XXIV General Assembly of the International Union of
Geodesy and Geophysics, Perugia, 2-13 July 2007
International Association of Hydrological Sciences

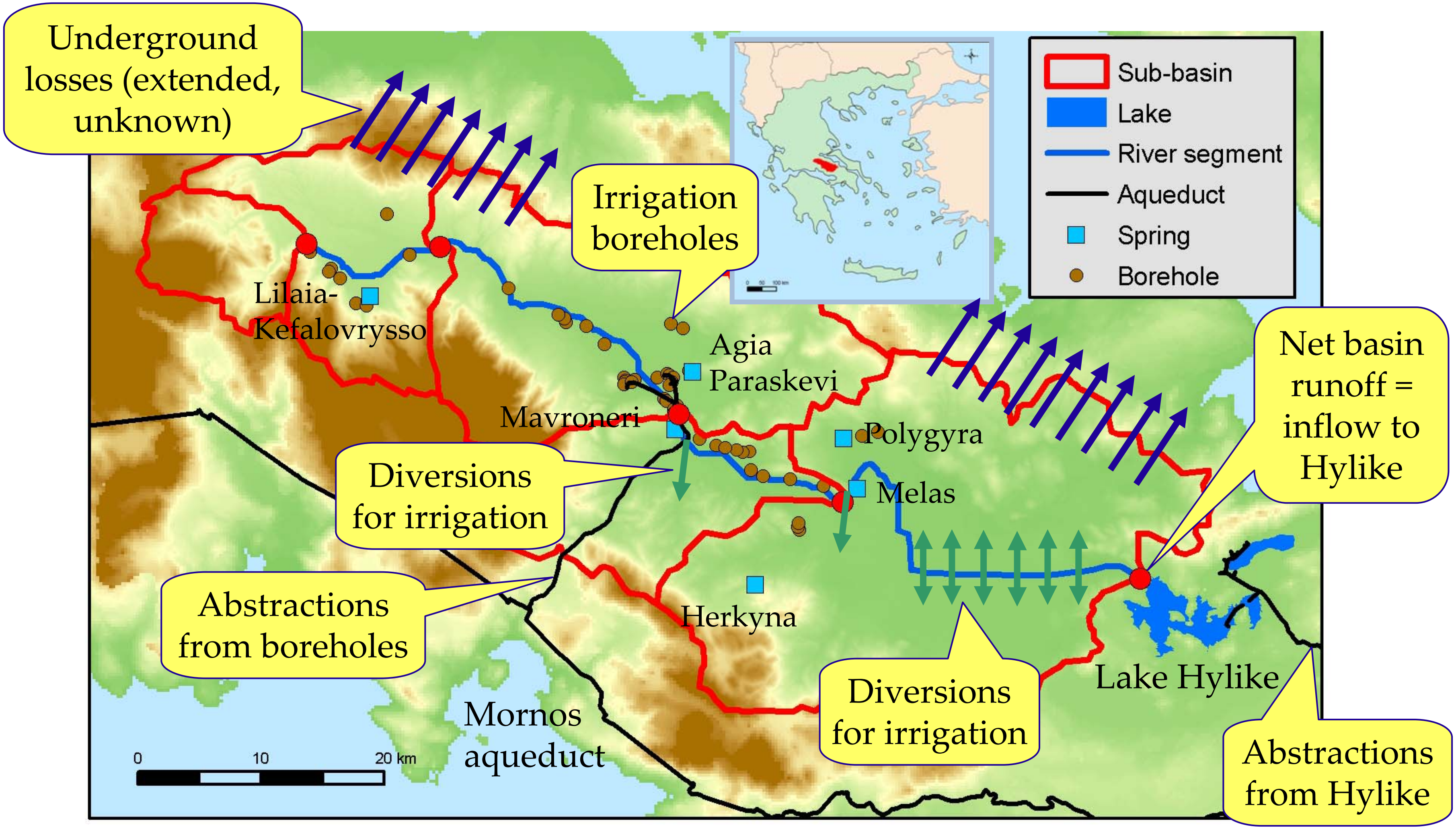
Session HW2005: From measurements and calibration to understanding and predictions

On the use and misuse of semi-distributed rainfall-runoff models

Ioannis Nalbantis, Andreas Efstratiadis and Demetris Koutsoyiannis

National Technical University of Athens

Study area: Boeoticos Kephisos river basin



Underground losses (extended, unknown)

