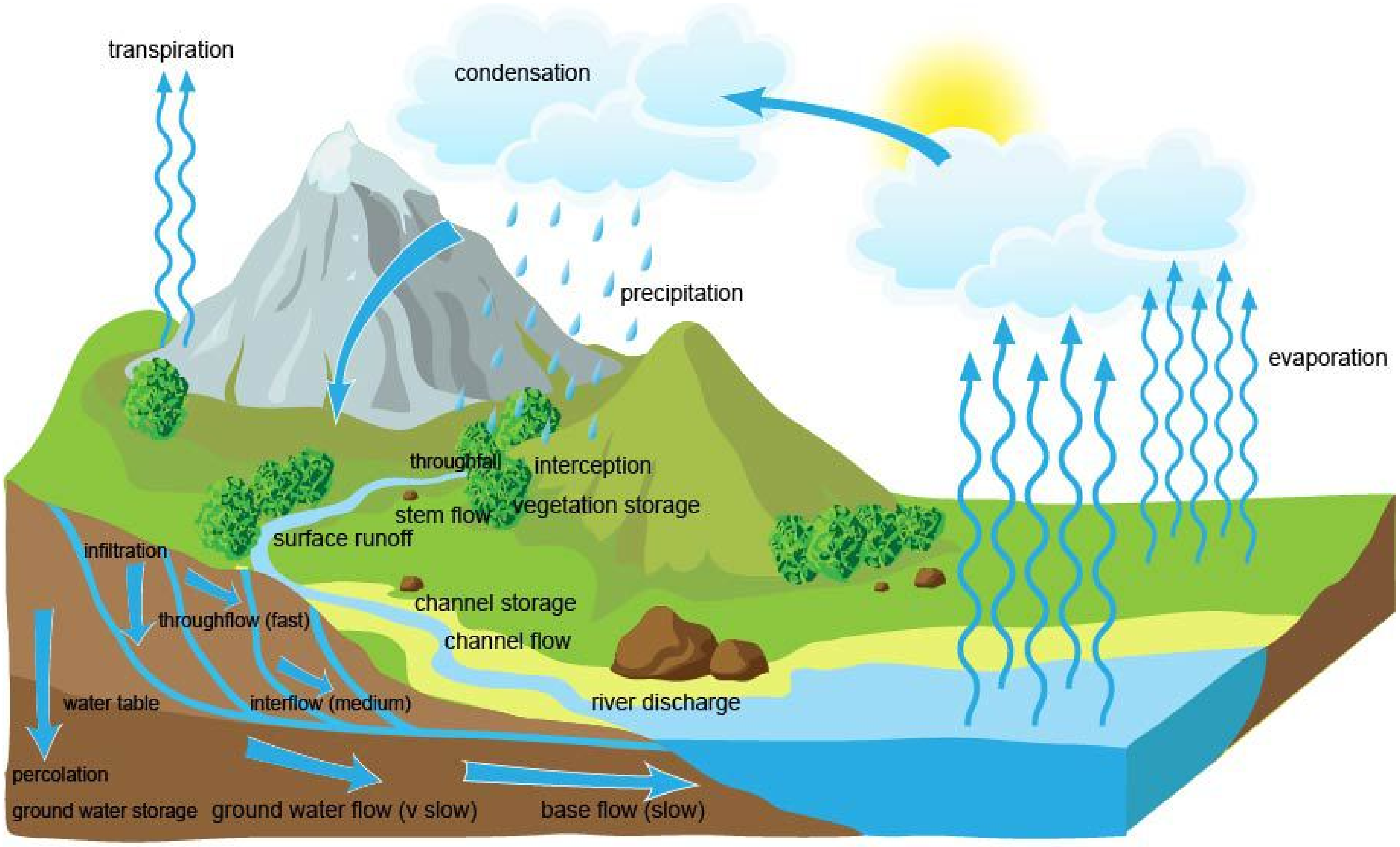
Calibration of hydrological models by PEST

Andrej Vidmar, Mitja Brilly

University of Ljubljana, Faculty of Civil and Geodetic Engineering,
Department of Environmental Civil Engineering, Ljubljana, Slovenia

Kyiv, Ukraine, November 6-8, 2019

The drainage basin hydrological cycle (4D space)



Water does not come into or leave planet earth. Water is continuously transferred between the atmosphere and the oceans. This is known as the global hydrological cycle.

Complex proceses:

- governed by nature
- essentially stochastic
- unique
- non-recurring
- changeable across space and time
- non-linear

Source: http://www.alevelgeography.com/drainage-basin-hydrological-system

(Mikhail Lomonosov law of mass conservation in 1756)

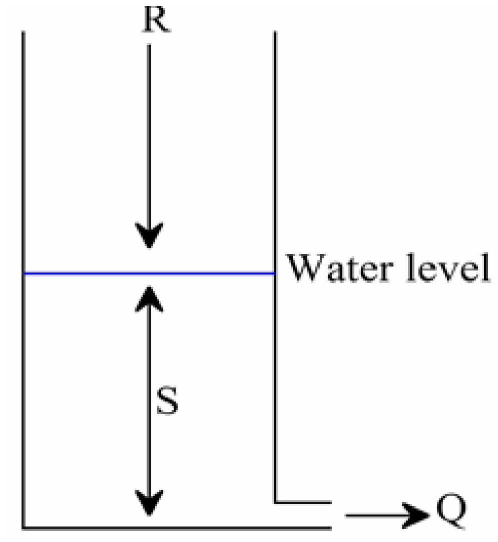
1st _ Why use a P-R modeling?

