Earth Syst. Sci., 3, 137–149, doi:10.5194/hess-3-137-1999, 1999
Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006.
Klemeš, V.: Conceptualization and scale in hydrology, J. Hydrol., 65, 1–23, 1983.
Klemeš, V.: Operational testing of hydrologic simulation models, Hydrolog. Sci. J., 31, 13–24, 1986.
Kopsiafti, M.: Investigation of parameterization strategies for the hydrogeological module of Hydrogeios software – Application to the Boeoticos Kephisos basin, Postgraduate thesis, 133 pages, Dept. of Water Resources, National Technical University of Athens, 2009.
Koutsoyiannis, D., Efstratiadis, A., and Georgakakos, K.: Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches, J. Hydrometeorol., 8(3), 261–281, 2007.
Koutsoyiannis, D.: Hurst-Kolmogorov dynamics and uncertainty, J. Amer. Wat. Res. Assoc., in press, 2010.

Koutsoyiannis, D., Karavokiros, G., Efstratiadis, A., Mamassis, N., Koukouvinos, A., and Christofides, A.: A decision support system for the management of the water resource system of Athens, Phys. Chem. Earth, 28(14–15), 599–609, 2003.

Koutsoyiannis, D., Makropoulos, C., Langousis, A., Baki, S., Efstratiadis, A., Christofides, A., Karavokiros, G., and Mamassis, N.: *HESS Opinions*: "Climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability", Hydrol. Earth Syst. Sci., 13, 247–257, doi:10.5194/hess-13-247-2009, 2009.

Koutsoyiannis, D.: Stochastic simulation of hydrosystems, in: Water Encyclopedia, Vol. 4, Surface and Agricultural Water, edited by: Lehr, J. H. and Keeley, J., Wiley, New York, 421–430, 2005.

Kuczera, G. and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, Water Resour. Res., 34(6), 1481–1489, 1998.

Nalbantis, I., and Rozos, E.: A system for the simulation of the hydrological cycle in the Boeoticos Kephisos basin, Modernisation of the supervision and management of the water resource system of Athens, Project report 10, Dept. of Water Resources, National Technical University of Athens, 2000.

Nalbantis, I., Rozos, E., Tentes, G., Efstratiadis, A., and Koutsoyiannis, D.: Integrating groundwater models within a decision support system, Proc. 5th Intern. Conf. of European Water Resources Association: "Water Resources Management in the Era of Transition", edited by: Tsakiris, G., Athens, 4–8 September 2002, EWRA, IAHR, 279–286, 2002.

Panday, S. and Huyakorn, P. S.: A fully coupled physically-based spatially distributed model for evaluating surface/subsurface flow, Adv. Water. Resour., 27, 361–382, 2004.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, 1997.

Rosberg, D. and Madsen, H.: Concepts of hydrologic modelling, in: Encyclopedia of Hydrological Sciences, Chapter 10, John Wiley & Sons, 2005.

Rozos, E. and Koutsoyiannis, D.: A multicell karstic aquifer model with alternative flow equations, J. Hydrol., 325(1–4), 340–355, 2006.

Rozos, E. and Koutsoyiannis, D.: Error analysis of a multi-cell groundwater model, J. Hydrol., 392(1–2), 22–30, 2010.

Rozos, E., Efstratiadis, A., Nalbantis, I., and Koutsoyiannis, D.: Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows, Hydrolog. Sci. J., 49(5), 819–842, 2004.

Savenije, H. H. G.: *HESS Opinions* "The art of hydrology", Hydrol. Earth Syst. Sci., 13, 157–161, doi:10.5194/hess-13-157-2009, 2009.

Savenije, H. H. G.: *HESS Opinions* "Topography driven conceptual modelling (FLEX-Topo)", Hydrol. Earth Syst. Sci. Discuss., 7, 4635–4656, doi:10.5194/hessd-7-4635-2010, 2010.

Seibert, J.: Reliability of model predictions outside calibration conditions, Nord. Hydrol., 34, 477–492, 2003.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration, Water Resour. Res., 38(11), 1241, doi:10.1029/2001WR000978, 2002.

Schoups, G., Lee Addams, C., and Gorelick, S. M.: Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico, Hydrol. Earth Syst. Sci., 9, 549–568, doi:10.5194/hess-9-549-2005, 2005.

Singh, V. and Bhallamudi, S. M.: Conjunctive surface-subsurface modeling of overland flow, Adv. Water Resour., 21, 567–579, 1998.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, Hydrolog. Sci. J., 48(6), 857–880, 2003.

Sivapalan, M., Zhang, L., Vertessy, R., and Blöschl, G.: Downward approach to hydrological prediction, Hydrol. Process., 17, 2101–2111, 2003.

Smith, M. B., Georgakakos, K. P., and Liang, X.: The distributed model intercomparison project (DMIP), J. Hydrol., 298(1–4), 1–335, 2004.

Srinivasan, R., Muttiah, R. S., Dyke, P. T., Walker, C., and Arnold, J.: Hydrologic unit model for the United States (HUMUS), Texas Agricultural Experiment Station, Blackland Research Center, Temple, TX, 2000.

Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H. H. G., Ayalew, S., Temesgen, M., and Wenninger, J.: Water balance modeling of Upper Blue Nile catchments using a top-down approach, Hydrol. Earth Syst. Sci. Discuss., 7, 6851–6886, doi:10.5194/hessd-7-6851-2010, 2010.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, Hydrol. Earth Syst. Sci., 5, 13–26, doi:10.5194/hess-5-13-2001, 2001.

Waterloo Hydrogeologic Inc.: Modflow packages – Reference Manual, 1999.

Zehe, E., Lee, H., and Sivapalan, M.: Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models, Hydrol. Earth Syst. Sci., 10, 981–996, doi:10.5194/hess-10-981-2006, 2006.

Zhang, G. P. and Savenije, H. H. G.: Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach, Hydrol. Earth Syst. Sci., 9, 243–261, doi:10.5194/hess-9-243-2005, 2005.

HESSD

7, 8265–8308, 2010

Holistic versus monomeric strategies

I. Nalbantis et al.

Table 1. Features of test modelling frameworks in regard to key modelling options.

Key modelling option	Modelling framework A	Modelling framework B
SW-GW	The surface hydrology model is separate from the groundwater flow model (MODFLOW).	Surface and groundwater models are integrated within a single computer package.
SW-GW-WM	No water management model is considered; groundwater abstractions are set equal to the theoretical water requirements.	A water management model, accounting for alternative sources and demand priorities through an optimization framework, is integrated with hydrological models.
SCALE-PARAM	The groundwater flow model (MODFLOW) is implemented as a fully distributed physics-based model, whereas the infiltration model used is semi-lumped.	A semi-distributed approach is used (combination of sub-basins and HRUs for surface water processes and a small number of non-rectangular cells for the aquifer).
SCHEM-PARAM	Parameterization of surface processes is simple due to the elementary model; for groundwater flow processes parameterization follows schematization; zonation is possible.	The HRU concept helps to decouple schematization and parameterization of the surface hydrology model. For the groundwater model decoupling is possible by grouping parameter values of several groundwater cells (zonation).
OPT	MODFLOW includes a module for automatic parameter optimization, which implements a deterministic local-search method without any possibility of intervention to guide solution. Parameter fitting is merely based on water table observations and cannot take into account other types of data (e.g. timing of interruption of spring hydrographs).	Calibration follows a hybrid (i.e., manual-automatic) procedure, where multiple criteria are embedded to control unmeasured responses and take advantage of the physical interpretation of parameters. The sparse water table measurements are taken into account to identify rising or falling trends in ground-water level. The global (i.e., evolutionary) character of the optimization method, supported by manual interventions during calibration, ensures finding a consistent solution.

HESSD
7, 8265–8308, 2010

Table 2. Coefficients of efficiency between computed and observed monthly flows for the period of calibration (October 1984–September 1990) and validation (October 1990–September 1994).

Monthly flow	Calibration		Validation	
of spring	Framework A	Framework B	Framework A	Framework B
Mavroneri	0.428	0.748	0.105	0.720
Melas	−1.712	0.251	−0.890	0.141
Polygyra	−1.245	0.193	Lack of data	Lack of data

HESSD
7, 8265–8308, 2010

Table 3. Bias in mean monthly flow (observed minus computed) in m^3/s for the period of calibration (October 1984–September 1990) and validation (October 1990–September 1994).

Monthly flow	Calibration		Validation	
of spring	Framework A	Framework B	Framework A	Framework B
Mavroneri	0.393	−0.117	0.250	0.172
Melas	0.494	−0.002	0.298	−0.001
Polygyra	0.151	0.011	Lack of data	Lack of data

HESSD
7, 8265–8308, 2010

Table 4. Comments on the effectiveness of the alternative modelling strategies as reflected in the research results of this work.

Key modelling option	Modelling framework A (strategy A)	Modelling framework B (strategy B)
SW-GW	The surface hydrology model proved too simplistic to feed MODFLOW with reliable inputs. Separate calibration of MODFLOW led to poor predictive capacity.	Integrating surface and groundwater models allowed for simultaneous calibration against basin and spring hydrographs within a single computer program.
SW-GW-WM	The absence of a water management model and the use of rough estimates of withdrawals produced errors due to drastic assumptions (satisfaction of water demand, time averaged values).	Model integration allowed for optimizing dynamic withdrawals and allocating targets fulfilled via different sources, which helped to improve overall model performance.
SCALE-PARAM	The coarse scale of the infiltration model decreased the value of the detailed information provided by MODFLOW.	Scale compatibility was guaranteed between surface and groundwater processes whereas respecting the principle of parsimony. The delineation of the aquifer to 40 cells (in contrast to the 3631 cells of strategy A) dramatically decreased the time of simulations.
SCHEM-PARAM	Surface processes were parameterized per sub-catchment as homogeneous areal units, i.e., system schematization dictated parameterization. Zonation was applied in groundwater flow modelling.	The use of HRUs helped decouple schematization and parameterization of the surface hydrology model. For the groundwater model, decoupling proved possible through parameter grouping, on the basis of both topographical and geological criteria (zonation).
OPT	The manual calibration was a tedious procedure. The model performance was much worse in calibration and rather unrealistic in stochastic simulation mode. The deterministic optimization of the local-search type used in MODFLOW certainly lies behind modern optimization methods.	Calibration was effectively guided towards a best compromise solution through proper formulation of the optimization problem, as explained in Table 1. Accounting for both the reproduction of observed data and the fluctuation of the groundwater levels ensured realistic responses against the two water management scenarios.

HESSD
7, 8265–8308, 2010

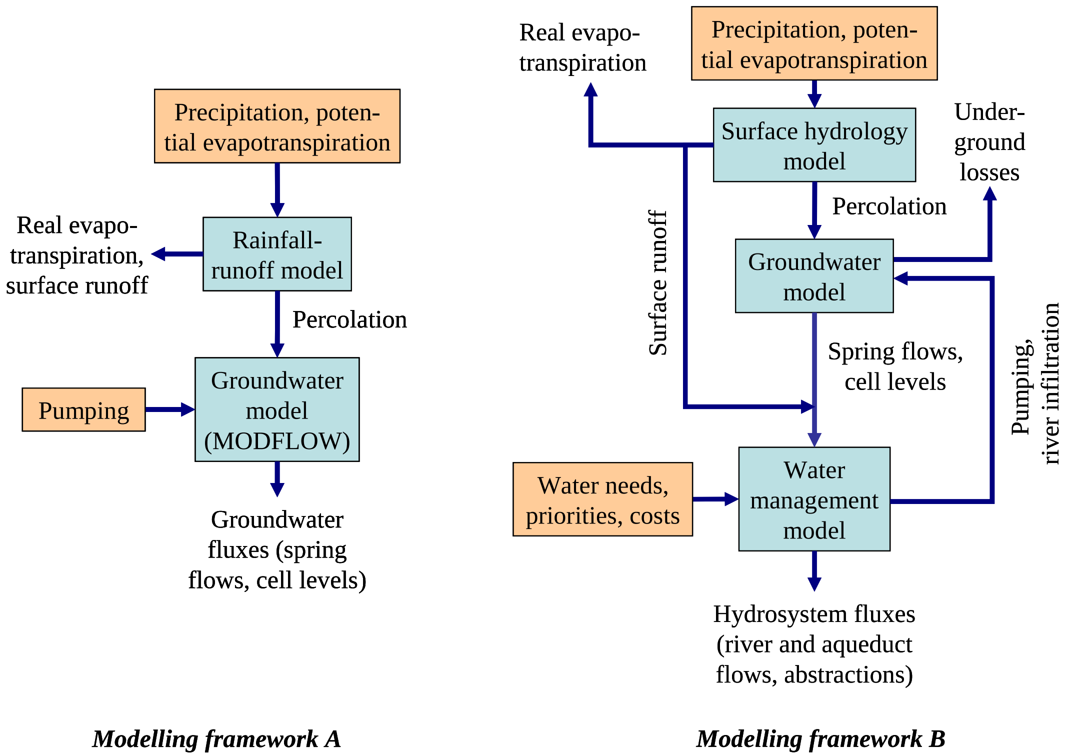


Fig. 1. Synoptic sketch of the two modelling frameworks.