Irrigation boreholes

Net basin runoff = inflow to Hylike

Diversions for irrigation

Abstractions from boreholes

Diversions for irrigation

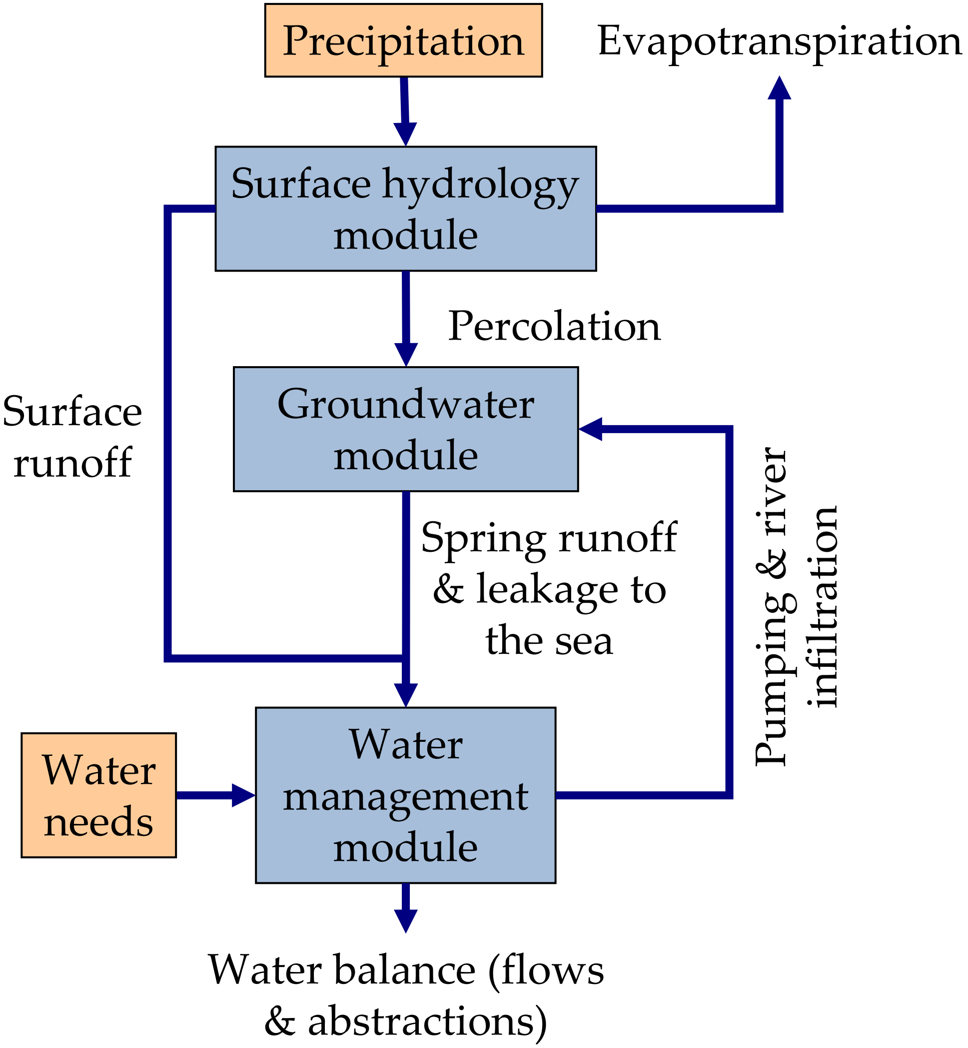
Abstractions from Hylike

Basin area 1956 km^2; Mean annual precipitation 875 mm (1710 hm^3); Mean annual runoff 146 mm (286 hm^3); Annual irrigation needs ~ 220 hm^3.

Basin peculiarities and modelling requirements

- Basin peculiarities:
  - highly non-linear interactions between surface and groundwater processes and human interventions;
  - significant contribution of baseflow to the total runoff;
  - combined uses from surface and groundwater resources, affecting drastically the hydrological regime of rivers and springs;
  - lack of real abstraction data;
  - extended leakage to the sea due to the karstic subsurface.
- Modelling requirements:
  - semi-distributed schematization;
  - representation of heterogeneities through a GIS-based approach;
  - consistency with physical characteristics;
  - reduction of computational effort;
  - exploitation of all types of information;
  - effective coupling with DSS for water management.

The modelling framework "HYDROGEIOS"



Precipitation

Evapotranspiration

Surface hydrology module

Percolation

Surface runoff

Groundwater module

Spring runoff & leakage to the sea

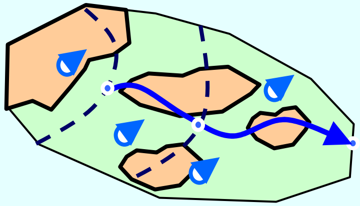
Pumping & river infiltration

Water needs

Water management module

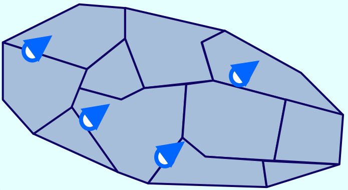
Water balance (flows & abstractions)

Surface hydrology module



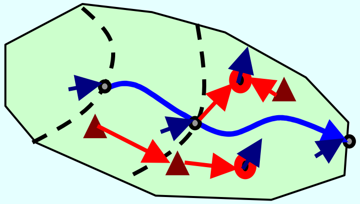
Conceptual soil moisture tanks, assigned to HRUs (hydrological response units)

Groundwater module



Multicell scheme of flexible geometry, based on the finite volume method

Water management module

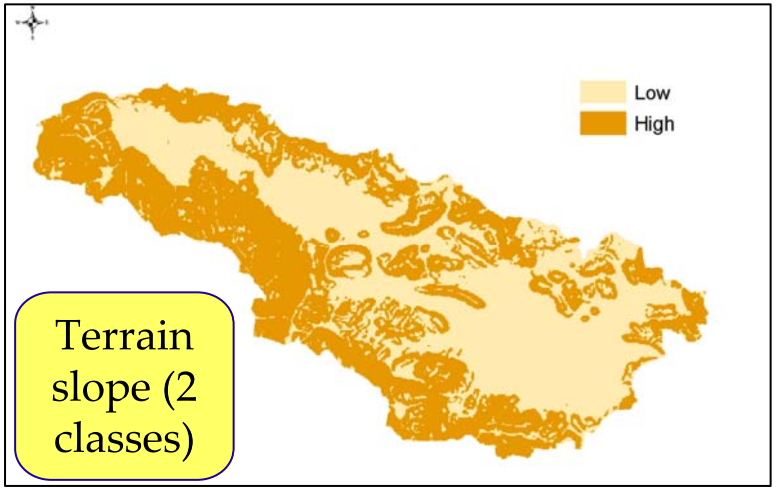


Representation of system operation via network programming

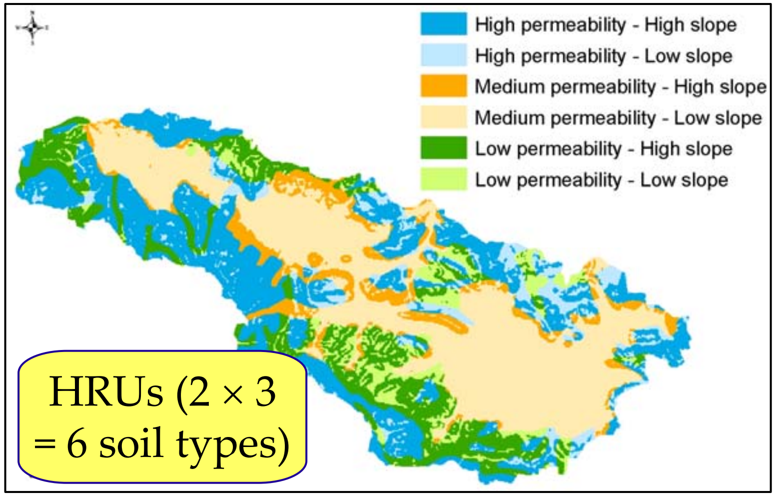
Representation of surface processes

- **Hydrographic network:**
  - The network is formulated on the basis of a digital terrain model, by adjusting the flow accumulation parameter and adding control points that correspond to flow measurement stations, diversion nodes, etc.;
  - Input data are precipitation and potential evapotranspiration, assigned to each sub-basin.
- **Hydrological response units (HRUs):**
  - Represent soil and land types, defining partitions of the basin, rather than "units" of contiguous geographical areas.
  - Defined as the product of separate partitions by different properties such as soil permeability, land cover, terrain slope, etc.
  - Through an appropriate classification of the above properties, one can adjust the number of HRUs and, consequently, the number of the parameters describing the surface hydrological mechanisms.
  - Through the HRU concept, parameters retain some physical consistency, allowing a better identification of their prior uncertainty.