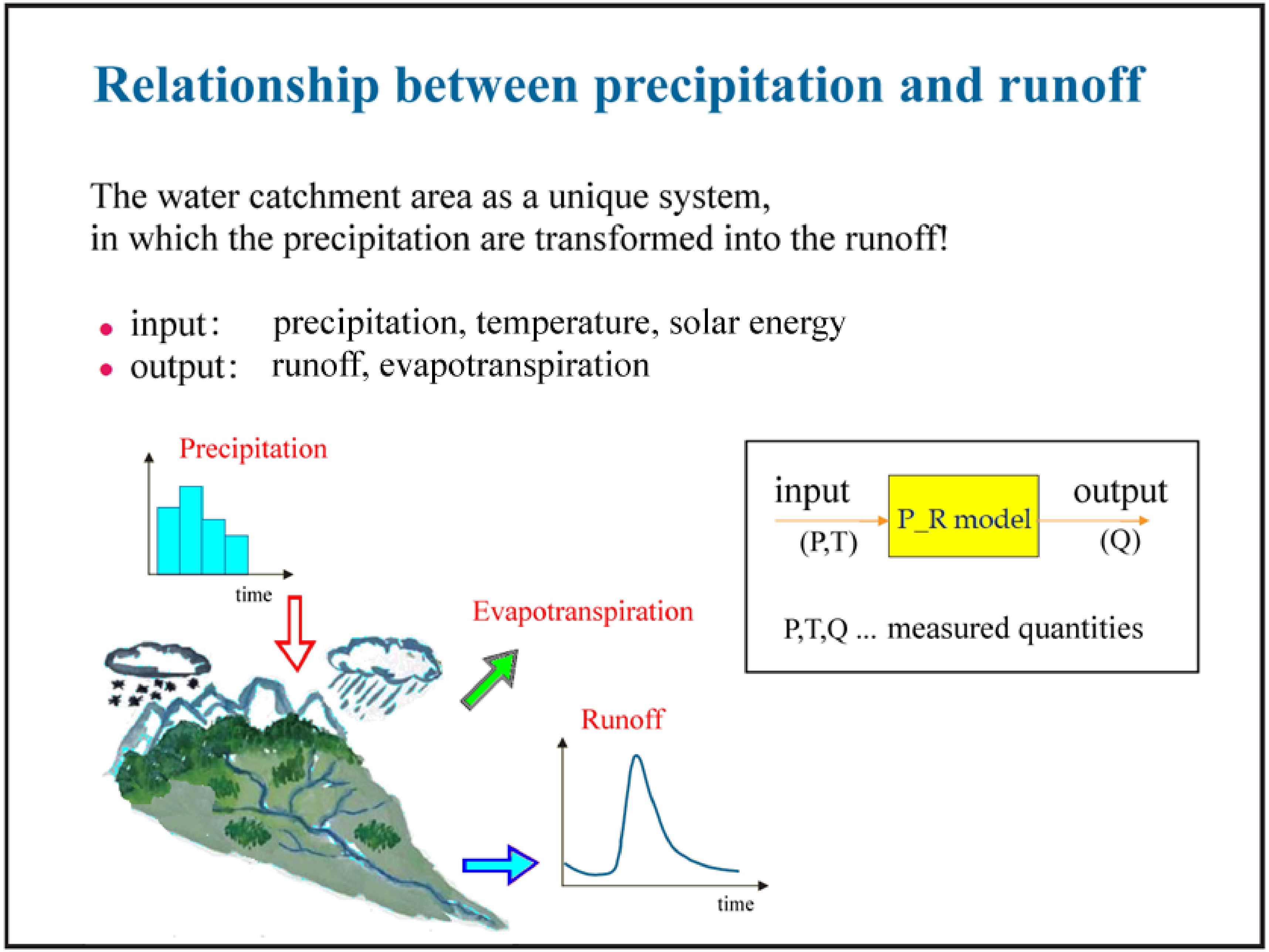
- for education
- for decision support
- for data quality control
- for water balance studies
- for drought runoff forecasting (irrigation)
- for fire risk warning
- for runoff forecasting/prediction (flood warning and reservoir operation)
- inflow forecast the most important input data in hydropower
- for what happens if' questions in watershed level

2^{nd} _ P-R modeling is used?