HESSD
7, 8265–8308, 2010

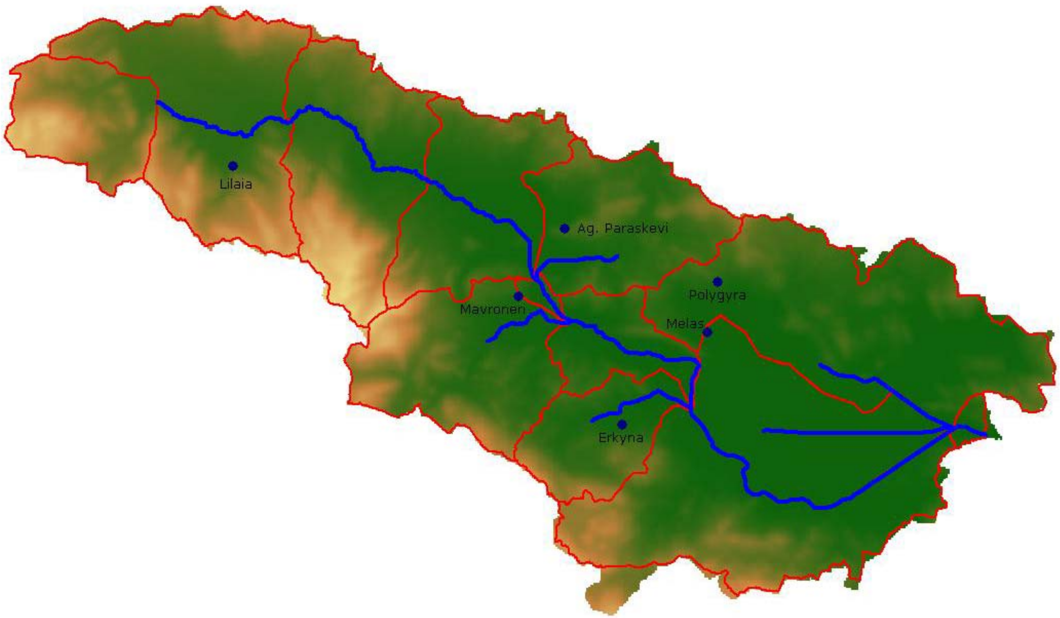


Fig. 2. The Boeoticos Kephisos river basin and the main hydrosystem components (sub-basins, river network, springs), according to the schematization of modelling framework B.

HESSD
7, 8265–8308, 2010

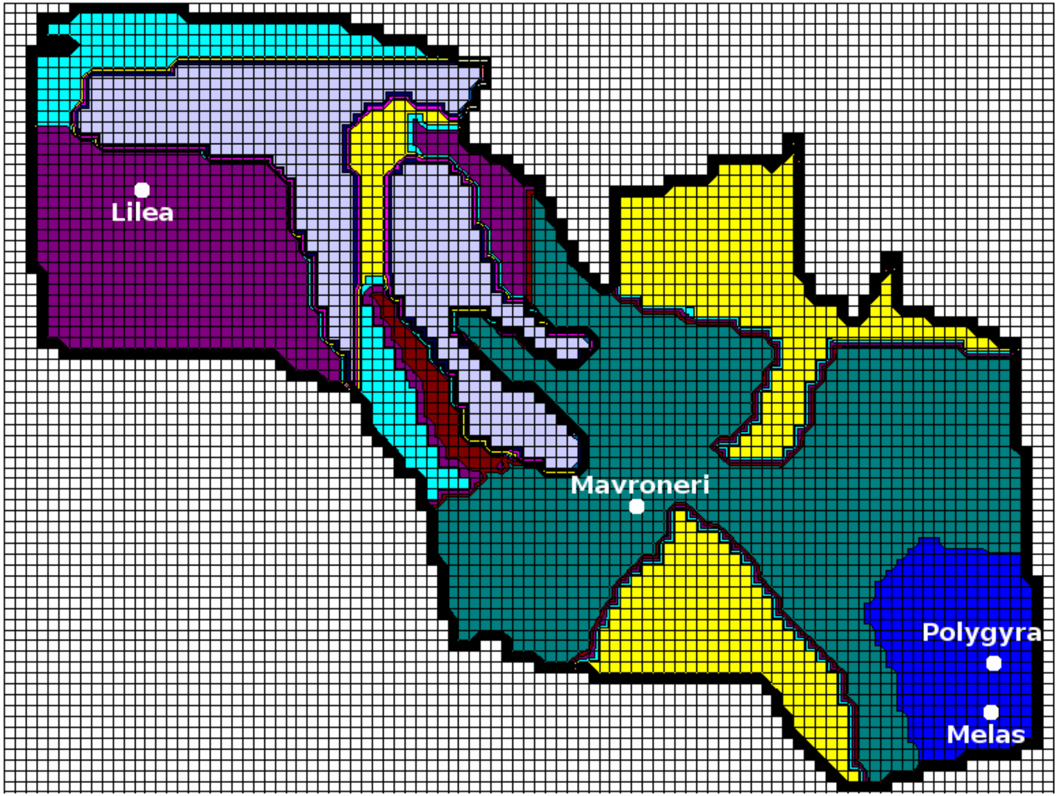


Fig. 3. The discretization of the karst aquifer (also indicating the springs) and the zonation approach (with zones in different colours), according to modelling framework A.

HESSD
7, 8265–8308, 2010

Holistic versus monomeric strategies

I. Nalbantis et al.

Fig. 4. The discretization of the entire groundwater system (also indicating the springs and the four dummy cells, accounting for underground losses) and the zonation approach, according to modelling framework B.

HESSD
7, 8265–8308, 2010

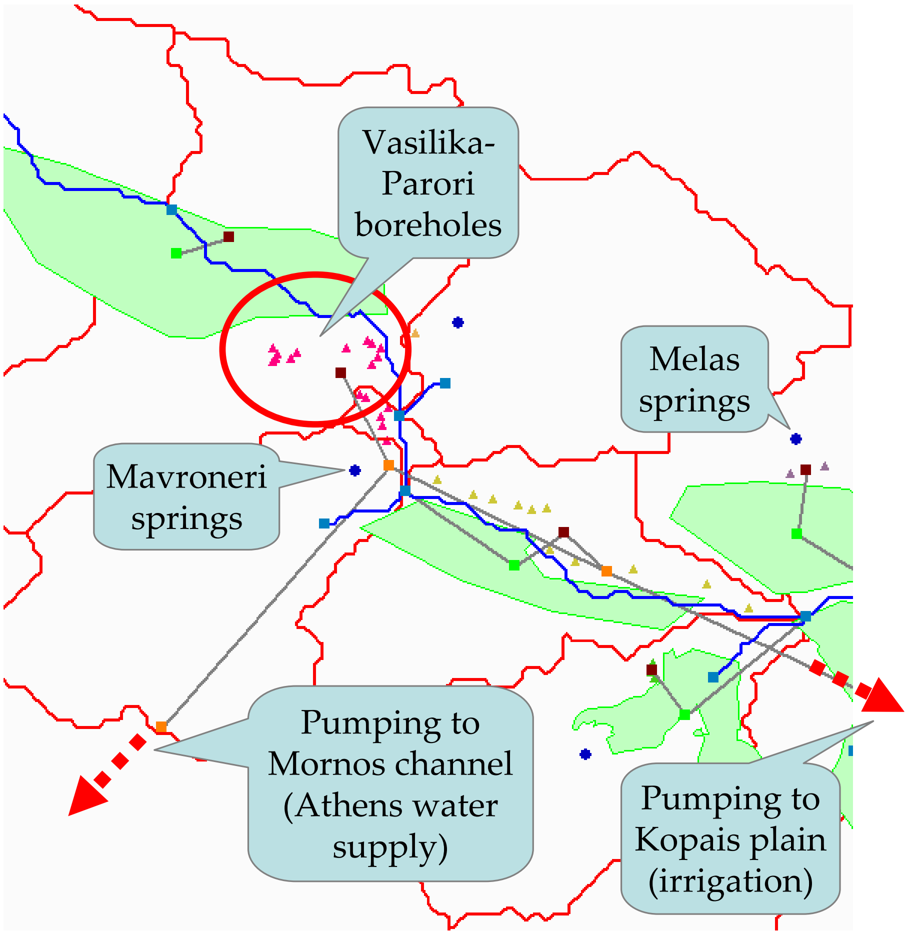


Fig. 5. A detailed depiction of the water management network in the middle part of the basin, according to modelling framework B.

HESSD
7, 8265–8308, 2010

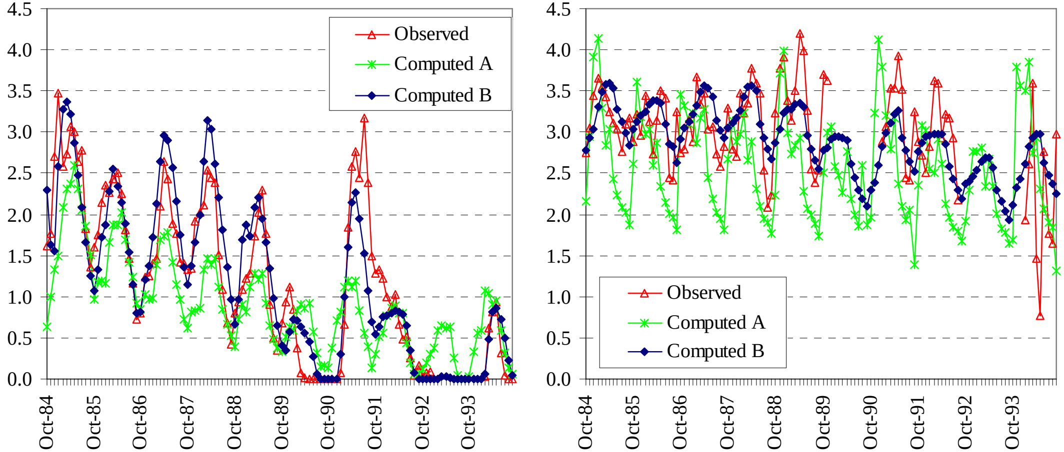


Fig. 6. Computed and observed discharge (m^3/s) at Mavroneri (left) and Melas (right) springs, for modelling frameworks A and B.

HESSD
7, 8265–8308, 2010

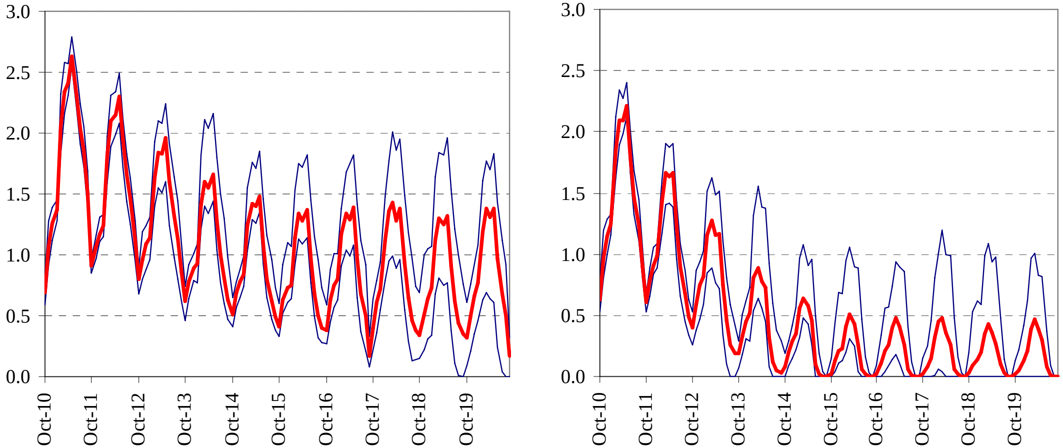


Fig. 7. Simulated discharge (m^3/s) at Mavroneri springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework A.