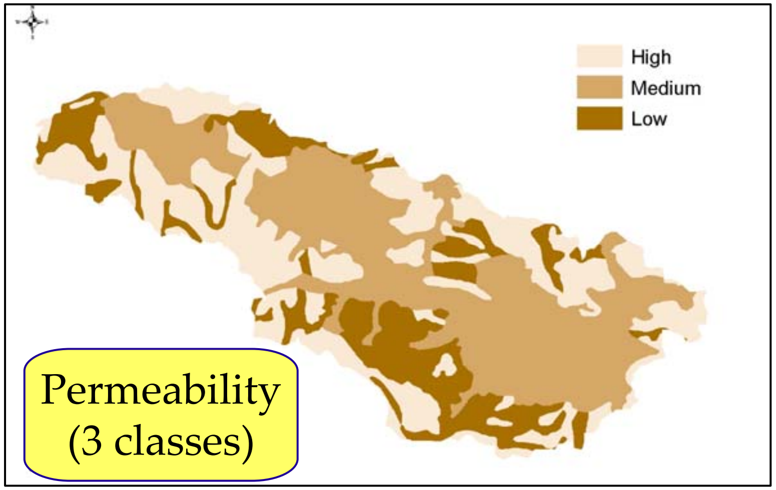
Formulation of hydrographic network and hydrological response units (HRUs)



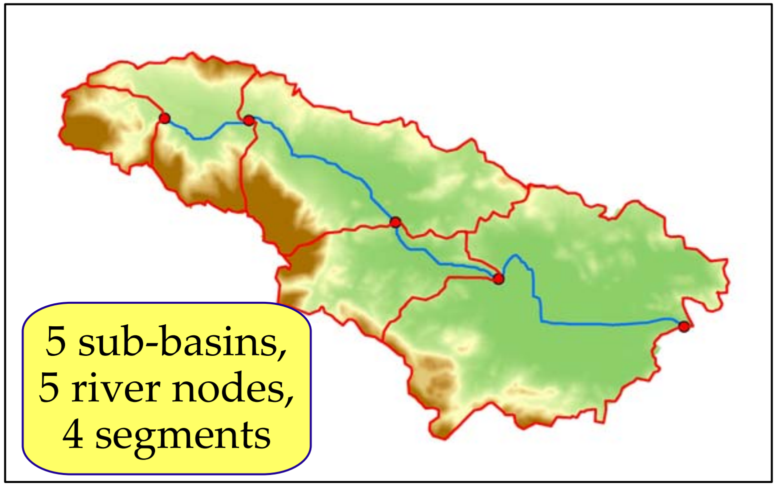
Terrain slope (2 classes)



HRUs (2 × 3 = 6 soil types)