- to compute design floods for flood risk detection
- to extend runoff data series (or filling gaps)
- to compute design floods for dam safety
- to investigate the effects of land-use changes within the catchment
- to simulate discharge from ungauged catchments
- to simulate climate change effects
- to compute energy production

The basic principle





Modeling noise

- Structural noise
  - Conceptual noise (semi-distributed,…)
  - Design noise
  - **Calibration noise (parameter estimation noise)**
- Measurements noise (P [rr], T, E, Evap)
- Computer noise (discrete numbers, double precisions, rounding…)

HBV - Hydrologiska Byråns Vattenbalansavdelning (Hydrological Agency Water Balance Department)

- The HBV model (Bergström, 1976, 1992) is a rainfall-runoff model, which includes conceptual numerical descriptions of hydrological processes at the catchment scale. The general water balance can be described as

$$P - E - Q = \frac{d}{dt}[SP + SM + TZ + UZ + LZ + lakes]$$

Where

P = precipitation

E = evapotranspiration

Q = runoff

SP = snow pack

SM = soil moisture

TZ = storage in soil top zone (introduced in HBV-light)

UZ = upper groundwater zone storage

LZ = lower groundwater zone storage

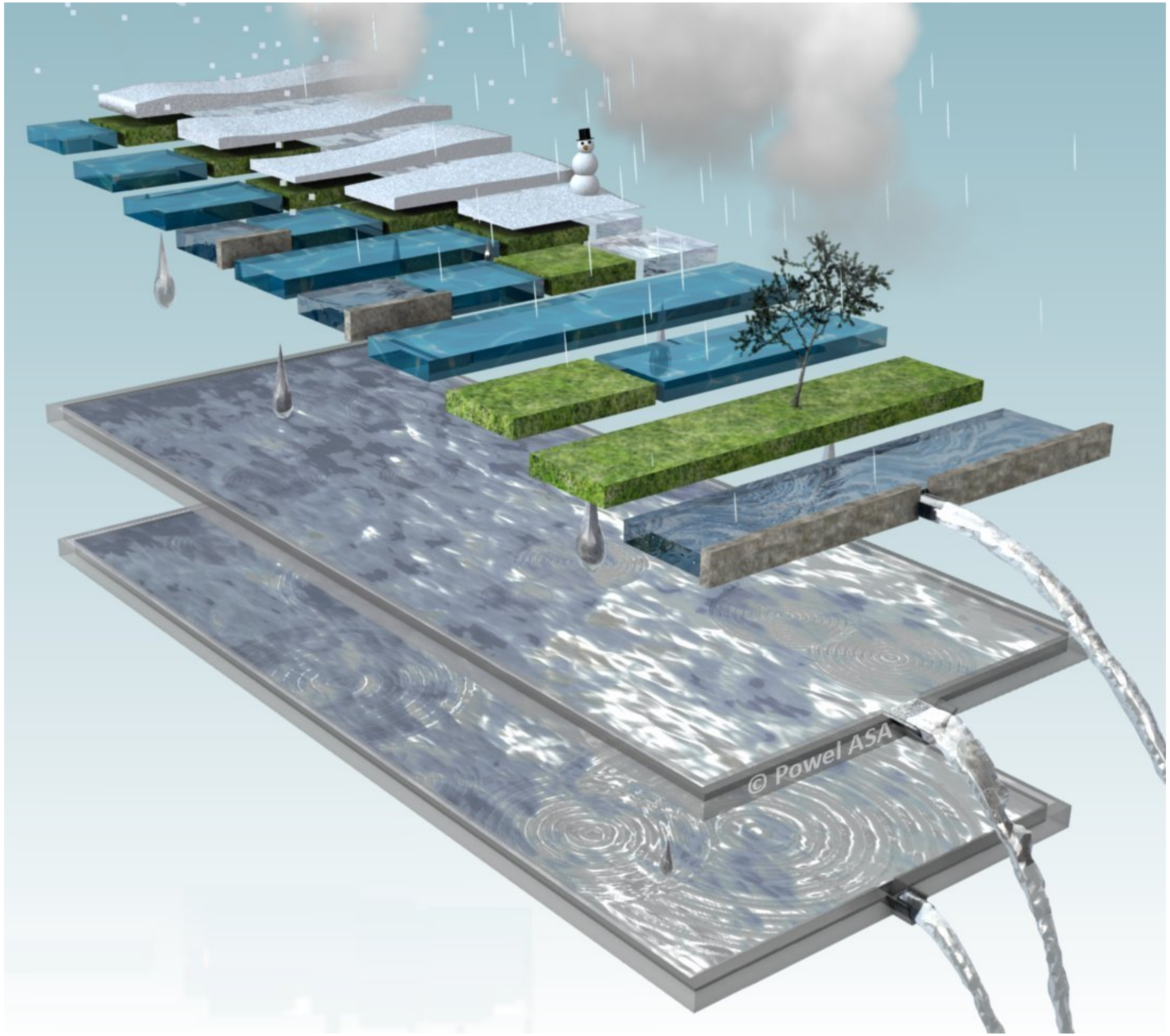
lakes = lake volume

Semi-distribution

Subdivides a large problem into smaller, simpler parts with unique characteristic