HESSD
7, 8265–8308, 2010

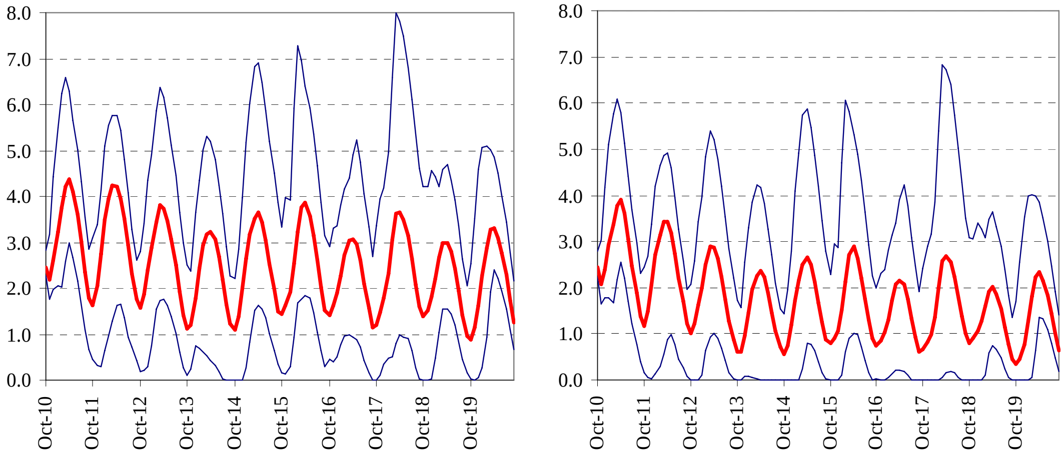


Fig. 8. Simulated discharge (m^3/s) at Mavroneri springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework B.

HESSD
7, 8265–8308, 2010

Holistic versus monomeric strategies

I. Nalbantis et al.

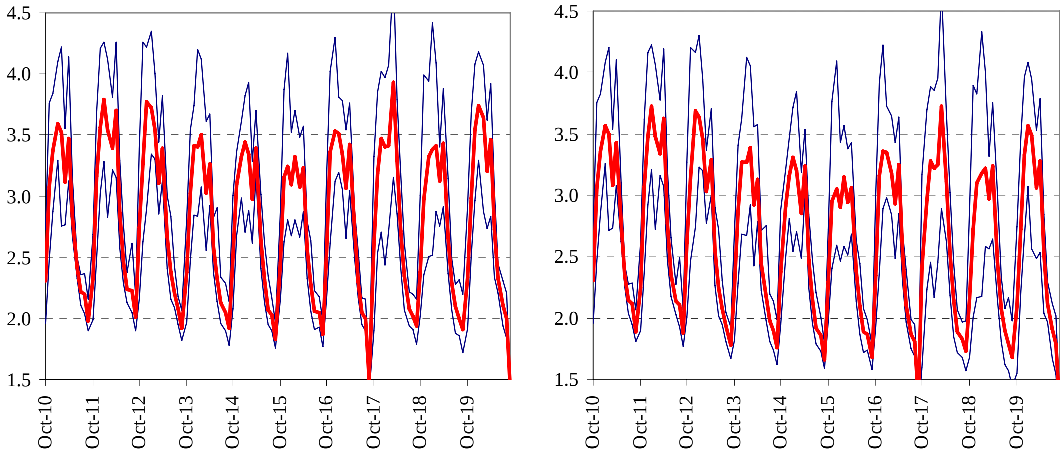


Fig. 9. Simulated discharge (m^3/s) at Melas springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework A.

HESSD
7, 8265–8308, 2010

Holistic versus monomeric strategies

I. Nalbantis et al.

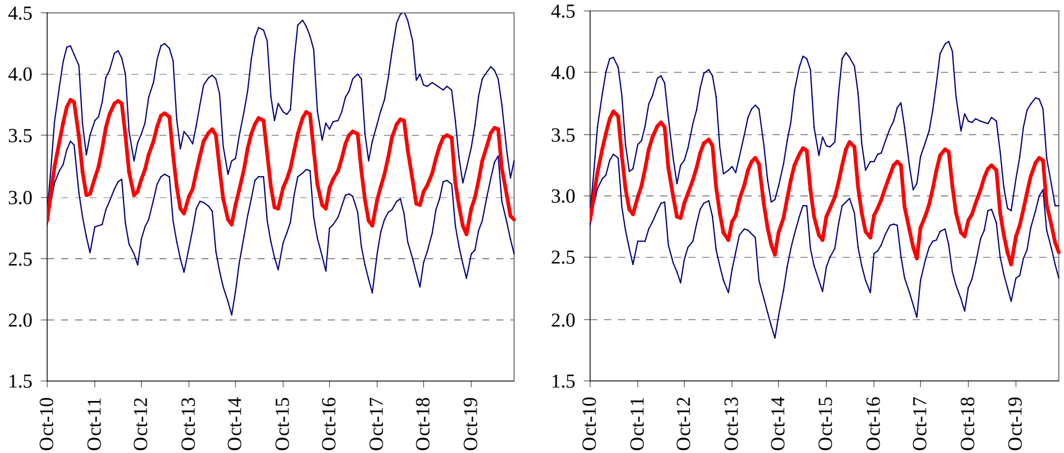


Fig. 10. Simulated discharge (m^3/s) at Melas springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework B.

HESSD
7, 8265–8308, 2010

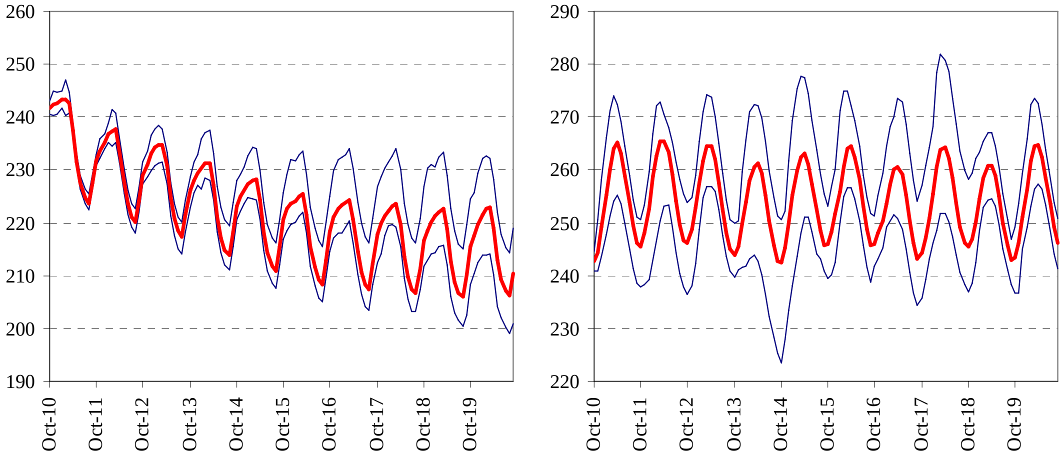


Fig. 11. Simulated level (m) at the upstream part of the aquifer (mean in red and 80% prediction limits in thin blue) under zero pumping for the water supply of Athens, according to modelling frameworks A (left) and B (right).

HESSD

7, 8265–8308, 2010

Holistic versus monomeric strategies

I. Nalbantis et al.

Hydrol. Earth Syst. Sci. Discuss., 7, 8265–8308, 2010
www.hydrol-earth-syst-sci-discuss.net/7/8265/2010/
doi:10.5194/hessd-7-8265-2010

Hydrology and Earth System Sciences Discussions

Holistic versus monomeric strategies for hydrological modelling of modified hydrosystems

I. Nalbantis[1], A. Efstratiadis[2], E. Rozos[2], M. Kopsiafti[2], and D. Koutsoyiannis[2]

[1]Laboratory of Reclamation Works and Water Resources Management, School of Rural and Surveying Engineering, National Technical University of Athens, Athens, Greece
[2]Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Athens, Greece

Received: 30 September 2010 – Accepted: 5 October 2010 – Published: 20 October 2010

Correspondence to: I. Nalbantis (nalbant@central.ntua.gr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The modelling of modified basins that are inadequately measured constitutes a challenge for hydrological science. Often, models for such systems are detailed and hydraulics-based for only one part of the system while for other parts oversimplified models or rough assumptions are used. This is typically a bottom-up approach, which seeks to exploit knowledge of hydrological processes at the micro-scale at some components of the system. Also, it is a monomeric approach in two ways: first, essential interactions among system components may be poorly represented or even omitted; second, differences in the level of detail of process representation can lead to uncontrolled errors. Additionally, the calibration procedure merely accounts for the reproduction of the observed responses using typical fitting criteria. The paper aims to raise some critical issues, regarding the entire modelling approach for such hydrosystems. For this, two alternative modelling strategies are examined that reflect two modelling approaches or philosophies: a dominant bottom-up approach, which is also monomeric and very often, based on output information and a top-down and holistic approach based on generalized information. Critical options are examined, which codify the differences between the two strategies: the representation of surface, groundwater and water management processes, the schematization and parameterization concepts and the parameter estimation methodology. The first strategy is based on stand-alone models for surface and groundwater processes and for water management, which are employed sequentially. For each model, a different (detailed or coarse) parameterization is used, which is dictated by the hydrosystem schematization. The second strategy involves model integration for all processes, parsimonious parameterization and hybrid manual-automatic parameter optimization based on multiple objectives. A test case is examined in a hydrosystem in Greece with high complexities, such as extended surface-groundwater interactions, ill-defined boundaries, sinks to the sea and anthropogenic intervention with unmeasured abstractions both from surface and groundwater. Criteria for comparison are the physical consistency of parameters, the reproduction

of runoff hydrographs at multiple sites within the basin, the likelihood of uncontrolled model outputs, the required amount of computational effort and the performance within a stochastic simulation setting.

1 Introduction and motivation

Two different general approaches or philosophies are applied in modelling of natural processes at large scales (in the order of at least a few km^2): The first approach, called bottom-up or upward (BU), seeks to exploit knowledge (typically physical laws) at the micro-scale (in the order of a few m^2) and then proceeds to larger scales through spatial aggregation. The second approach, called top-down or downward (TD), examines processes directly at the large scale and then eventually proceeds to making inferences about processes at smaller scales (Klemeš, 1983; Sivapalan et al., 2003b). Apart from this categorization of modelling approaches whose criterion is the base spatial scale of process representation, another categorization arises when the criterion is the level of modelling detail: Very often, some parts of the studied system are modelled in detail (in space-time) while for other parts simplified models are employed; in that way essential interactions among system components may be poorly represented or even omitted; we will call this approach monomeric (M), which originates from the Greek words "*μόνος*" and "*μέρος*" respectively denoting "solely" and "part". Conversely, when all parts of the studied system are modelled in the same detail and are linked via feedback mechanisms the approach will be called holistic (H) from the Greek word "*όλον*", which means "whole". The distinction between the monomeric and the holistic character of the approach can be extended to the type of information about the system that is exploited in modelling; in a monomeric approach often limited information is used, which encompasses a small number of measured system outputs or responses whereas in an holistic approach often one seeks to exploit information that is more general. Apart from observations, any qualitative information about the system

responses also can be taken into account together with the empirical interpretation of the unmeasured variables (whether output or internal ones).

Focusing on hydrological modelling, Savenije (2009) pointed out that "*... the dominant paradigm of hydraulics is reductionism, or a bottom-up approach, whereas in hydrology it is (or should be) empiricism and a top-down approach looking for links with fundamental laws of physics.*" The implementation of the BU approach into hydrology has led to modelling of hydrological processes at the small scale (e.g., local, plot or hillslope), which has been an active research area in recent years (Zhang and Savenije, 2005; Zehe et al., 2006; Bárdossy, 2007). The practical usefulness of such models lies in allowing hydrological predictions at the catchment scale, supposedly without using any information on hydrological responses (Kilsby et al., 1999). Essentially this was the initial central focus of the "Predictions in Ungauged Basins (PUB)" initiative (Sivapalan et al., 2003a).

Critiques on the fundamental limitations of this approach, promising substantial reduction of uncertainty through reduction (i.e., theoretical explanation of small-scale processes) rather than deduction (i.e., explanation based on "lumped" response data) have appeared recently (Koutsoyiannis et al., 2009), but the underlying idea has been criticised from its early steps (Beven, 1989). Savenije (2009) reports examples where the BU approach fails while taking a broader perspective of the system under study through a top-down (TD) approach manages to better explain reality. Applications of the latter approach, which is rather macroscopic and, in this sense, holistic, are few (e.g., Tekleab, 2010).