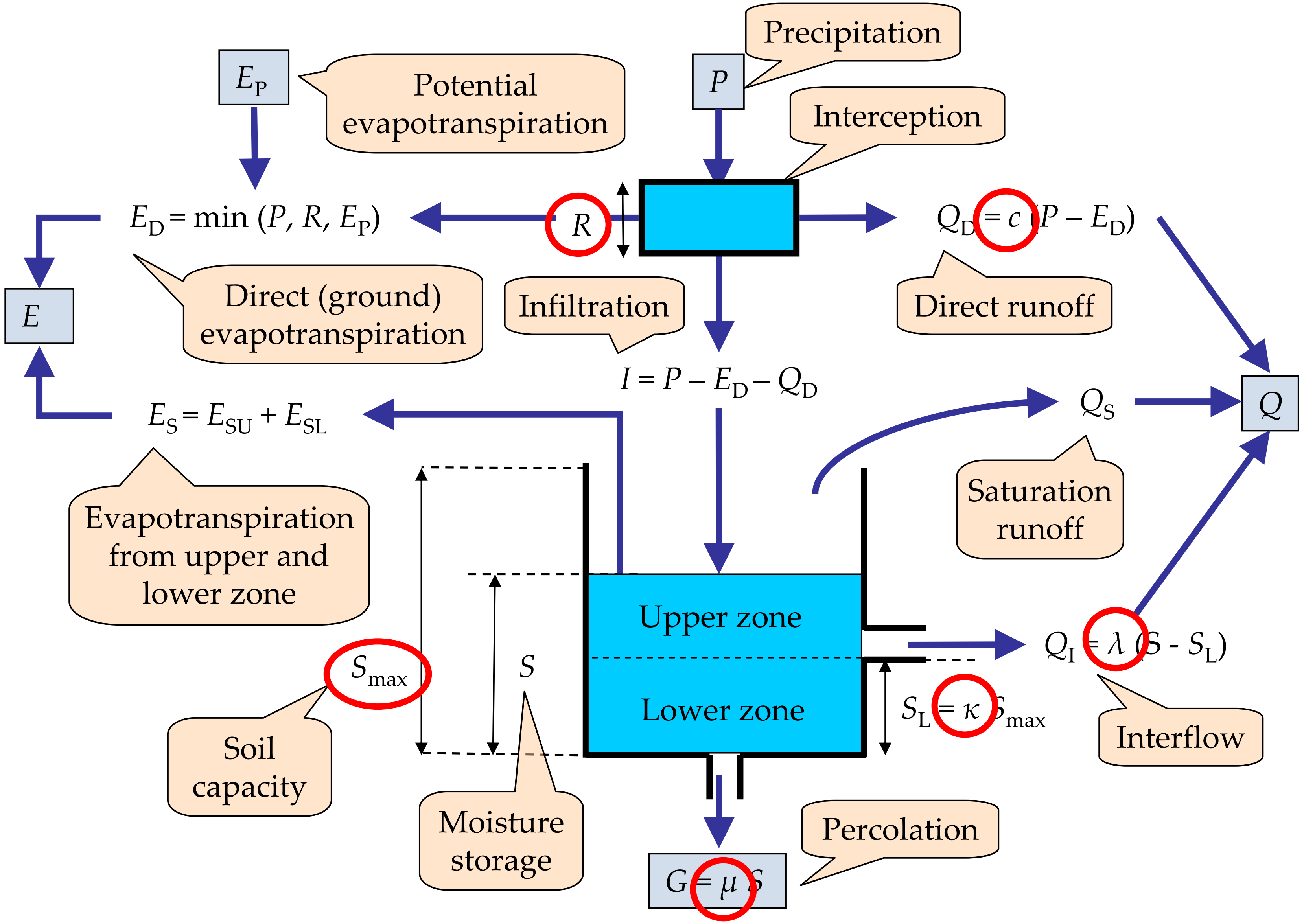


Permeability (3 classes)



5 sub-basins, 5 river nodes, 4 segments

Surface hydrology module (assigned to HRUs)



Precipitation

Potential evapotranspiration

E_P

P

Interception

$E_D = \min(P, R, E_P)$

R

$Q_D = c\,(P - E_D)$

Direct (ground) evapotranspiration

Infiltration

Direct runoff

E

$I = P - E_D - Q_D$

Q_S

Q

$E_S = E_{SU} + E_{SL}$

Saturation runoff

Evapotranspiration from upper and lower zone

Upper zone

Lower zone

$Q_I = \lambda\,(S - S_L)$

S_{max}

S

$S_L = \kappa\,S_{max}$

Interflow

Soil capacity

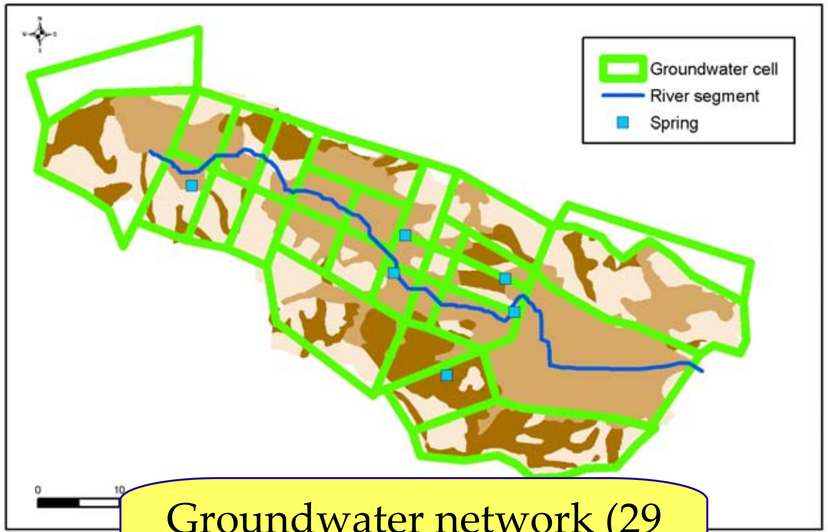
Moisture storage

Percolation

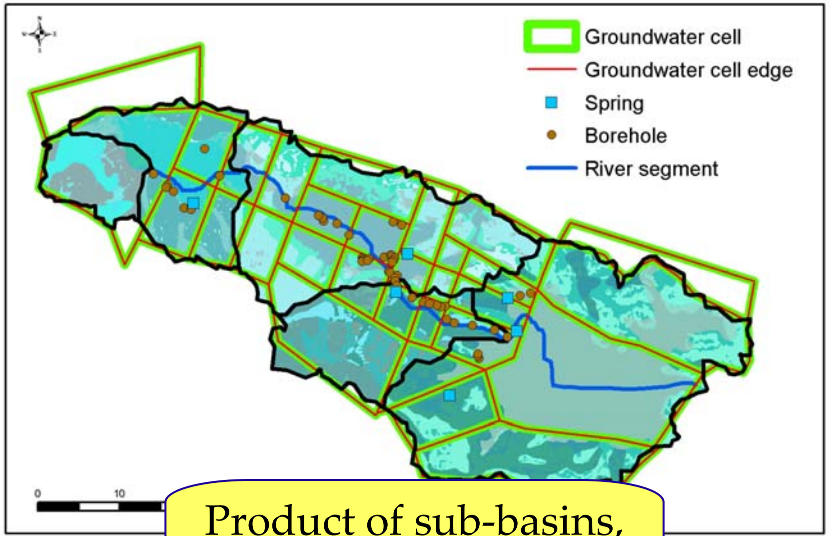
$G = \mu\,S$

Groundwater module

- Follows a multicell approach with non-rectangular discretization of the aquifer and implements the finite volumes method with simplified integration.
- 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 per cell (conductivity, specific yield).
- Springs and underground losses are assumed virtual cells of very large base.
- Model inputs are: (a) areal inflows, due to percolation through each sub-basin and HRU combination; (b) inflows due to infiltration underneath each river segment; (c) point outflows due to pumping from each borehole.