- **Elevation zones**
- **Vegetation zones**

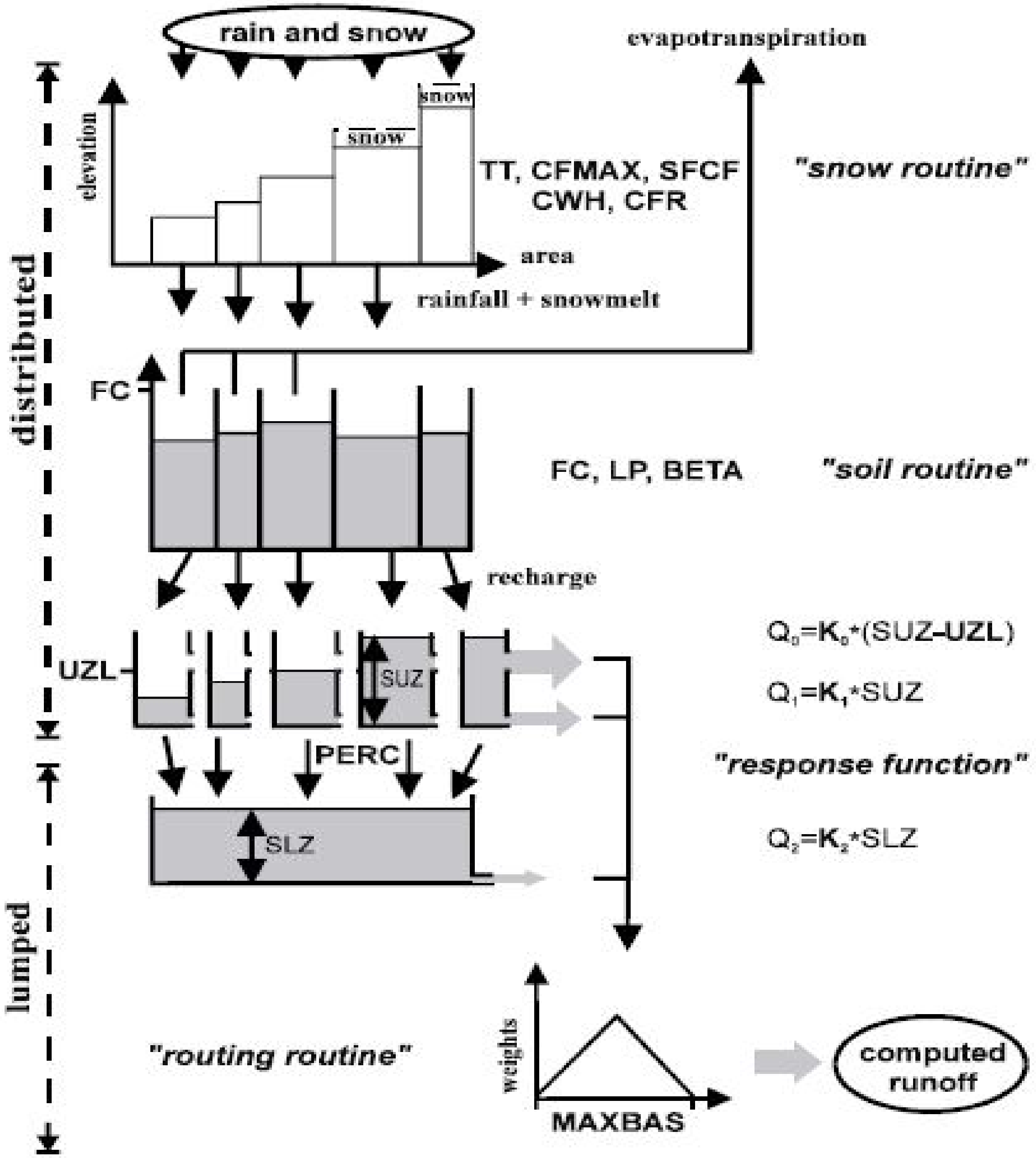
Conceptual noise



HBV overview

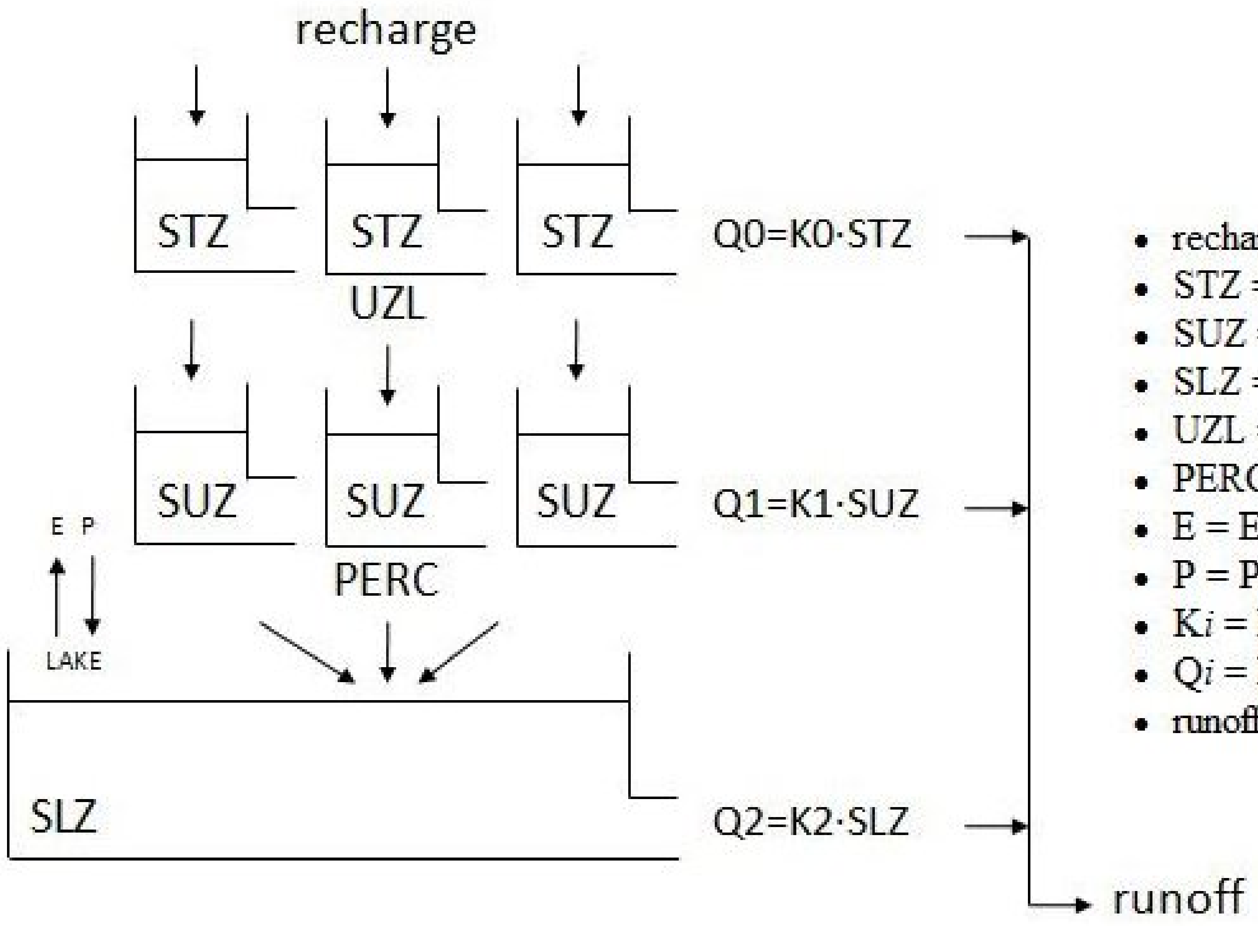
- The HBV model is a **simple multi-tank-type model** for simulating runoff.

Conceptual noise



Model of Computed Runoff

Design noise



- recharge = Input from soil routine [mm/Δt]
- STZ = Storage in soil top zone [mm]
- SUZ = Storage in soil upper zone [mm]
- SLZ = Storage in soil lower zone [mm]
- UZL = Maximum percolation to the soil upper zone [mm/Δt]
- PERC = Maximum percolation to the soil lower zone [mm/Δt]
- E = Evaporation from the lake
- P = Precipitation into the lake
- K_i = Recession coefficient [1/Δt]
- Q_i = Runoff component [mm/Δt]
- runoff = Total amount of generated runoff [mm/Δt]

Source: Help HBV-light – Three GW Box (Distributed STZ and SUZ) Model

Equations Overview

Conceptual noise

$$melt = CFMAX(T(t) - TT)$$

$$refreezing = CFR \; CFMAX(TT - T(t))$$

$$\frac{recharge}{P(t)} = \left(\frac{SM(t)}{FC}\right)^{BETA}$$

$$E_{act} = E_{pot} \; \min\left(\frac{SM(t)}{FC \cdot LP}, \; 1\right)$$

$$Q_{GW}(t) = K_2 SLZ + K_1 SUZ + K_0 \max(SUZ - UZL, 0)$$

$$Q_{sim}(t) = \sum_{i=1}^{MAXBAS} c(i) \; Q_{GW}(t - i + 1)$$

$$\text{where } \; c(i) = \int_{i-1}^{i} \frac{2}{MAXBAS} - \left| u - \frac{MAXBAS}{2} \right| \frac{4}{MAXBAS^2} du$$