The problems of the bottom-up approach become apparent when hydrological models are called to support engineering and management decisions, i.e., meet their major role (Efstratiadis and Mamassis, 2009). Supporting of decisions often requires modelling hydrosystems that involve extended surface-groundwater interactions and extended anthropogenic interventions in the hydrological cycle, such as abstractions from surface water bodies, pumping, and returns through artificial drainage systems. Theoretically, applying the BU approach for such modified hydrosystems would necessitate putting together all physical processes and process interactions. Obviously, this would require a tremendous amount of information, which is absent in every real-world application. What is very frequently encountered within the BU approach is the monomeric character of modelling as this is defined earlier in this section. For example, a very detailed model is often formulated for one part of the system (or subsystem), while using oversimplified models for other parts or even ignoring dynamic links between subsystems. More often than not, the focus is on the detailed hydraulic model of a specific subsystem, such as an aquifer. According to the two categorization criteria the approach will be characterized as bottom-up/monomeric (BU-M). Although one may say in advance that this is simply a bad modelling practice, which merits no further study, the use of such approach is still so widespread that analysis of its implications is, to our view, justified.

We will concentrate our effort to modelling of complex, human-modified hydrosystems, which is a practical problem of high interest. To represent the BU-M approach we will consider a particular modelling strategy, called here strategy A. This focuses on hydraulic modelling of one natural subsystem only, which is the basin aquifer. To cope with the system complexities, a multi-stage modelling process is used that involves five stages: (1) splitting the hydrosystem into a number of natural sub-systems (sub-basins and the aquifer) and one man-made sub-system; (2) modelling natural sub-systems individually; (3) transferring predictions from the natural sub-systems to the man-made sub-system; and (4) optimizing the operation of the latter sub-system to represent as close as possible the observed conditions of the past (calibration). This is typically the strategy followed in engineering studies with the aid of popular commercial computer packages for water resources management. It presupposes that: (a) pure natural sub-systems can be effectively found, and (b) sufficient information is available for each sub-system modelled.

Strategy A may lead to erroneous predictions in complex basins where no simple natural sub-systems can be identified due to complex water exchanges. Moreover, inadequate information on some sub-systems may prohibit successful modelling. Data

inadequacy involves, among others, ill-defined system boundaries due to unknown leakages, sinks to the sea and anthropogenic intervention with unmeasured abstractions both from surface and groundwater. In the last years some approaches have appeared that cope with some of the above problems, although they do not cover the case of modified basins (e.g., Panday and Huyacorn, 2004; Bari and Smettem, 2004; Singh and Bhallamudi, 1998; Gauthier et al., 2009) neither do they treat the case with unknown abstractions (e.g., Schoups et al., 2005).

An alternative modelling strategy, called here strategy B, will be used to represent a top-down/holistic approach. The hydrosystem is viewed as a whole, having the input and the required information as guides to formulate spatial modelling units and process models. Ultimately, this approach leads to model integration, parsimonious parameterization and simultaneous optimization of all model parameters. All these provide flexibility to strategy B, which may be critical in cases with modified but poorly measured hydrosystems.

The motivation for this work is to test the applicability of modelling strategy A when the latter is used for modified hydrosystems. The target is precisely to examine the every-day modelling strategies in a critical spirit. It is the effects of these strategies that are investigated and not the value of the models used therein. In this respect, our work differs from the few comparative studies reported in the literature, such as the distributed model intercomparison project (Smith et al., 2004). The potential benefit when the problems of strategy A are faced is evaluated through applying strategy B. To achieve this, we extended the typical split-sample procedure for model building (i.e., calibration/validation based on historical data) to examining also the system response under hypothetical future conditions, using synthetically-generated forcing (stochastic simulation). This offers advantages over the typical model validation procedure: (1) it helps testing a modelling strategy within a framework that is similar to an operational one; (2) it can help examining future water management scenarios that are different from the historical ones; (3) it provides an opportunity to check for unreasonable long-term statistical trends or jumps of any model variable, which tests model credibility (e.g., the model should normally generate stationary outputs, if fed with stationary forcing); (4) it can provide estimates of the uncertainty of model predictions against synthetic forcing data (e.g., precipitation, temperature).

To implement the two strategies, two modelling frameworks were chosen. The choice merely reflects the authors' experience on specific models. Modelling framework A implements strategy A and is based on the well-known groundwater package MODFLOW, coupled with a simple infiltration scheme. Modelling framework B, which is chosen to implement strategy B, uses a recently proposed framework that integrates a semi-distributed rainfall-runoff model, a coarse groundwater model and a network-type water allocation model (Efstratiadis et al., 2008).

A challenging operational case study was chosen involving the Boeoticos Kephisos river basin, Greece. This comprises all complexities described above and has been studied by the authors in the past (Nalbantis and Rozos, 2000; Nalbantis et al., 2002; Rozos et al., 2004; Efstratiadis et al., 2008). All above works present sequentially improved modelling strategies, from relatively simple to more detailed ones, which are consistent with the top-down/holistic type of approach. Taking advantage of this effort, we detected and investigated five key modelling options within the selected modelling strategies, which are discussed hereinafter.

2 Key modelling options in hydrological modelling strategies

When formulating a modelling strategy, critical decisions are made in regard to selecting, formulating and fitting hydrological models. These decisions lead to defining key modelling options that constitute the "ingredients" of the formulated strategy; these options are described next.

2.1 Key modelling option SW-GW: link between models for surface and groundwater processes

In strategy A, different stand-alone models are used for surface and groundwater hydrology, which precludes accounting for feedback interaction between the corresponding processes. Very often, the two models differ in the spatial and temporal scale used. They may also differ in their modelling approach or philosophy. For, the surface processes are usually represented via conceptual (hydrological) approaches; even the fully-distributed physically-based schemes are considered conceptual at the grid scale (Beven, 1989). Yet, groundwater modelling typically follows a hydraulic rather than a hydrologic rationale. All these aspects affect the parameter estimation procedure, which requires either to provide unrealistic simplifications (e.g., assume that the entire runoff is derived from the surface system, thus omitting the contribution of groundwater runoff) or (rarely) make successive approximations, i.e., calibrate the one component after the other, which is computationally inefficient. In strategy B, the main hydrological interactions are explicitly represented, and thus model parameters can be simultaneously optimized, taking advantage of the available measurements across both components (e.g., flow and piezometric data).

2.2 Key modelling option SW-GW-WM: link between models for hydrological processes and water management

In the staged modelling procedure of strategy A, hydrological models are constructed exclusively for undisturbed parts of the system (e.g., sub-basins) and the outputs thereof (e.g. river flows) are transferred as inputs to the water management model of the man-made sub-system (usually implemented within a Decision Support System or DSS). This serial operation, apart from being computationally inefficient, suffers from a number of drawbacks: (a) it is unrealistic when decision-related interactions between hydrological and man-made sub-systems are significant; (b) it is infeasible when real abstractions are unknown, since it precludes the assessment of boundary conditions for hydrological and hydrogeological models; (c) it may impose serious limitations to the representation of the water allocation problem (e.g., a "first-come, first-served" management policy may be mandatory); (d) it requires data transfer between models, which puts stringent requirements regarding space-time scale compatibility of hydrological model outputs and inputs to the water management model; (e) it precludes automatic calibration of models, which, in presence of the above problems, is at least questionable (Efstratiadis et al., 2010). Attempts to cope with the above problem are rare in literature (Fredericks et al., 1998; Dai and Labadie, 2001). Strategy B adopts model integration, which copes with the problem.

2.3 Key modelling option SCALE-PARAM: link of spatial scale and model parameterization

Since parameterization is designed to represent factors that influence the spatial variability of hydrological processes, it is naturally linked to the spatial scale (Klemeš, 1983). The large heterogeneity of mechanisms and properties makes it difficult to achieve compatibility between measurements made at the local scale and model predictions. Quite often, in strategy A, very detailed models are chosen in the hope to achieve scale compatibility between data and predicted variables. Yet, the resulting high dimensionality leads to extremely time-consuming schemes, which is a major restricting factor affecting not only calibration but also the operational applicability of models; the latter arises because models have to co-operate with DSS that run in forecast mode, using synthetic forcings for long time horizons (Nalbantis et al., 2002).

2.4 Key modelling option SCHEM-PARAM: link between hydrosystem schematization and parameterization

Inevitably, in strategy A, the hydrosystem schematization (i.e., the simplification of the process representation in space) dictates parameterization. Parameters are assigned to individual spatial elements (e.g., sub-basins, grid cells), thus having limited

physical meaning. Moreover, due to the detailed spatial scale adopted, the principle of parsimony is broken, which results to poorly identified parameters (Kuczera and Mroczkowski, 1998) and increased predictive uncertainty (Efstratiadis and Koutsoyiannis, 2010; Savenije, 2010). Attempts to reduce the number of control variables of the optimization problem require hybridized strategies, such as detecting only the most important parameters while estimating the rest of them on the basis of field data (Refsgaard, 1997; Eckhardt and Arnold, 2001) or using zonation approaches (i.e. spatial grouping of parameters). Contrary to the above, in strategy B schematization and parameterization are disconnected, thus ensuring that models are by construction parsimonious. In this approach, the schematization is adapted to the engineering objectives (i.e., which processes should be simulated and where), while the parameterization is only linked to the available information (cf. Dehotin and Braud, 2008).

2.5 Key modelling option OPT: appropriate use of optimization in calibration

In theory, physically-based approaches enable their free variables to be derived from field measurements. Yet, in practice, their applicability is significantly restrained not only by the heterogeneity of processes and the unknown scale dependencies of parameters (Beven, 2001; Wagener et al., 2001; Rosberg and Madsen, 2005), but also by the high computational effort and the subsequent inability to co-operate with DSSs (Nalbantis et al., 2002). Hence, optimization is always required for at least some of model parameters (Refsgaard, 1997; Eckhardt and Arnold, 2001). However, within strategy A, non-expert users frequently adopt a "black-box" approach, seeking for a "global optimal" parameter set (typically against a single, quantitative performance criterion), through an automatic algorithmic procedure. Very often, this leads to: (a) parameter values that are inconsistent with their physical interpretation; (b) poor predictive capacity against independent control data (validation); (c) unreasonable values of uncontrolled responses (e.g., evapotranspiration, underground losses) and internal model variables (e.g., soil and groundwater storage) (Rozos et al., 2004; Efstratiadis et al., 2008). On the other hand, strategy B emphasizes less on the optimization task and more on the comprehensive understanding of the problem components (real system, model and data), to ensure reliable results. Particularly in models of complex parameterization, it aims to increase the information that is exploited in calibration as much as possible, by means of both multi-response measurements and empirical metrics ("soft" data), which account for the hydrological expertise (Efstratiadis and Koutsoyiannis, 2010).

3 Overview of alternative modelling frameworks

3.1 Modelling framework A

Within this framework an off-line coupling of MODFLOW is made with a simple rainfall-runoff model, as illustrated in Fig. 1, left. MODFLOW is a classical tool for 3-D simulation of groundwater flow, where the flow field is discretized into a number of rectangular cells and all quantities are referred to cell centres. The 3-D continuity equation and the Darcy's law written in finite differences form provide one final flow equation for each cell as a function of the unknown hydraulic head h and other known variables. The latter are either parameters of the aquifer (conductances for the three axis directions and the specific yield) or external stresses of the cell (percolation from soil or the river bed, pumped water, water outflow to the sea). After defining initial and boundary conditions, the final system of equations on h is solved via the Preconditioned Conjugate Gradient method.

For model implementation, we used the computer package by Waterloo Hydrogeologic Inc. (1999). This includes an optimization module, employing a deterministic local-search type method for the automatic calibration of model parameters.

The rainfall-runoff process is represented through two modelling components; the first calculates the effective rainfall on the basis of precipitation and potential evapotranspiration, whereas the second divides the effective rainfall into runoff and percolation, assuming a constant ratio between these two variables, which stands as the single

model parameter. The above scheme runs independently, to provide external stresses to MODFLOW cells (due to percolation).

3.2 Modelling framework B

In modelling framework B the computer package HYDROGEIOS 2.0 is used. This is a GIS-based model suite, which allows for flexible representation of hydrological processes, as summarized below and more analytically described by Efstratiadis et al. (2008). A synoptic description of the modelling framework is illustrated in Fig. 1, right. More precisely, surface flow is considered within the hydrographic network, which is extracted from a digital terrain model through adjusting the flow accumulation parameter and adding control points that correspond to flow measurement stations or diversion nodes. The network conveys flow of sub-basins, which are subject to different hydrological stresses (precipitation and potential evapotranspiration). Surface processes are considered homogeneous within partitioned areas of the basin termed as the Hydrological Response Units (HRUs). This idea has found a limited number of applications to distributed modelling also (Flügel, 1995; Srinivasan et al., 2000; Gurtz et al., 2003). The HRUs represent soil and land use types within portions (patches) of the studied area, which are not necessarily contiguous. They are defined as the product of partitions of the basin on the basis of different properties, such as soil permeability, land cover and topography (i.e. terrain slope). Through classification of the above properties, one can adjust the number of HRUs and, consequently, the number of the parameters of the surface hydrological model; the latter comprises six parameters per HRU and transforms precipitation into real evapotranspiration, percolation and surface runoff (direct, overland, lateral) for each sub-basin and HRU combination.