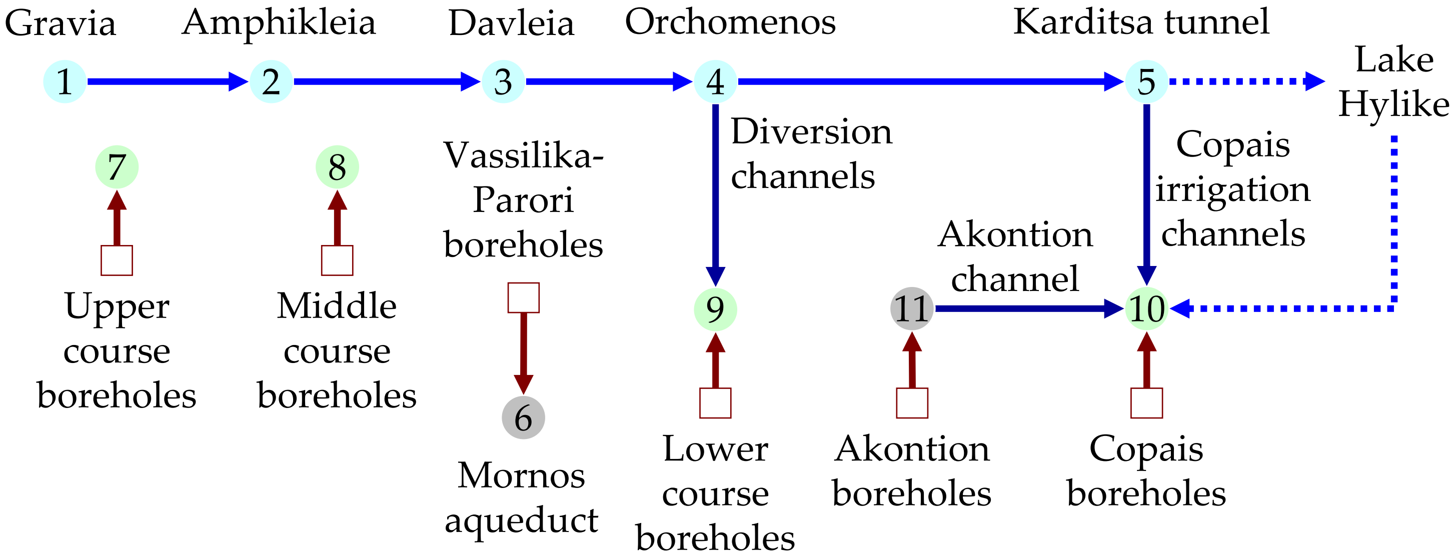
Groundwater network (29 cells, 6 springs, 2 leakages)



Product of sub-basins, HRUs and cells

Water management module

- Coarse depiction of the major hydraulic works, the corresponding water uses and constraints and their interactions with the physical system.
- Properties of network components are discharge and pumping capacities, target priorities, demand time series and unit transportation costs.
- Dynamic inputs are surface and groundwater runoff, assumed nodal inflows.
- 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 to fulfil the hydrosystem targets.
- Model outputs are river and aqueduct flows, infiltrations and abstractions.



The hybrid calibration strategy

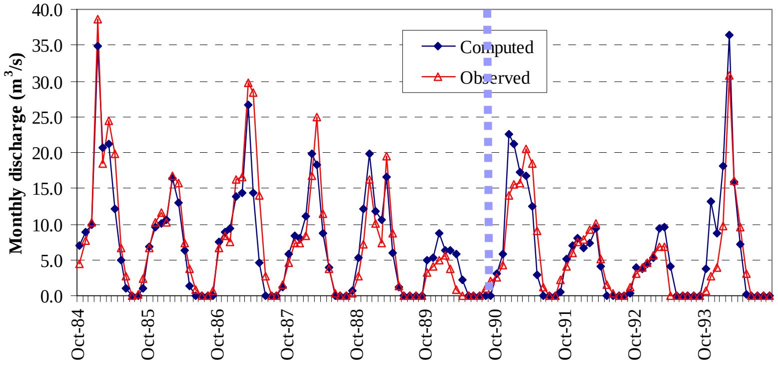
- The entire model uses only 9 parameter types.
- However, the total number of parameters is large (~ 100; 40 and 60 respectively for the surface and ground water modules).
- Multiple responses and criteria are required for the calibration problem.
- The objective function embraces a number of statistical and empirical criteria:
  - efficiency and bias of the monthly hydrographs at the basin outlet and downstream of the six springs;
  - penalties for not reproducing flow intermittencies;
  - penalties for unrealistic trends of groundwater levels.
- Calibration was based on a combined strategy, utilizing hydrological experience and automatic multiobjective optimization.
- Optimization was carried out through a fast evolutionary annealing-simplex method.
- Apart from the model performance during the calibration period, the solutions provided by optimization were also evaluated on the basis of efficiency values in validation and the consistency of parameters.
- By manually changing the parameter bounds and the weights of criteria, we guided the search towards a realistic, best-compromise parameter set.

Synoptic results: Performance indices

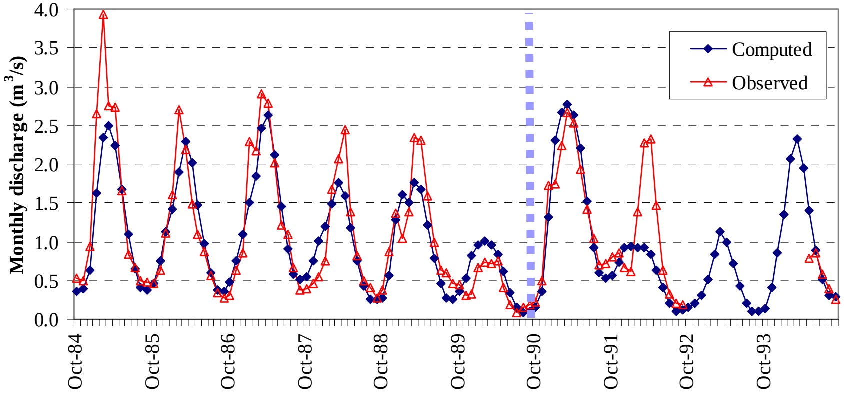
	Calibration period		Validation period	
Monthly runoff	Efficiency	Average bias	Efficiency	Average bias
Basin outlet	0.870	0.058	0.761	-0.116
Lilea-Kefalovrysso springs	0.809	0.069	0.605	0.180
Agia Paraskevi springs	0.707	0.106	–	–
Mavroneri springs	0.724	0.038	0.601	0.480
Herkyna springs	0.446	-0.040	0.403	-0.049
Melas springs	0.265	0.036	0.028	0.231
Polygyra springs	0.372	0.024	–	–

Calibration period: October 1984 - September 1990
Validation period: October 1990 - September 1994