$$P(h) = P_0 \left(1 + \frac{PCALT\,(h - h_0)}{10000}\right)$$

$$T(h) = T_0 - \frac{TCALT\,(h - h_0)}{100}$$

$$E_{pot}(t) = \left(1 + C_{ET}\left(T(t) - T_M\right)\right) E_{pot,M}$$

$$\text{but } \; 0 \le E_{pot}(t) \le 2 \; E_{pot,M}$$

Source: Help HBV-light – An Overview of the HBV Model

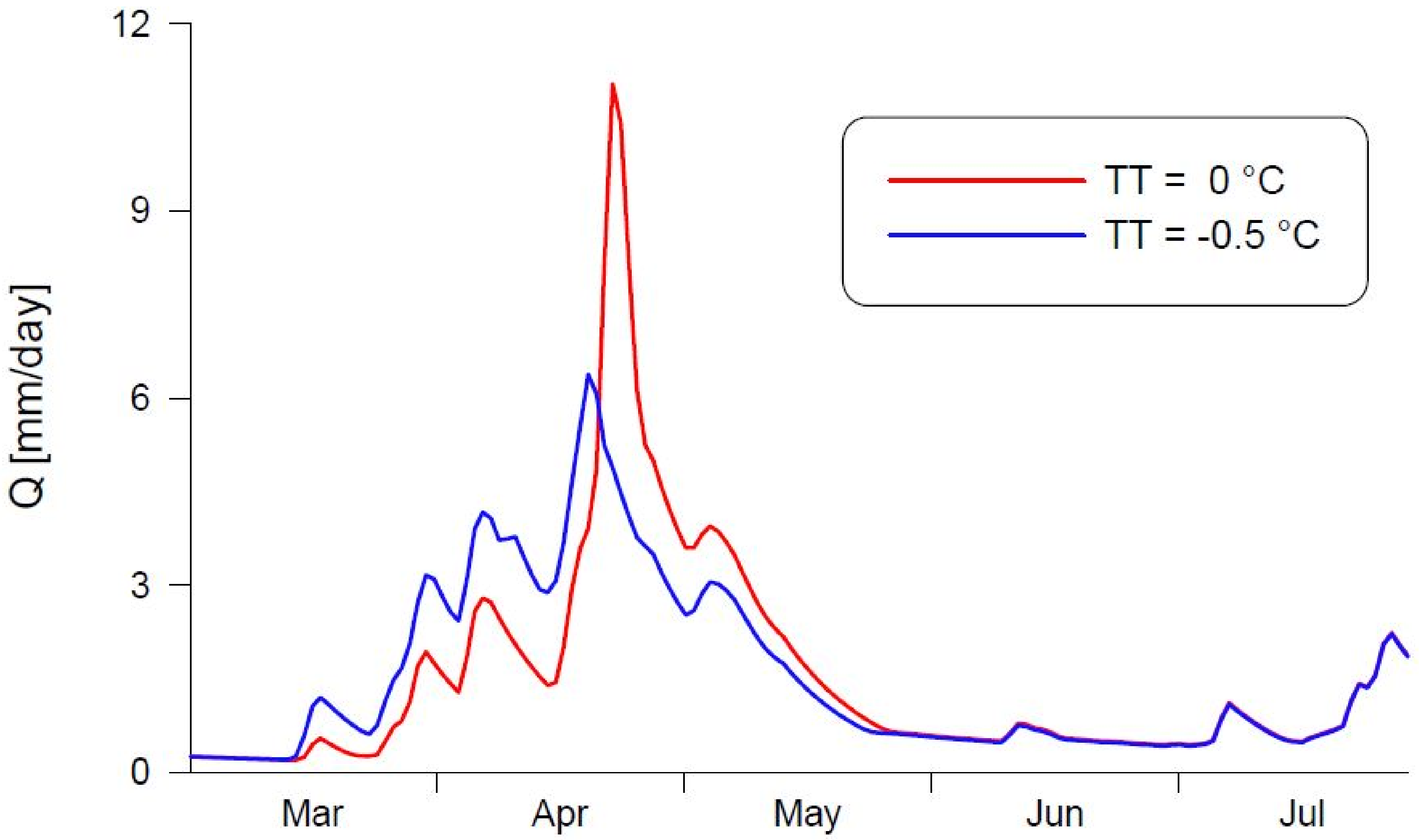
Vegetation Zone Parameters

Name	Unit	Valid range	Default value	Description	See also
TT	°C	(-inf,inf)	0	threshold temperature	Snow Routine
CFMAX	mm/Δt °C	[0,inf)	3	degree-Δt factor	Snow Routine
SFCF	-	[0,inf)	1	snowfall correction factor	Snow Routine
CFR	-	[0,inf)	0.05	refreezing coefficient	Snow Routine
CWH	-	[0,inf)	0.1	water holding capacity	Snow Routine
CFGlacier	-	[0,inf)	1	glacier correction factor	Glacier Model
CFSlope	-	(0,inf)	1	slope correction factor	Aspect Model Glacier Model
FC	mm	(0,inf)	200	maximum soil moisture storage	Soil Moisture Routine
LP	-	[0,1]	1	soil moisture value above which AET reaches PET	Soil Moisture Routine
BETA	-	(0,inf)	1	parameter that determines the relative contribution to runoff from rain or snowmelt	Soil Moisture Routine

Catchment Parameters

Name	Unit	Valid range	Default value	Description	See also
PERC	mm/Δt	[0,inf)	1	treshold parameter	Response Function
Alpha	-	[0,inf)	0	non-linearity coefficient	Response Function
UZL	mm	[0,inf)	20	treshold parameter	Response Function
K0	1/Δt	[0,1)	0.2	storage (or recession) coefficient 0	Response Function
K1	1/Δt	[0,1)	0.1	storage (or recession) coefficient 1	Response Function
K2	1/Δt	[0,1)	0.05	storage (or recession) coefficient 2	Response Function
MAXBAS	Δt	[1,100]	1	length of triangular weighting function	Routing Routine
Cet	1/°C	[0,1]	0	potential evaporation correction factor	An Overview of the HBV Model
PCALT	%/100m	(-inf,inf)	10	change of precipitation with elevation	Height Increment Variables
TCALT	°C/100m	(-inf,inf)	0.6	change of temperature with elevation	Height Increment Variables
Pelev	m	(-inf,inf)	0	elevation of precipitation data in the PTQ file	Height Increment Variables
Telev	m	(-inf,inf)	0	elevation of temperature data in the PTQ file	Height Increment Variables
PART	-	[0,1]	0.5	portion of the recharge which is added to groundwater box 1	Response Routine With Delay
DELAY	Δt	[0,inf)	1	time period over which recharge is evenly distributed	Response Routine With Delay

Effect of T_T



Model Calibration

The calibration of the model is usually made by manual try and error technique (Bergström, 1992).