Groundwater flow is modelled through a multi-cell approach involving discretization of the aquifer into non-rectangular cells (Rozos and Koutsoyiannis, 2006, 2010). This allows the description of complex geometries on the basis of the physical characteristics of the aquifer (e.g., geology), through parsimonious structures. Two parameters are assigned to each cell (conductivity and specific yield). Springs and underground losses are implemented as virtual cells of very large base, which allows keeping almost constant hydraulic head regardless storage variations. Model stresses are: (a) areal inflows due to percolation through each sub-basin and HRU combination; (b) inflows due to infiltration underneath each river segment; (c) outflows due to pumping from each borehole.

Regulated flow through man-made structures and portions of the hydrographic network is modelled with the aid of a water management network, which is an extension of the scheme described by Efstratiadis et al. (2004). This has as nodal inflows the surface and the groundwater runoff, as nodal outflows the withdrawals for water uses, and as distributed fluxes the water losses due to infiltration and river discharge. Major hydraulic works are also represented along with the corresponding water uses and constraints as well as their interactions with the natural system. Model properties are discharge and pumping capacities, target priorities, demand time series and unit transportation costs of water. The allocation of flows is based on a linear programming approach where virtual unit costs, positive or negative, are assigned either to prohibit undesirable fluxes or to force the model fulfil the hydrosystem targets (Efstratiadis et al., 2010).

HYDROGEIOS embeds an advanced calibration module that provides a number of statistical and empirical criteria for model fitting on multiple responses (river and spring discharge, hydraulic heads) and various options regarding the delineation of the feasible search space. Optimization is carried out through the evolutionary annealing-simplex method (Efstratiadis and Koutsoyiannis, 2002; Rozos et al., 2004).

4 Case study

4.1 The study area

The Boeoticos Kephisos river basin lies on the Eastern Sterea Hellas, to the north of Athens, and drains a closed area of 1930 km^2 (Fig. 2). The catchment is formed on

heavily karstified limestone covered with alluvial deposits in plain areas. Considerable groundwater amount (more than half of the annual catchment runoff) is discharged through large springs in the upper and middle part of the basin, whereas an unknown amount is leaking to the sea.

The surface runoff reaching the basin outlet is conducted to the neighbouring Lake Hylike, which is one of the major reservoirs of Athens water supply system. Groundwater of the middle part is also considered as an emergency resource. Finally, a significant part of the surface and groundwater resources of the basin is used in agriculture; more specifically, river abstractions are mainly implemented in the lower part of the basin, practically eliminating flow availability during the summer months. For a more detailed description of the basin features, the reader may refer to previous publications (Rozos et al., 2004; Efstratiadis et al., 2008).

For the study area, a major question is about the impacts of abstractions though the Vasilika-Parori boreholes to the overall hydrological regime of the basin. These boreholes were drilled within the frame of emergency measures taken during a severe drought in the period from 1989 to 1994, at the end of which almost all surface resources dried out. Yet, due to the considerable reduction of precipitation and the intense pumping for providing drinking water to Athens, the discharge of the neighbouring Mavroneri springs was interrupted twice during 1990 and 1993, thus resulting to various social and environmental problems (these springs account for 15% of the total basin runoff). In the last 15 years, several engineering studies were carried out to investigate the complex dynamics of the hydrosystem and provide reliable estimations regarding the consequences of water supply abstractions to the system responses. Modelling framework A originates from an earlier engineering study (Nalbantis and Rozos, 2000) while framework B capitalizes the experience gained after continuous research attempts (Nalbantis et al., 2002; Rozos et al., 2004; Efstratiadis et al., 2008).

4.2 Model schematization and parameterization

Within strategy A, the karst aquifer is considered as a single layer with free surface flow and cell size varying from $800 \times 800\,\mathrm{m}^2$ near the boundaries to $150 \times 150\,\mathrm{m}^2$ in central areas (near the springs); this resulted to a total number of cells equal to 3631. To reduce the number of control variables, the flow field was divided into zones (Fig. 3), which resulted to totally 18 parameters to calibrate. The alluvial aquifer was not simulated since its water yield is low and of limited interest. In the whole perimeter of the karst aquifer no-flow boundaries were assigned, apart from two areas: a small area in SW, where lateral recharge was considered using dummy recharge wells, and the NE boundary, where small-scale outflows to the sea were simulated through a cluster of pumping wells. Unsteady flow was assumed whereas external stresses due to percolation were modelled in a simple way as explained in Sect. 3.1. Precipitation was considered homogenous in the interior of three sub-basins, which represent different hydrogeologic units (low, middle and upper course). Infiltration depth was taken as a constant fraction of the effective precipitation, which varied from one type of surface geological formation to another; three types were assumed, corresponding to limestone, alluvial and flysch (impermeable) formations. Stream-aquifer interactions were represented using the River module of MODFLOW assuming zero infiltration during the April–August period. No water management model was explicitly considered; to estimate unmeasured withdrawals from groundwater; the irrigated area that is supplied by each well component and the related water demand per unit area were used by assuming (wrongly) that the entire demand is fulfilled via pumping (Nalbantis and Rozos, 2000; Nalbantis et al., 2002).

Strategy B followed a semi-distributed schematization to 15 sub-basins, as shown in Fig. 2. The spatial average of precipitation of each sub-basin was calculated through the Thiessen method, whereas potential evapotranspiration series was estimated via the Penman-Monteith method, by adjusting air temperatures to the mean elevation

of each sub-basin. For the definition of the HRUs, three categories of permeability (low, medium, high) and two categories of terrain slope were taken. The groundwater flow field was discretized into 40 non-rectangular cells, four of which represent underground leakages to neighbouring basins and the sea. In addition, six dummy cells were used to model surface outflows through the major karstic springs (Fig. 4). This spatial discretization is two orders of magnitude coarser than the one implemented within strategy A, thus making the computational effort for groundwater simulation almost negligible, in comparison to that of strategy A. For the parameterization of the groundwater flow field, we used three categories of permeability and porosity, which reflect topography and geology. Moreover, we used particular permeability values for the rest of cells representing springs and underground losses (ten parameters in total). In combination with the parameterization of the surface processes via the six HRUs, the total number of model parameters was 52 (36 for the surface model and 16 for the groundwater model). We note that both the schematization and parameterization differ from those reported by Efstratiadis et al. (2008), in an attempt to significantly reduce the number of groundwater parameters; this ensures a more parsimonious modelling approach (Kopsiafti, 2009).

The water management network includes consumption nodes for agricultural areas, abstraction nodes (river diversion works and borehole groups) and aqueducts. Water management policy was implemented through assigning water supply targets to a number of consumption nodes (seven for irrigation and one for water supply) and virtual costs to the system aqueducts with the purpose to represent a realistic and close to the actual abstraction priority for water uses (e.g., in case of combined abstractions, priority was given to river abstractions instead of pumping, which is the historical practice). This is a critical difference between the two strategies, since strategy A requires known abstractions whereas strategy B is much more flexible, since it is based on theoretical demands, costs and priorities, thus allowing to choose among different abstraction policies to fulfil demands. For instance, pumped water through Vasilika-Parori boreholes is either conducted downstream, for the irrigation of Kopais plain, or diverted

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

for the water supply of Athens. Withdrawals from other resources, both surface and groundwater, are also implemented to serve local agricultural demands (Fig. 5).

4.3 Calibration strategies and data for parameter estimation

Both strategies were tested against the observed hydrological responses for a 10-year control horizon (October 1984–September 1994), employing a monthly time step. For this period, discharge series at seven locations are available, precisely at the basin outlet (Karditsa tunnel) and downstream of the six karst springs, illustrated in Fig. 2. With regard to groundwater, several level gauges were available, mostly located in the vicinity of the main river branch.

Regarding the rainfall-runoff component of the framework A, no attempt was made for parameter optimization, since infiltration fractions were estimated empirically, on the basis of the main geological formations of the basin. The MODFLOW parameters were manually optimized on the basis of 18 observed level series and through visual inspection of the closeness of observed and simulated spring hydrographs.

In strategy B, as thoroughly explained by Efstratiadis et al. (2008), the parameters of the surface and the groundwater models were simultaneously calibrated. Since the number of parameters was large, to cope with the resulted uncertainty which is directly related to equifinality (Freer et al., 1996), multiple criteria were taken into account, including "soft" data (Seibert and McDonnell, 2002; Efstratiadis and Koutsoyiannis, 2010). Thus, a weighted objective function was formulated comprising the following statistical and empirical measures: (a) efficiency and bias of the monthly hydrographs at the seven locations mentioned above; (b) penalties for not reproducing flow intermittency; and (c) penalties for generation of unrealistic trends regarding the groundwater levels. The first group of criteria accounts for the so-called "hard data" (measurements); this is essential for reproducing the global water balance and the spring mechanisms, but not sufficient for representing the regime of the groundwater fluxes; these criteria will be later referred to as performance indices. The second group simply accounts for the information "zero" or "non-zero discharge", which is easily observable, reliable and

is of major interest in water management. Finally, the third group of criteria is a kind of "soft" data, ensuring reasonable fluctuation of the non-observable internal variables of the model. Optimization was carried out through a hybrid strategy, which combines human experience and automatic tools (Boyle et al., 2000; Rozos et al., 2004). In that manner, search was guided towards a realistic, best-compromise parameter set, ensuring satisfactory predictive capacity for all model responses.

4.4 Operational use of models through stochastic simulation

As mentioned in the Introduction, calibration is essentially an intermediate step in the modelling procedure, which allows for optimising the predictive capacity of the model on the basis of observed data. Yet, the reproduction of the past responses has limited interest, if not accompanied by further analyses with "projected" inputs, thus providing support to decisions for future. In this respect, the two strategies are evaluated within a stochastic (Monte Carlo) simulation framework, aiming to examine the system response under different stress conditions, comprising both natural (precipitation, potential evapotranspiration) and anthropogenic (abstractions from surface and groundwater recourses) forcing.

For the representation of rainfall, a multivariate stochastic scheme was used to generate point series of 1000-year length, which preserve the essential statistical characteristics of the observed samples of 12 rain gauges across the basin, at the annual and monthly time scales (Koutsoyiannis et al., 2003). Next, the point series were aggregated to the appropriate spatial scale, thus providing areal rainfall series for the 3 and 15 sub-basins, which correspond to modelling frameworks A and B respectively.

The synthetic rainfall records were divided into clusters of ten-year length, to formulate a hundred of statistically equivalent forcing scenarios. Each modelling scheme ran 100 times, each one under different stochastic forcing, whereas for all runs the same initial conditions were supplied. For instance, for modelling framework B we used the soil moisture depths and groundwater levels of the beginning of calibration (October 1984). The model outputs within this configuration represent statistically consistent trajectories of the system responses, for a ten-year horizon. Such type of Monte Carlo analysis is typically used in operational applications (e.g., forecasts), and is also known as terminating simulation (Koutsoyiannis et al., 2003; Koutsoyiannis, 2005).

Regarding the potential evapotranspiration throughout each sub-basin, we assigned the same areal values with the control period 1984–1994, since this is a forcing variable of very low interannual variability. Finally, for the anthropogenic forcing, we examined two alternative water management scenarios, to evaluate the impacts of water supply abstractions to the system responses (especially in the middle course of the basin). For both policies, we assigned the actual agricultural demands across the basin (223 hm^3/yr) and assumed either zero or extensive (46 hm^3/yr, equal to that of the water year 1993–1994) demand through the Vassilika-Parori boreholes, for providing drinking water to Athens (Fig. 5). We remind that in modelling strategy A, the water requirements are by definition fulfilled via pumping (which is an erroneous yet obligatory assumption), while in strategy B the demands can be satisfied through multiple sources, following the cost optimization approach.

5 Results

5.1 The testing framework

Ideally, testing the effect of adopting modelling strategy A for modified hydrosystems, would require a complex computer experiment based on a series of alternative frameworks implementing strategy A, each one differing from framework B in only one key modelling option. Careful examination of the commonly available models shows that it is almost impossible to build frameworks that perfectly fulfil the above requirements. So, we simplified our experiment by considering the two frameworks presented in Sect. 3. Thus, the combined effect of all key modelling options is tested. Table 1 summarises the a priori knowledge about the relation of each key modelling option to each modelling framework. Application to a common data set of hydrological variables of the

aforementioned test basin allowed an objective evaluation of the effect of adopting strategy A. For this, both numerical and empirical criteria are used for the comparison of the two approaches, in calibration and simulation mode.