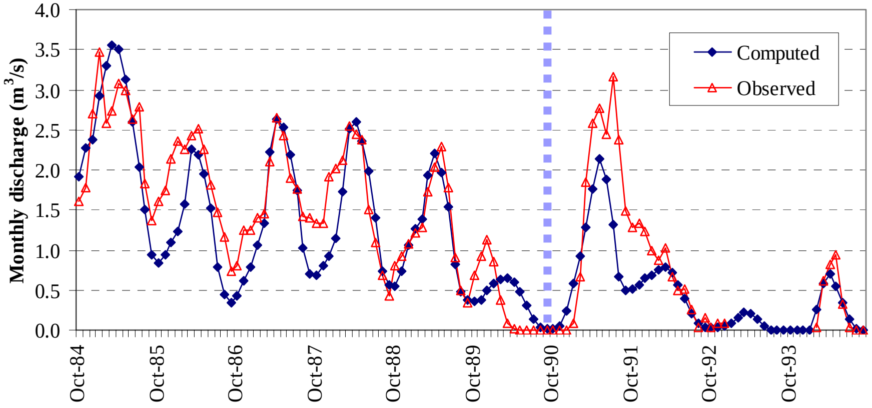
Synoptic results: Characteristic hydrographs



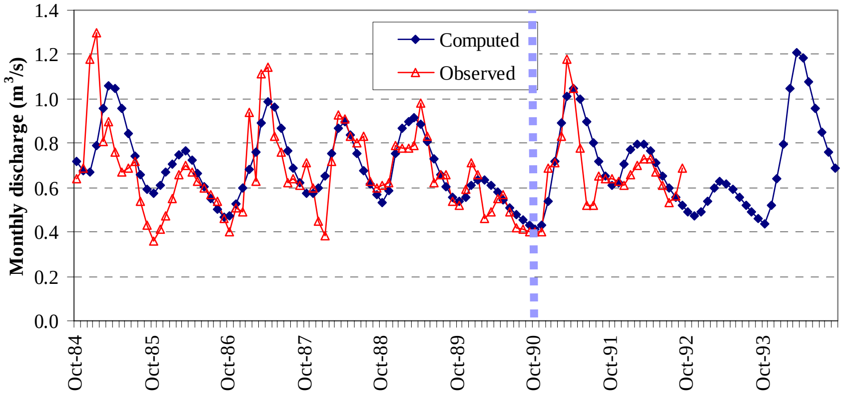
Basin outlet



Lilaia-Kefalovrysso springs

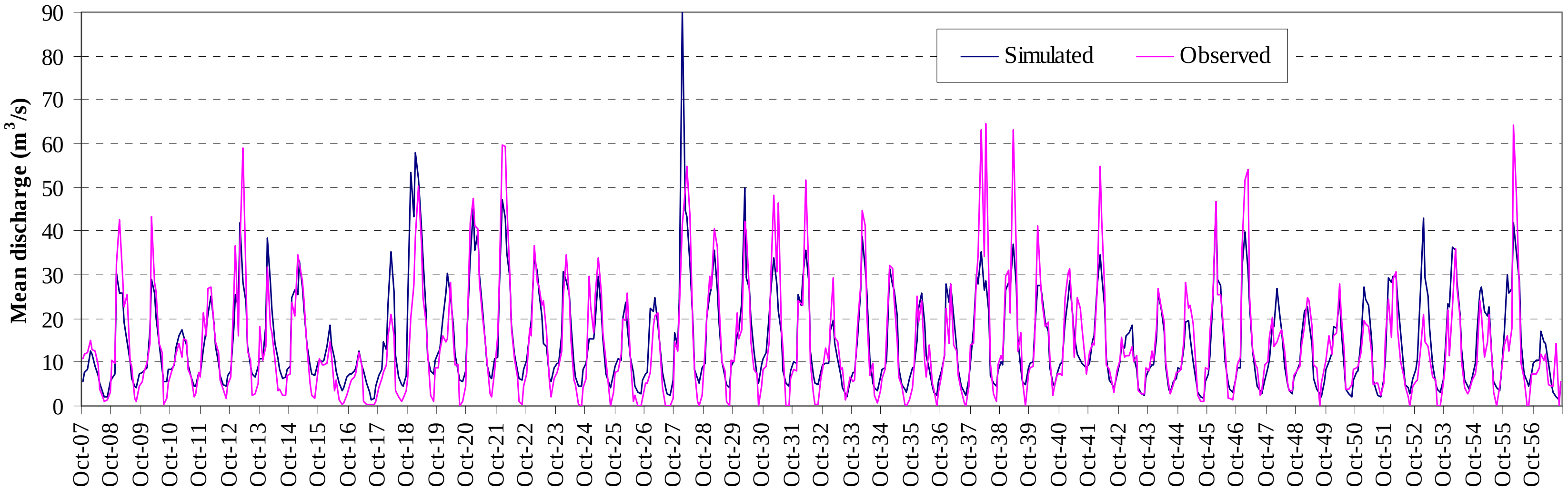


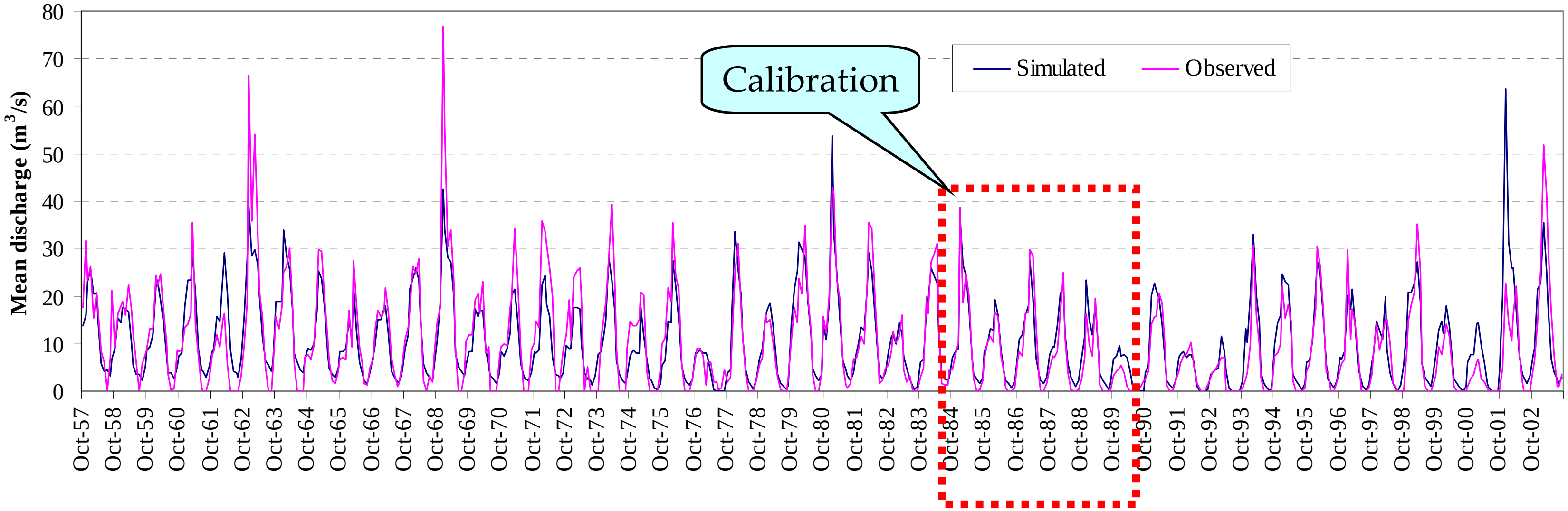
Mavroneri springs



Herkyna springs

Long-period validation (1907-2003)

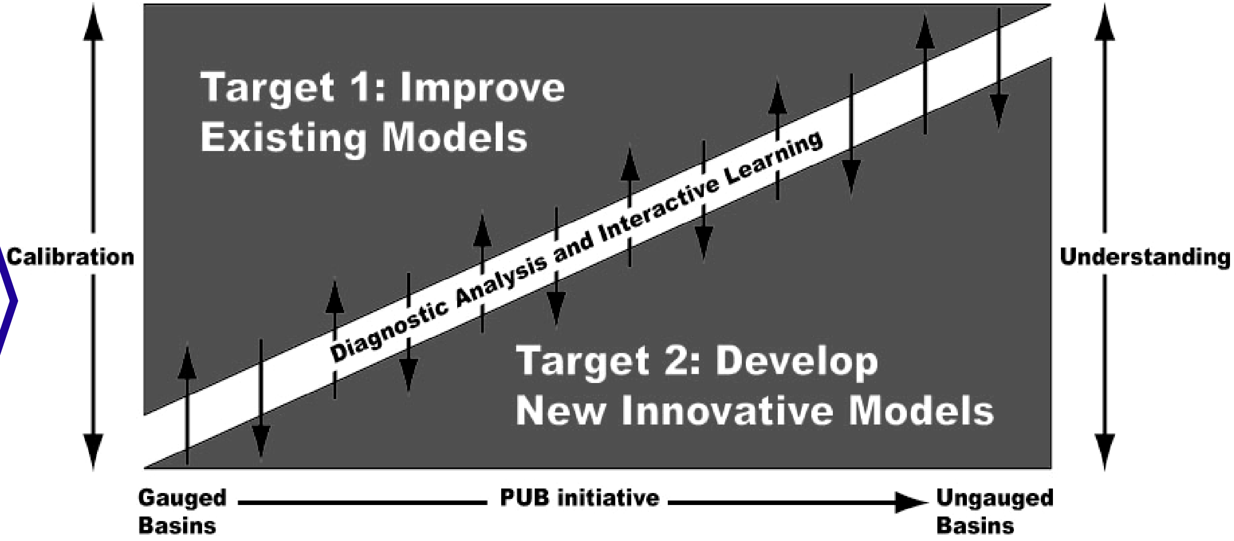




Misuse case 1: Depreciation of measurements and calibration; hope to replace them with better understanding

PUB has prompted towards paradigm change: from models based on calibration to models based on increased understanding

Source: M. Sivapalan et al. (2003), IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.*, 48(6), 857-880



- Can we make an accurate model even for a simple system with well understood hydraulics, e.g. river discharge vs. river stage, without flow measurements and without calibration based upon the measurements?
- Can we ever know even the geometry of a karst aquifer or of the leaves of plants in a basin microscopically? (and is it useful at all?)
- Can we built a macroscopic karst model or an evapotranspiration model without measurements and calibration?

Misuse case 1: measurements and calibration vs. understanding (2)

- In complex systems understanding of physical behaviours absolutely relies on understanding of probability, statistics and stochastics.
- Deterministic thinking and mechanistic analogues may be obstacles in understanding.
- Examples: quantum physics, thermophysics, dynamical systems.
- The mechanistic caloric theory, according to which heat consists of a weightless fluid called "caloric" that flows from hotter to colder bodies, did not help understanding thermodynamics.
- Knowledge of complex systems relies on probability and statistics:
  - *“Even if we would know everything, we should still have to derive statistical information from this knowledge in order to answer what are essentially statistical problems, such as explaining gas pressure or the intensity of spectral lines”*. (Karl Popper to Albert Einstein)
- Even the definition of fundamental physical quantities relies on probability:
  - The definition of entropy ($\phi = E[-\ln p(X)]$) is based on probability.

Misuse case 1: measurements and calibration vs. understanding (3)