The coefficient of efficiency, R_{eff}, is normally used for assessment of simulations by the HBV model.

$$R_{eff} = 1 - \frac{\sum (Q_{Sim}(t) - Q_{Obs}(t))^2}{\sum (Q_{Obs}(t) - \overline{Q}_{Obs})^2}$$

Different criteria can be used to assess the fit of simulated runoff to observed runoff:

- visual inspection of plots with Qsim and Qobs
- accumulated difference
- statistical criteria

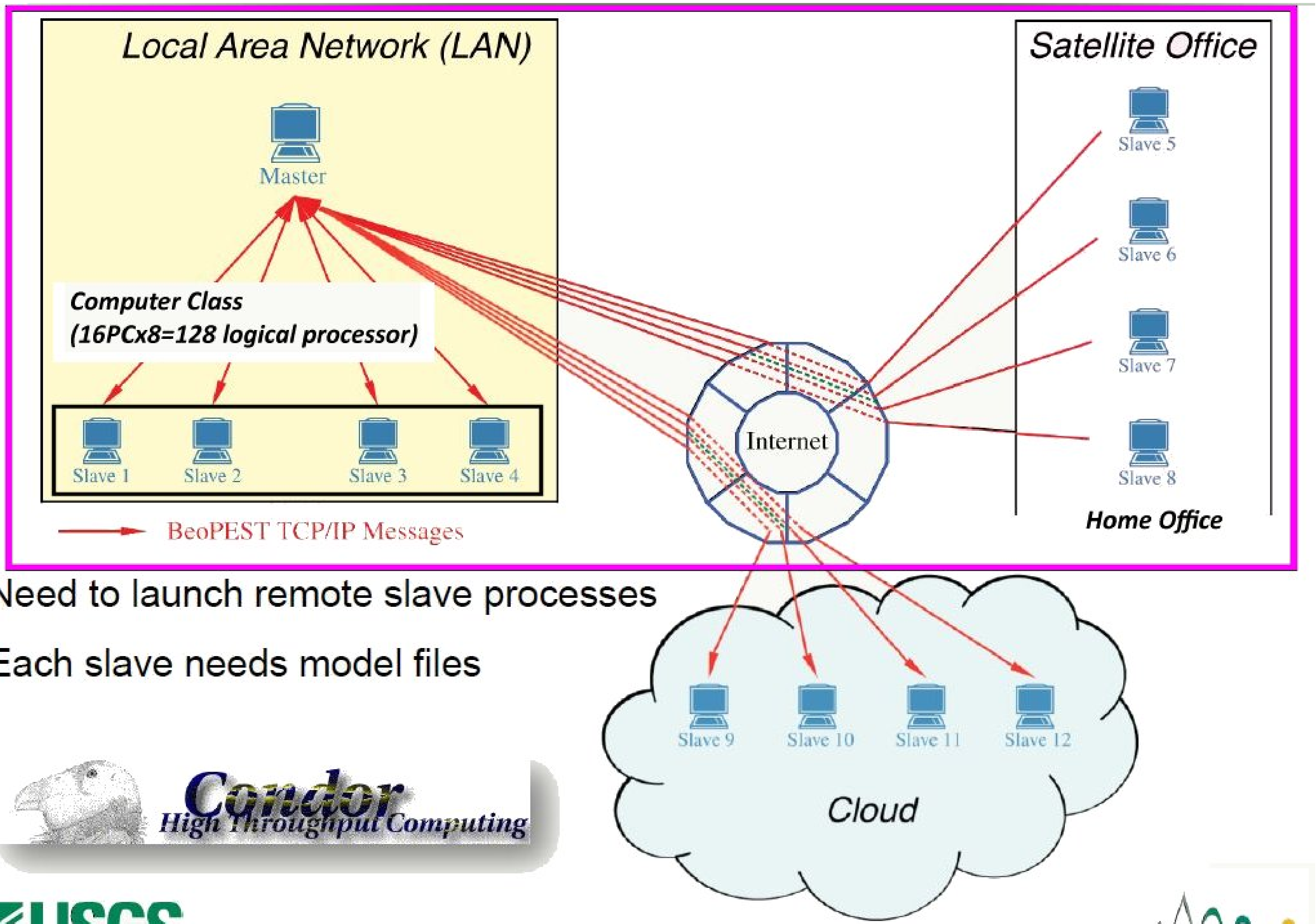
R_{eff} compares the prediction by the model with the simplest possible prediction, a constant value of the observed mean value over the entire period.