We emphasize on the monthly flows at control points (i.e., the main karstic springs), whose hydrographs are also depicted for visual comparison. Comparisons on the river flow at the basin outlet were not feasible since framework A did not provide such output. Regarding the groundwater levels, we only examined their long-term behaviour rather than their actual values, since a direct comparison of aquifer water levels was meaningless, due to enormous differences of scale between the two alternative groundwater models used.

5.2 Comparison of model performance in calibration

The model performance during the calibration and the validation periods is evaluated on the basis of two criteria, efficiency values and bias in the mean. Due to the different assumptions regarding the system delineation (e.g., in strategy A, the river network and the alluvial areas of the aquifer were not simulated), comparisons were possible only at three observation sites, namely downstream of Mavroneri, Melas and Polygyra springs. In Tables 2 and 3, values of the corresponding performance indices are provided, which show a clear improvement of model performance in spring flow predictions when passing from strategy A to strategy B for both the calibration and the validation periods. This is confirmed by hydrograph comparisons on Fig. 6 for the two most important springs of the basin (Mavroneri and Melas). We note that the first approach, although concentrated on the detailed representation of the aquifer dynamics, fails to reproduce the key characteristics of the observed flows, namely the monthly variability, which is overestimated for Melas and underestimated for Mavroneri springs; it also fails to reproduce the interruption of the discharge of the latter, during 1990 and 1993.

On the other hand, the advantages of model integration as offered by strategy B are not restricted to flow predictions only. Other improvements are equally significant, which are commented in Table 4 and provide explanation of the superior performance

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

indices regarding observed groundwater variables, despite the fact that framework B follows a much simpler modelling approach, which focuses on the surface processes and the water management practices.

5.3 Comparison of model performance in simulation

Obviously, when comparing two modelling approaches in a stochastic simulation setting, it is impossible to use quantitative criteria (e.g., goodness-of-fit measures), as in calibration. Therefore, the evaluations are based on the grounds of common sense, i.e. testing whether the model provides the right answers for the right reasons (cf. Kirchner, 2006), taking advantage of the hydrological experience.

The implementation of the two frameworks under synthetically generated inputs further reveals the drawbacks of strategy A. In Fig. 7 we plot the projected discharge at Mavroneri springs for a ten-year horizon (mean value of 100 flow scenarios and 80% prediction limits), under zero and intensive pumping through the neighbouring boreholes at Vasilika-Parori (respectively referred to as "actual abstraction policy" and "intensive abstraction policy"). For both management policies, there is a sharp decrease of discharge, which is inconsistent with the experience so far. Although one could expect that under an extensive pumping a systematic negative trend could be possible, it is unlikely that such trend is encountered under the actual abstraction policy. The differences with the respective results obtained through the modelling approach B (Fig. 8) are substantial; here, the mean projection for the spring outflows (which corresponds to average rainfall conditions) follows a stationary pattern under the actual abstraction policy, while there is a progressive decrease of the spring resources under the intensive abstraction policy (Fig. 8). This indicates that, in a long-term perspective, the intensive use of the Vasilika-Parori boreholes for the water supply of Athens is not a sustainable option.

Regarding the response of the Melas springs according to modelling approach A, there are negligible differences between the two abstraction policies (Fig. 9). This is also an unexpected result, since the entire karst aquifer should be affected by the water

supply abstractions, also because of the fact that the irrigation needs in the downstream part of the basin should be fulfilled though increased pumping in the vicinity of the aforementioned springs. Yet, this presupposes a proper description of the water management alternatives, which is only feasible thorough framework B. Following strategy B, it is projected that Melas springs are obviously affected by the upstream abstractions, and thus the slightly negative trend of the average flow trajectory is reasonable (Fig. 10).

Attempting to investigate the reasons for the unrealistic performance of modelling framework A, we concluded that we should revisit the calibration procedure. In contrast to strategy B, where we accounted for the internal variables of the model by assigning trend penalties on groundwater levels, in strategy A, we just attempted to fit the model responses (spring flows) on the observations. In that manner, we allowed systematic drains and fillings of the cells lying near the boundaries of the aquifer where spring flow observations are not available. Actually, we left calibration to assign unrealistic conductivity values in order to maximize the model performance against the efficiency indices. On the other hand, within framework B soft criteria are also employed but this required a more time-consuming and hard to automate calibration strategy. In Fig. 11 we compare the synthetically generated levels obtained through the two approaches, which correspond to the most upstream part of the karst system, assuming the actual abstraction policy. We observe a questionable behaviour (i.e., negative trend) of the projected level when applying modelling framework A, whereas for framework B the groundwater system exhibits stationary behaviour, which is more reasonable.

6 Conclusions

Our investigations have shown that in watersheds that are modified by human interventions the classical modelling strategy based on the monomeric bottom-up approach is, in general, inefficient. It makes use of a detailed hydraulic model for only a part of the studied system and of separate models for surface hydrological processes,

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

groundwater flow processes and water management processes. Such serial use of models prohibits the modelling of the process interactions and suffers from increased computational burden. The monomeric character of this approach reflects other modelling aspects, including the assignment of parameters (coincidence of spatial scales of schematization and parameterization). The calibration procedure is based only on measured system outputs; moreover, fully automatic search is employed and internal variable dynamics is ignored. All the above misuse practices are reflected to the predictive capacity of the model, which proved disappointedly poor for such an exhaustive effort (in calibration). This behaviour was identified by employing the model in stochastic simulation mode for operational use; the obtained projections are far from being realistic.

Conversely, a holistic top-down approach allows for model schematization and parameterization that respects the principle of parsimony and ensures computational efficiency by means of both simulation and optimization/calibration. This precludes taking advantage of the power of fully distributed models while, at the same time, favouring the use of the semi-distributed approach. An effective way to reduce the size of the parameter set is to decouple parameterization and model schematization through using Hydrological Response Units (HRUs). Through the HRU concept, the model structure depends on a limited number of landscape classes, whose parameters retain some physical consistency thus allowing for a better identification of their prior uncertainty (cf. Savenije, 2010). Last, model integration allows simultaneous calibration of all models through exploiting all kinds of information and not only information about some basic output variables. A hybrid process of manual and automatic optimization proved very effective in finding a best compromise solution while, at the same time, respecting some physical interpretation of parameters. In this approach all available pieces of information, including hydrological experience, are exploited in model calibration within a multiobjective framework.

Our tests proved that running models in stochastic simulation mode can be a useful tool for their testing and validation since this augments information supplied by typical

calibration and validation procedures, while, at the same time, addressing some of the intrinsic restrictions of the available data (Kirchner, 2006). In fact, the stochastic framework offers new perspectives towards multiple paths. First, it allows for assessing the performance of hydrological models in situations that are consistent to those in which it is supposed to be used in practice (Klemeš, 1986). In addition, it provides a "crash-test" for evaluating the model transposability in time, which is a necessary condition for their operational adequacy (Andréassian et al., 2009). Finally, through a proper representation of the varying character of climate and the related processes in stochastic terms, it ensures reliable estimation of the uncertainty and the long-term risk in hydrological studies and water resources management (Koutsoyiannis et al., 2007; Koutsoyiannis, 2010). This can also include the evaluation of extremes, which are not represented in the (usually limited) calibration data (Seibert, 2003).

We believe that the research presented in this paper can contribute in (1) formulating specifications for model packages applicable to modified basins, and (2) opening new research routes regarding different types of approach followed in hydrological modelling.

Acknowledgements. The second version of the HYDROGEIOS package was carried out within the research project "Development of Database and software applications in a web platform for the National Databank for Hydrological and Meteorological Information". The software is freely available at http://www.hydroscope.gr/software/hydrogeios.htm.

References

Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M.-H., and Valéry, A.: *HESS Opinions* "Crash tests for a standardized evaluation of hydrological models", Hydrol. Earth Syst. Sci., 13, 1757–1764, doi:10.5194/hess-13-1757-2009, 2009.

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, doi:10.5194/hess-11-703-2007, 2007.

Bari, M. and Smettem, K. R. J.: Modelling monthly runoff generation processes following land use changes: groundwatersurface runoff interactions, Hydrol. Earth Syst. Sci., 8, 903–922, doi:10.5194/hess-8-903-2004, 2004.
Beven, K. J. : Changing ideas in hydrology – The case of physically-based models, J. Hydrol., 105(1–2), 157–172, 1989.
Beven, K.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, doi:10.5194/hess-5-1-2001, 2001.
Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, Water Resour. Res., 36(12), 3663–3674, 2000.
Dai, T. and Labadie, J. W.: River basin network model for integrated water quantity/quality management, ASCE J. Water Res. Plan. Manag., 127(5), 295–305, 2001.
Dehotin, J. and Braud, I.: Which spatial discretization for distributed hydrological models? Proposition of a methodology and illustration for medium to large-scale catchments, Hydrol. Earth Syst. Sci., 12, 769–796, doi:10.5194/hess-12-769-2008, 2008.
Eckhardt, K. and Arnold, J. G.: Automatic calibration of a distributed catchment model, J. Hydrol. 251, 103–109, 2001.
Efstratiadis, A., and Mamassis, N.: Evaluating models or evaluating modelling practices? – Interactive comment on HESS Opinions "Crash tests for a standardized evaluation of hydrological models", Hydrol. Earth Syst. Sci. Disc., 6, C1404–C1409, 2009.
Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E., and Koutsoyiannis, D.: HYDROGEIOS: a semi-distributed GIS-based hydrological model for modified river basins, Hydrol. Earth Syst. Sci., 12, 989–1006, doi:10.5194/hess-12-989-2008, 2008.
Efstratiadis, A. and Koutsoyiannis, D.: An evolutionary annealing-simplex algorithm for global optimisation of water resource systems, Proc. Fifth Intern. Conf. on Hydroinformatics, IWA, Cardiff, UK, 1423–1428, 2002.
Efstratiadis, A. and Koutsoyiannis, D.: One decade of multiobjective calibration approaches in hydrological modelling: a review, Hydrolog. Sci. J., 55(1), 58–78, 2010.
Efstratiadis, A., Koutsoyiannis, D., and Xenos, D.: Minimising water cost in the water resource management of Athens, Urban Water J., 1(1), 3–15, 2004.
Efstratiadis, A., Nalbantis, I., Rozos, E., and Koutsoyiannis, D.: Accounting for water management issues within hydrological simulation: Alternative modelling options and a network optimization approach, EGU General Assembly 2010, Vienna, Austria, 2–7 May 2010, Geophys.

Res. Abstr., 12, EGU2010-10085, 2010.
Flügel, W. A.: Delineating Hydrological Response Units by Geographical Information System analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the river Bröl in Germany, Hydrol. Process., 9(3–4), 423–436, 1995.
Fredericks, J. W., Labadie, J. W., and Altenhofen, J. M.: Decision support system for stream-aquifer management, J. Water Res. Pl.-ASCE, 124(2), 69–78, 1998.
Freer, J., Beven, K. J., and Ambroise, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach, Water Resour. Res. 32(7), 2161–2173, 1996.
Gauthier, M. J., Camporese, M., Rivard, C., Paniconi, C., and Larocque, M.: A modeling study of heterogeneity and surface water-groundwater interactions in the Thomas Brook catchment, Annapolis Valley (Nova Scotia, Canada), Hydrol. Earth Syst. Sci., 13, 1583–1596, doi:10.5194/hess-13-1583-2009, 2009.
Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., and Vitvar, T.: A comparative study in modelling runoff and its components in two mountainous catchments, Hydrol. Process., 17, 297–311, 2003.
Kilsby, C. G., Ewen, J., Sloan, W. T., Burton, A., Fallows, C. S., and O'Connell, P. E.: The UP modelling system for large scale hydrology: simulation of the Arkansas-Red River basin, Hydrol. Earth Syst. Sci., 3, 137–149, doi:10.5194/hess-3-137-1999, 1999
Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006.
Klemeš, V.: Conceptualization and scale in hydrology, J. Hydrol., 65, 1–23, 1983.
Klemeš, V.: Operational testing of hydrologic simulation models, Hydrolog. Sci. J., 31, 13–24, 1986.
Kopsiafti, M.: Investigation of parameterization strategies for the hydrogeological module of Hydrogeios software – Application to the Boeoticos Kephisos basin, Postgraduate thesis, 133 pages, Dept. of Water Resources, National Technical University of Athens, 2009.
Koutsoyiannis, D., Efstratiadis, A., and Georgakakos, K.: Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches, J. Hydrometeorol., 8(3), 261–281, 2007.
Koutsoyiannis, D.: Hurst-Kolmogorov dynamics and uncertainty, J. Amer. Wat. Res. Assoc., in press, 2010.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Koutsoyiannis, D., Karavokiros, G., Efstratiadis, A., Mamassis, N., Koukouvinos, A., and Christofides, A.: A decision support system for the management of the water resource system of Athens, Phys. Chem. Earth, 28(14–15), 599–609, 2003.