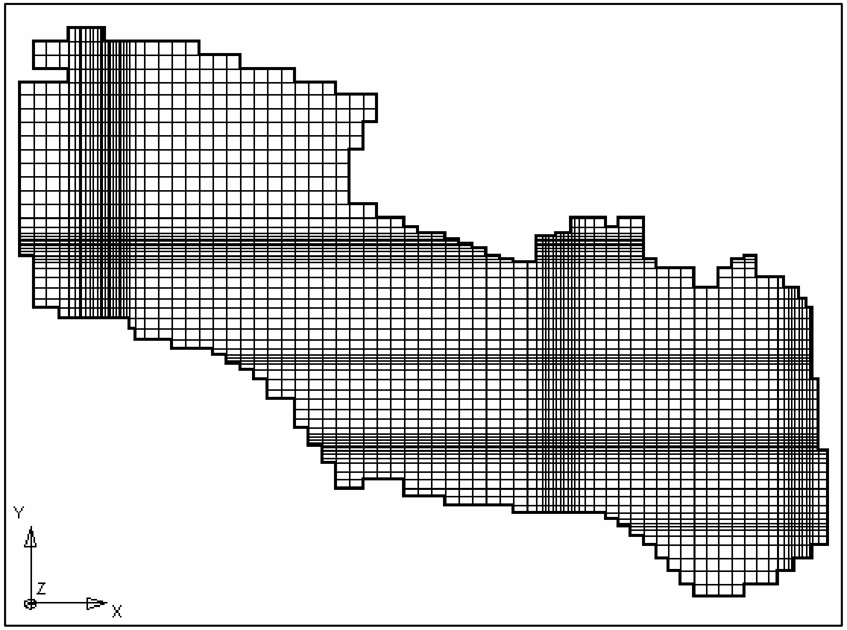
- In probability and statistics there are two ways of developing laws and models:
  - deduction, as in mathematical theorems – use of probability theory;
  - induction from observations (data) – use of statistical theory.
- Statistical thermodynamical laws are mostly produced by deduction.
  - This was possible because a gas or a fluid consists of only a few types of molecules; in each type (e.g. H_2O) all molecules are precisely identical.
- For the same reason, fundamental quantities (e.g. specific heat) are representative for huge systems and can be inferred with deduction.
- In hydrological systems/subsystems all components are unique – they are not identical to each other (e.g. trees, leaves of a a single tree, karstic pipes).
- Thus deduction is not effective; we should use induction instead.
- Induction requires data and quantities should be estimated from data.

One measurement is worth a thousand models.

Misuse case 2: Employing different models for surface and groundwater processes

- One-way linking of a surface water model and a ground water model is a common method in hydrological practice.
- It is often based on simplified assumptions for estimating groundwater inflows (e.g. percolation is assumed a constant ratio of precipitation).
- Presupposes that surface and ground water abstractions are known.
- It is infeasible in the case of interactions between the surface and the groundwater processes.

In the study basin, a major part of runoff originates from spring outflows, whereas significant part of groundwater comes from river infiltration.



An earlier MODFLOW approach, with 500×500 m^2 grid

Misuse case 3: Employing different models for natural processes and water management practices

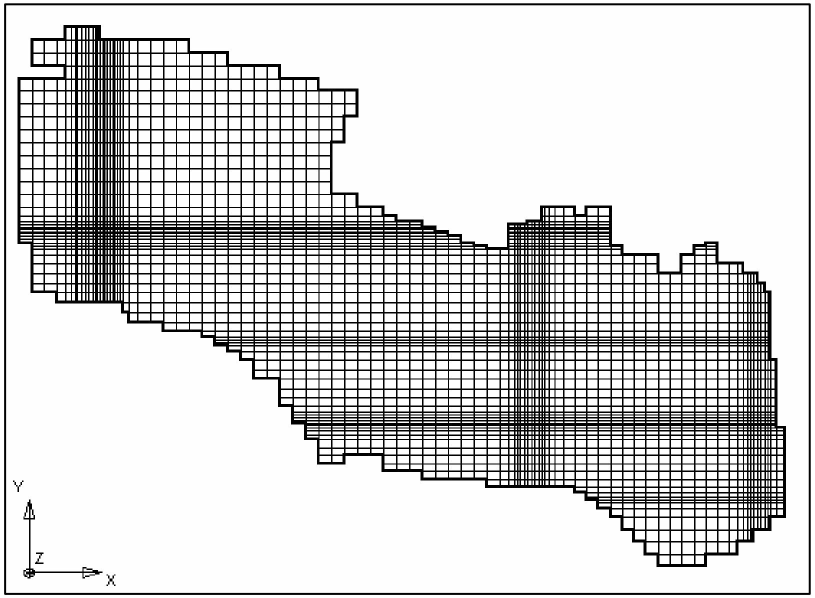
- A typical two-stage procedure in hydrological practice is to build a model for the natural processes whose output is used as input in a water management model.
- However, in disturbed catchments, engineering structures and management practices affect (interact with) the natural processes (surface and underground). Thus, the method is inappropriate.
- Also, it is infeasible when real abstractions are unknown.
- The implementation of automatic calibration within a two-stage approach is questionable.

In the study basin, a water management module is necessary to reconstruct real abstractions from surface and ground water resources, on the basis of theoretical needs, costs and priorities; however, these abstractions have significant impacts to the river basin dynamics, making an off-line co-operation of hydrological and water management models infeasible.

Misuse case 4: Employing too detailed models

- In theory, detailed physically-based approaches provide the best representation of natural systems.
- However, in large-scale systems their hydraulic parameters usually cannot be inferred from field data due to heterogeneity.
- A detailed parameterization and parameter estimation by optimization is infeasible (curse of dimensionality).
- The computational burden precludes their use in decision support systems that require thousands of simulations

In the study basin, a lumped model had better performance in terms of reproducing river flow than a detailed grid-based model (because of more efficient parameter optimization)



Misuse case 5: Blind use of optimization in calibration – Fitting on a single hydrograph

- In a naïve view, calibration is a “black-box” procedure, generating a “global optimal” parameter set through an automatic algorithmic procedure.
- This usually may lead to:
  - parameter values inconsistent with their physical explanation;
  - poor predictive capacity on validation;
  - non reasonable behaviour of unobservable responses (e.g. actual evapotranspiration, underground losses) and state variables of models (e.g. soil and groundwater storage).
- Given a semi-distributed schematization and a relatively large number of parameters, it is necessary to:
  - calibrate parameters on multiple observed responses;
  - insert empirical criteria to control unmeasured responses
  - take advantage of the interpretation of parameters.

Conclusions

- Key characteristics of the HYDROGEIOS approach are:
    - incorporation into the same computer application of a surface water module, a ground water module and a water management module;
    - computational efficiency for use with a DSS;
    - use of GIS for automating processing of spatial information;
    - use of automatic multiobjective optimization algorithms;
    - parsimonious use of parameters;
    - physical consistency of modelling components and parameters.
- The hydrologist’s experience plays a very important role, with respect to:
    - the model schematization and parameterization;
    - the effective use of optimization algorithms to guide calibration towards realistic parameter sets;
    - the testing of consistency and interpretation of all simulated responses (not only the calibrated ones);
    - the selection of the best compromise solution.