Koutsoyiannis, D., Makropoulos, C., Langousis, A., Baki, S., Efstratiadis, A., Christofides, A., Karavokiros, G., and Mamassis, N.: *HESS Opinions*: "Climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability", Hydrol. Earth Syst. Sci., 13, 247–257, doi:10.5194/hess-13-247-2009, 2009.

Koutsoyiannis, D.: Stochastic simulation of hydrosystems, in: Water Encyclopedia, Vol. 4, Surface and Agricultural Water, edited by: Lehr, J. H. and Keeley, J., Wiley, New York, 421–430, 2005.

Kuczera, G. and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, Water Resour. Res., 34(6), 1481–1489, 1998.

Nalbantis, I., and Rozos, E.: A system for the simulation of the hydrological cycle in the Boeoticos Kephisos basin, Modernisation of the supervision and management of the water resource system of Athens, Project report 10, Dept. of Water Resources, National Technical University of Athens, 2000.

Nalbantis, I., Rozos, E., Tentes, G., Efstratiadis, A., and Koutsoyiannis, D.: Integrating groundwater models within a decision support system, Proc. 5th Intern. Conf. of European Water Resources Association: "Water Resources Management in the Era of Transition", edited by: Tsakiris, G., Athens, 4–8 September 2002, EWRA, IAHR, 279–286, 2002.

Panday, S. and Huyakorn, P. S.: A fully coupled physically-based spatially distributed model for evaluating surface/subsurface flow, Adv. Water. Resour., 27, 361–382, 2004.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, 1997.

Rosberg, D. and Madsen, H.: Concepts of hydrologic modelling, in: Encyclopedia of Hydrological Sciences, Chapter 10, John Wiley & Sons, 2005.

Rozos, E. and Koutsoyiannis, D.: A multicell karstic aquifer model with alternative flow equations, J. Hydrol., 325(1–4), 340–355, 2006.

Rozos, E. and Koutsoyiannis, D.: Error analysis of a multi-cell groundwater model, J. Hydrol., 392(1–2), 22–30, 2010.

Rozos, E., Efstratiadis, A., Nalbantis, I., and Koutsoyiannis, D.: Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows, Hydrolog. Sci. J., 49(5), 819–842, 2004.

Savenije, H. H. G.: *HESS Opinions* "The art of hydrology", Hydrol. Earth Syst. Sci., 13, 157–161, doi:10.5194/hess-13-157-2009, 2009.

Savenije, H. H. G.: *HESS Opinions* "Topography driven conceptual modelling (FLEX-Topo)", Hydrol. Earth Syst. Sci. Discuss., 7, 4635–4656, doi:10.5194/hessd-7-4635-2010, 2010.

Seibert, J.: Reliability of model predictions outside calibration conditions, Nord. Hydrol., 34, 477–492, 2003.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration, Water Resour. Res., 38(11), 1241, doi:10.1029/2001WR000978, 2002.

Schoups, G., Lee Addams, C., and Gorelick, S. M.: Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico, Hydrol. Earth Syst. Sci., 9, 549–568, doi:10.5194/hess-9-549-2005, 2005.

Singh, V. and Bhallamudi, S. M.: Conjunctive surface-subsurface modeling of overland flow, Adv. Water Resour., 21, 567–579, 1998.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, Hydrolog. Sci. J., 48(6), 857–880, 2003.

Sivapalan, M., Zhang, L., Vertessy, R., and Blöschl, G.: Downward approach to hydrological prediction, Hydrol. Process., 17, 2101–2111, 2003.

Smith, M. B., Georgakakos, K. P., and Liang, X.: The distributed model intercomparison project (DMIP), J. Hydrol., 298(1–4), 1–335, 2004.

Srinivasan, R., Muttiah, R. S., Dyke, P. T., Walker, C., and Arnold, J.: Hydrologic unit model for the United States (HUMUS), Texas Agricultural Experiment Station, Blackland Research Center, Temple, TX, 2000.

Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H. H. G., Ayalew, S., Temesgen, M., and Wenninger, J.: Water balance modeling of Upper Blue Nile catchments using a top-down approach, Hydrol. Earth Syst. Sci. Discuss., 7, 6851–6886, doi:10.5194/hessd-7-6851-2010, 2010.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, Hydrol. Earth Syst. Sci., 5, 13–26, doi:10.5194/hess-5-13-2001, 2001.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Waterloo Hydrogeologic Inc.: Modflow packages – Reference Manual, 1999.
Zehe, E., Lee, H., and Sivapalan, M.: Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models, Hydrol. Earth Syst. Sci., 10, 981–996, doi:10.5194/hess-10-981-2006, 2006.
Zhang, G. P. and Savenije, H. H. G.: Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach, Hydrol. Earth Syst. Sci., 9, 243–261, doi:10.5194/hess-9-243-2005, 2005.

Table 1. Features of test modelling frameworks in regard to key modelling options.

Key modelling option	Modelling framework A	Modelling framework B
SW-GW	The surface hydrology model is separate from the groundwater flow model (MODFLOW).	Surface and groundwater models are integrated within a single computer package.
SW-GW-WM	No water management model is considered; groundwater abstractions are set equal to the theoretical water requirements.	A water management model, accounting for alternative sources and demand priorities through an optimization framework, is integrated with hydrological models.
SCALE-PARAM	The groundwater flow model (MODFLOW) is implemented as a fully distributed physics-based model, whereas the infiltration model used is semi-lumped.	A semi-distributed approach is used (combination of sub-basins and HRUs for surface water processes and a small number of non-rectangular cells for the aquifer).
SCHEM-PARAM	Parameterization of surface processes is simple due to the elementary model; for groundwater flow processes parameterization follows schematization; zonation is possible.	The HRU concept helps to decouple schematization and parameterization of the surface hydrology model. For the groundwater model decoupling is possible by grouping parameter values of several groundwater cells (zonation).
OPT	MODFLOW includes a module for automatic parameter optimization, which implements a deterministic local-search method without any possibility of intervention to guide solution. Parameter fitting is merely based on water table observations and cannot take into account other types of data (e.g. timing of interruption of spring hydrographs).	Calibration follows a hybrid (i.e., manual-automatic) procedure, where multiple criteria are embedded to control unmeasured responses and take advantage of the physical interpretation of parameters. The sparse water table measurements are taken into account to identify rising or falling trends in ground-water level. The global (i.e., evolutionary) character of the optimization method, supported by manual interventions during calibration, ensures finding a consistent solution.

Table 2. Coefficients of efficiency between computed and observed monthly flows for the period of calibration (October 1984–September 1990) and validation (October 1990–September 1994).

Monthly flow of spring	Calibration Framework A	Calibration Framework B	Validation Framework A	Validation Framework B
Mavroneri	0.428	0.748	0.105	0.720
Melas	−1.712	0.251	−0.890	0.141
Polygyra	−1.245	0.193	Lack of data	Lack of data

Table 3. Bias in mean monthly flow (observed minus computed) in m^3/s for the period of calibration (October 1984–September 1990) and validation (October 1990–September 1994).

Monthly flow of spring	Calibration Framework A	Calibration Framework B	Validation Framework A	Validation Framework B
Mavroneri	0.393	−0.117	0.250	0.172
Melas	0.494	−0.002	0.298	−0.001
Polygyra	0.151	0.011	Lack of data	Lack of data

Table 4. Comments on the effectiveness of the alternative modelling strategies as reflected in the research results of this work.

Key modelling option	Modelling framework A (strategy A)	Modelling framework B (strategy B)
SW-GW	The surface hydrology model proved too simplistic to feed MODFLOW with reliable inputs. Separate calibration of MODFLOW led to poor predictive capacity.	Integrating surface and groundwater models allowed for simultaneous calibration against basin and spring hydrographs within a single computer program.
SW-GW-WM	The absence of a water management model and the use of rough estimates of withdrawals produced errors due to drastic assumptions (satisfaction of water demand, time averaged values).	Model integration allowed for optimizing dynamic withdrawals and allocating targets fulfilled via different sources, which helped to improve overall model performance.
SCALE-PARAM	The coarse scale of the infiltration model decreased the value of the detailed information provided by MODFLOW.	Scale compatibility was guaranteed between surface and groundwater processes whereas respecting the principle of parsimony. The delineation of the aquifer to 40 cells (in contrast to the 3631 cells of strategy A) dramatically decreased the time of simulations.
SCHEM-PARAM	Surface processes were parameterized per sub-catchment as homogeneous areal units, i.e., system schematization dictated parameterization. Zonation was applied in groundwater flow modelling.	The use of HRUs helped decouple schematization and parameterization of the surface hydrology model. For the groundwater model, decoupling proved possible through parameter grouping, on the basis of both topographical and geological criteria (zonation).
OPT	The manual calibration was a tedious procedure. The model performance was much worse in calibration and rather unrealistic in stochastic simulation mode. The deterministic optimization of the local-search type used in MODFLOW certainly lies behind modern optimization methods.	Calibration was effectively guided towards a best compromise solution through proper formulation of the optimization problem, as explained in Table 1. Accounting for both the reproduction of observed data and the fluctuation of the groundwater levels ensured realistic responses against the two water management scenarios.

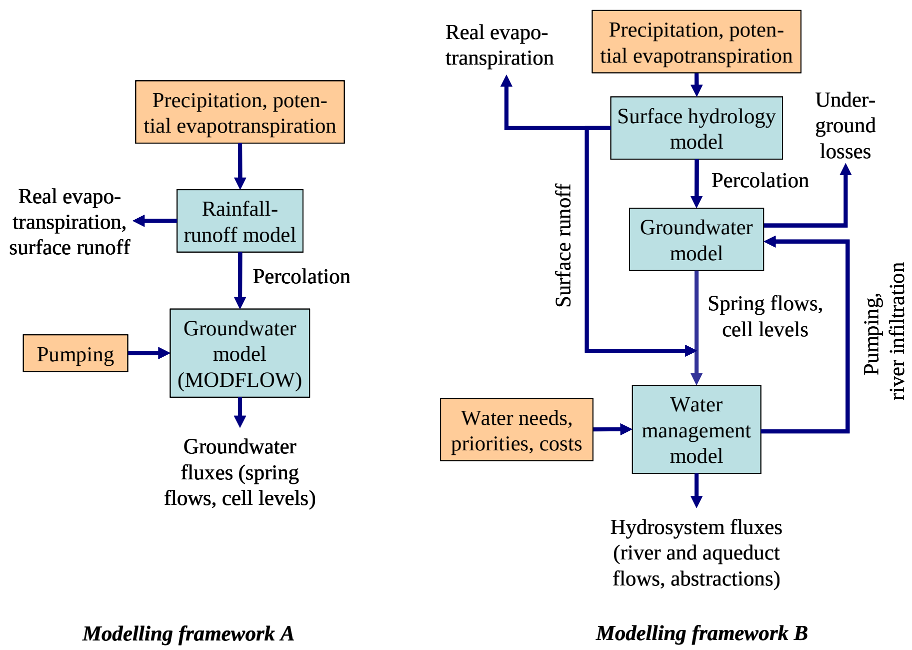


Fig. 1. Synoptic sketch of the two modelling frameworks.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

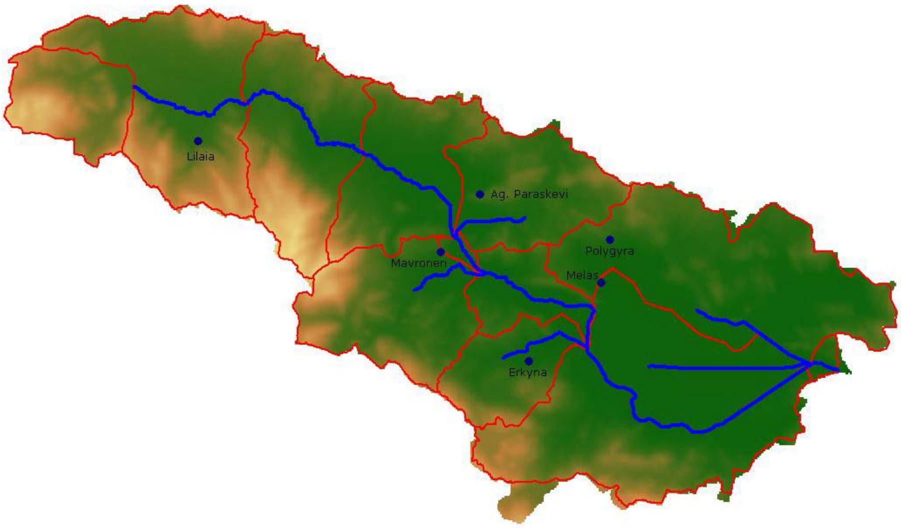


Fig. 2. The Boeoticos Kephisos river basin and the main hydrosystem components (sub-basins, river network, springs), according to the schematization of modelling framework B.

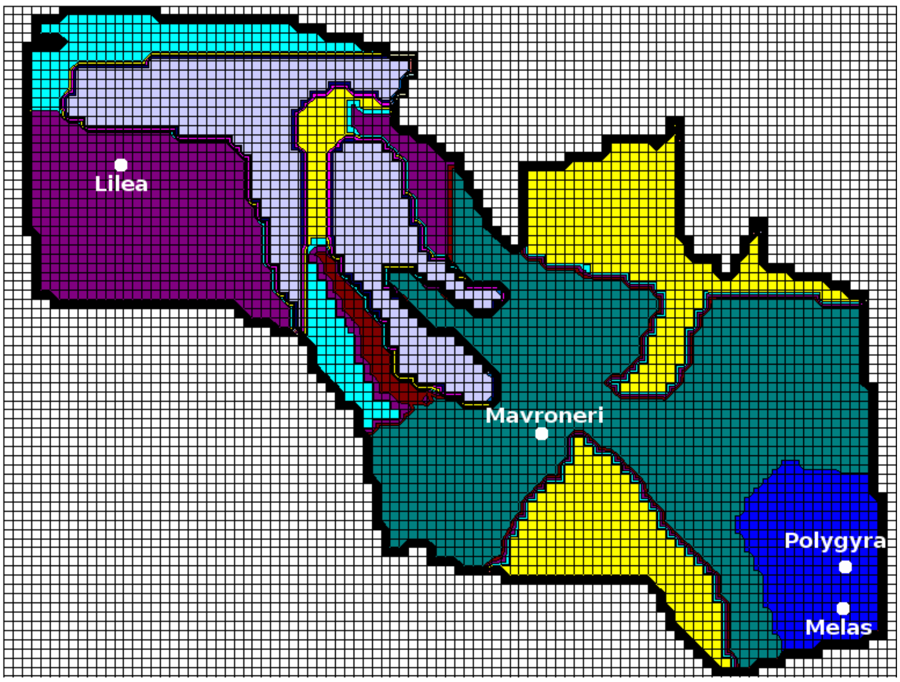


Fig. 3. The discretization of the karst aquifer (also indicating the springs) and the zonation approach (with zones in different colours), according to modelling framework A.

Fig. 4. The discretization of the entire groundwater system (also indicating the springs and the four dummy cells, accounting for underground losses) and the zonation approach, according to modelling framework B.

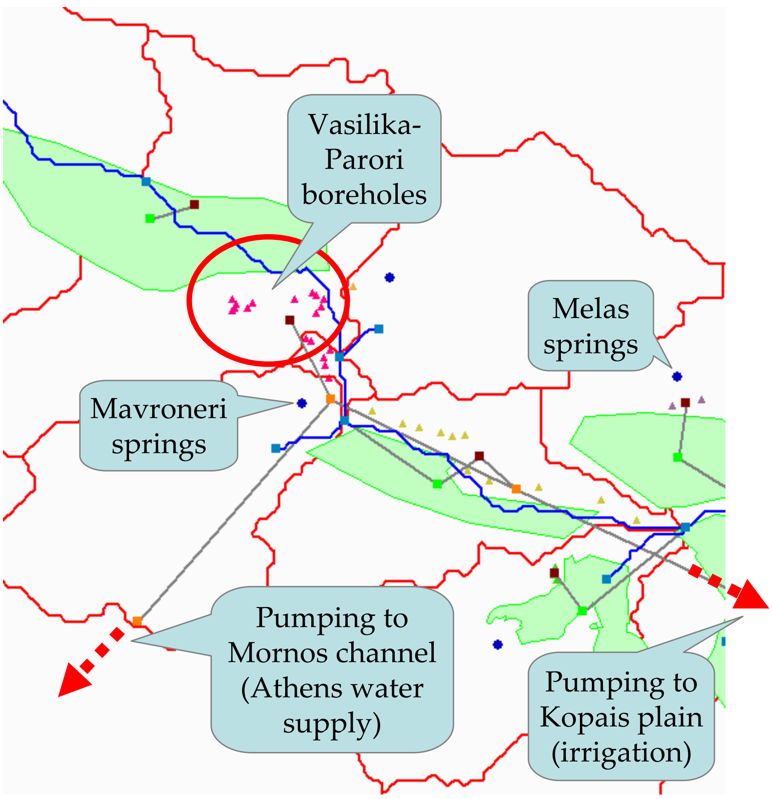
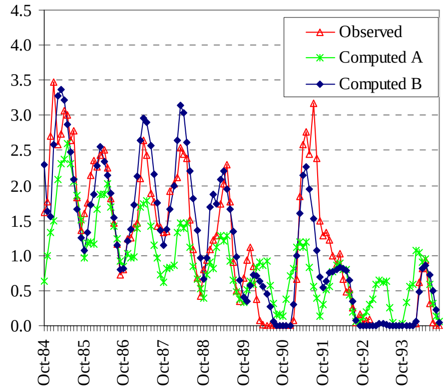


Fig. 5. A detailed depiction of the water management network in the middle part of the basin, according to modelling framework B.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |



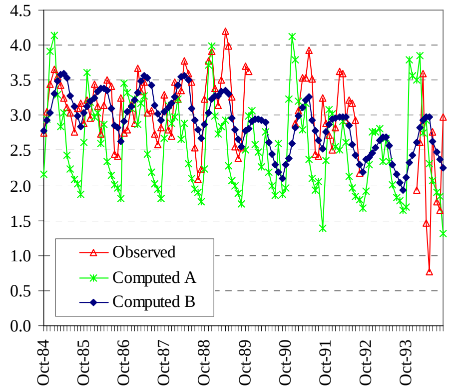
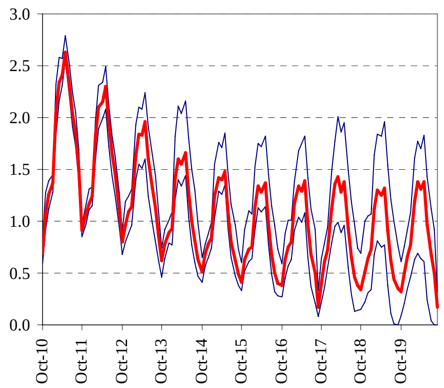


Fig. 6. Computed and observed discharge (m^3/s) at Mavroneri (left) and Melas (right) springs, for modelling frameworks A and B.



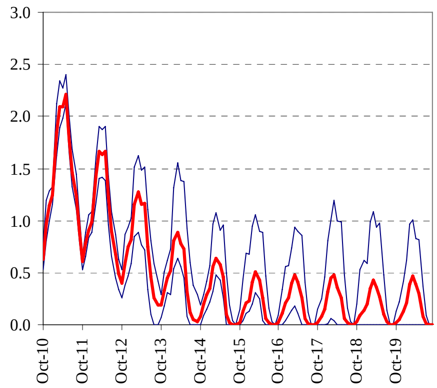


Fig. 7. Simulated discharge (m^3/s) at Mavroneri springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework A.

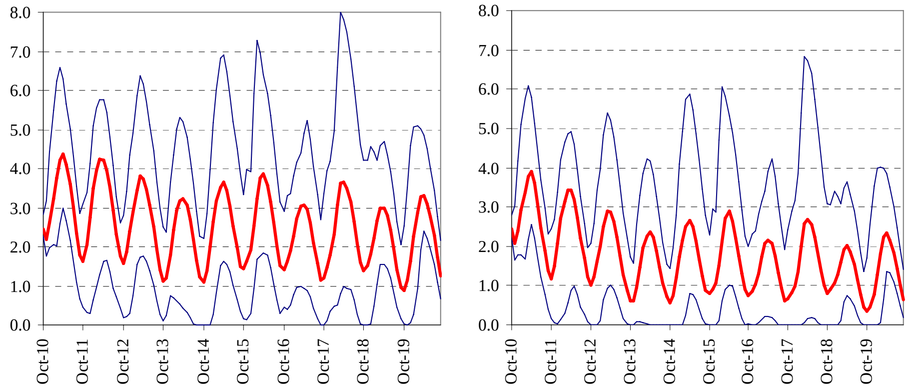


Fig. 8. Simulated discharge (m^3/s) at Mavroneri springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework B.

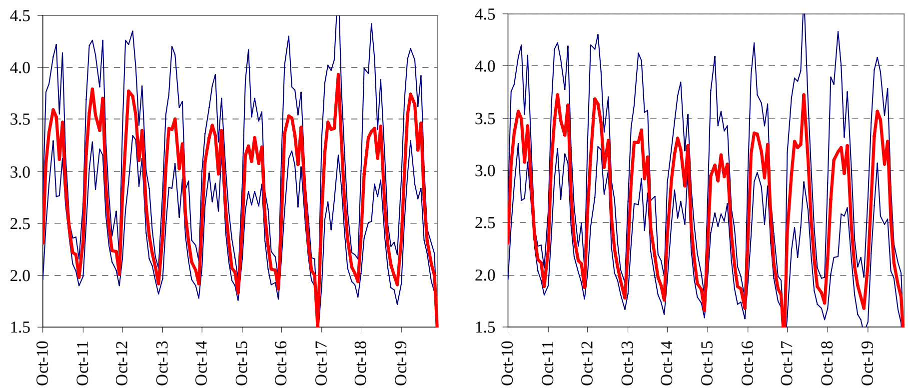


Fig. 9. Simulated discharge (m^3/s) at Melas springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework A.

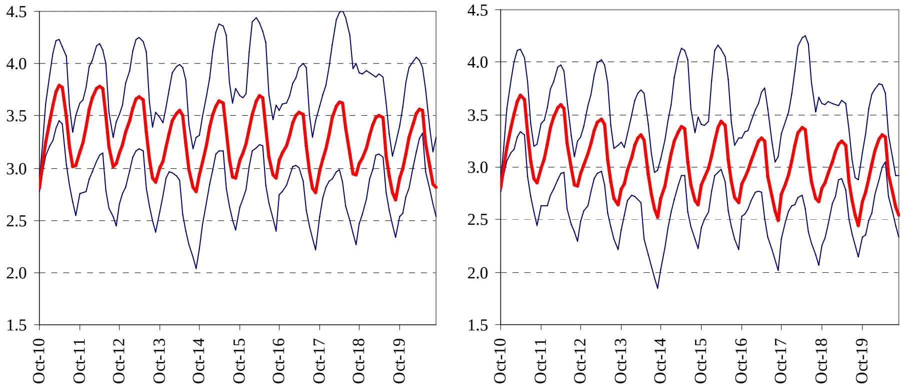


Fig. 10. Simulated discharge (m^3/s) at Melas springs (mean in red and 80% prediction limits in thin blue) under zero (left) and intensive (right) pumping for the water supply of Athens, according to modelling framework B.

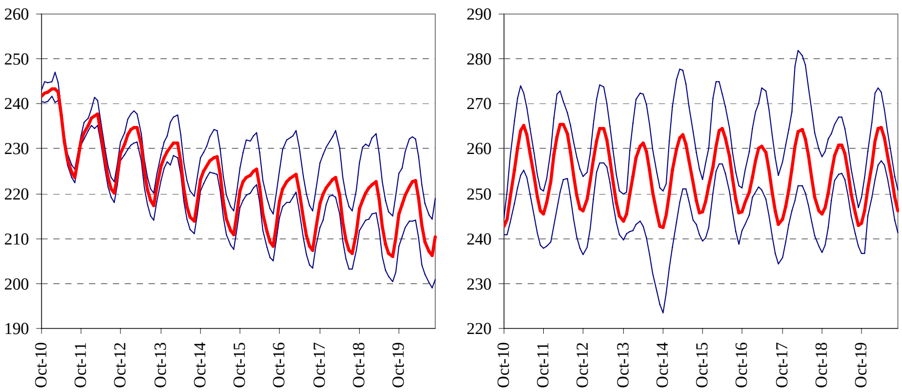


Fig. 11. Simulated level (m) at the upstream part of the aquifer (mean in red and 80% prediction limits in thin blue) under zero pumping for the water supply of Athens, according to modelling frameworks A (left) and B (right).