$R_{eff} = 1$	Perfect fit, QSim(t)= QObs(t)
$R_{eff} = 0$	Simulation as good (or poor) as the constant-value prediction
$R_{eff} < 0$	Very poor fit

Note:

- the calibration period should include a variety of hydrological events
- normally 5 to 10 years sufficient to calibrate the model
- validation: test of model performance with calibrated parameters for an independent period

BeoPEST is an tool for model calibration

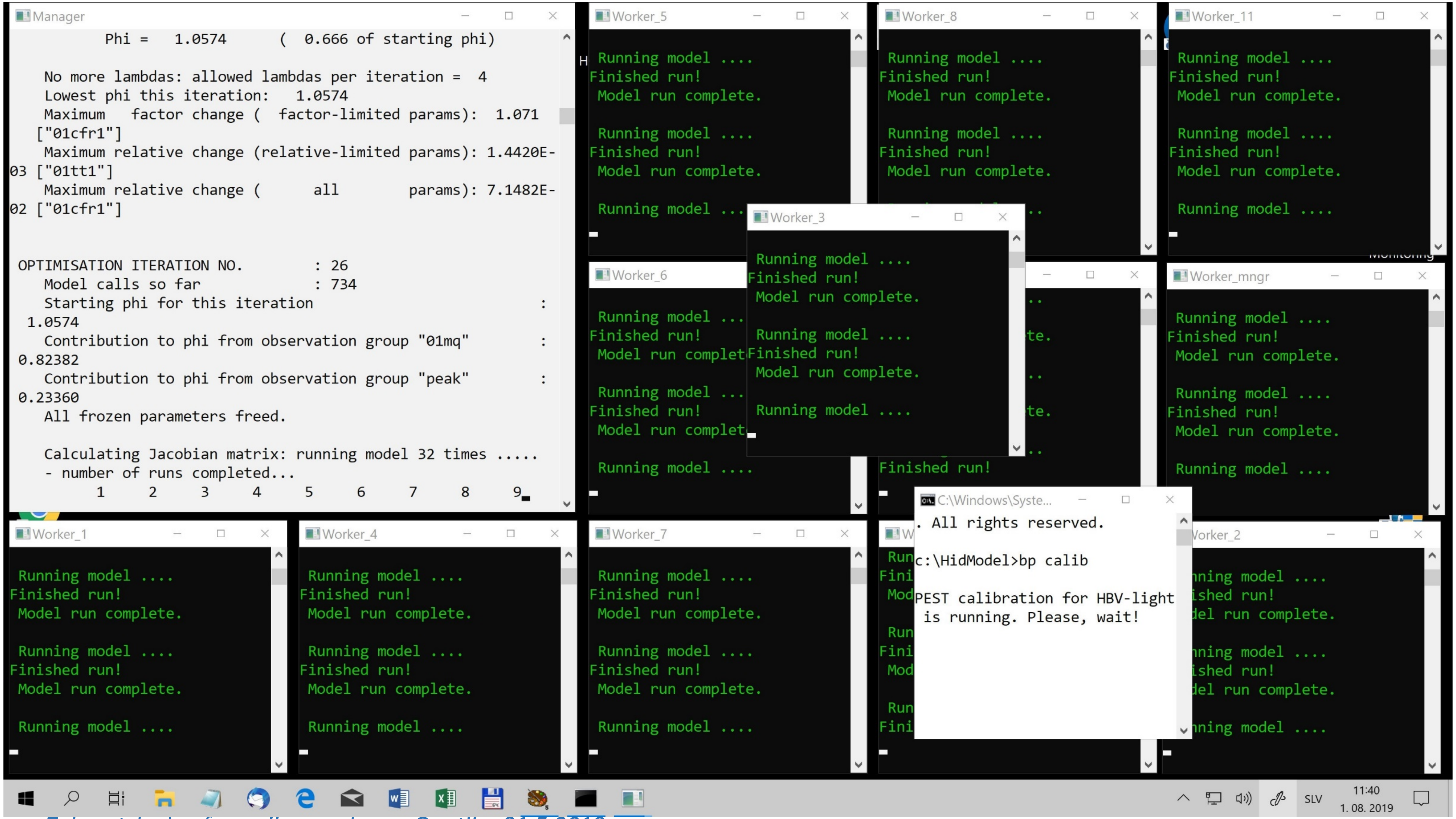


Need to launch remote slave processes

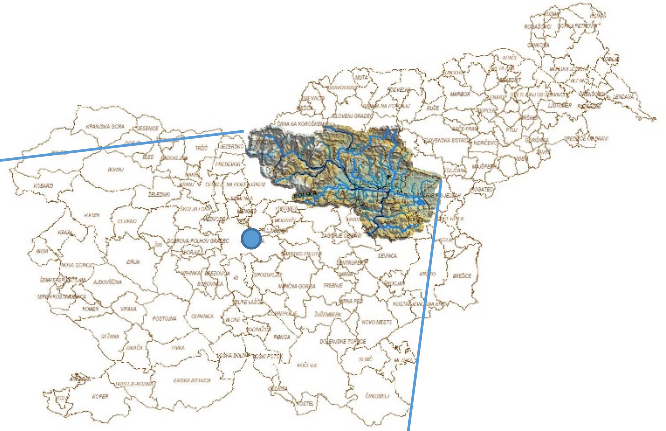
Each slave needs model files

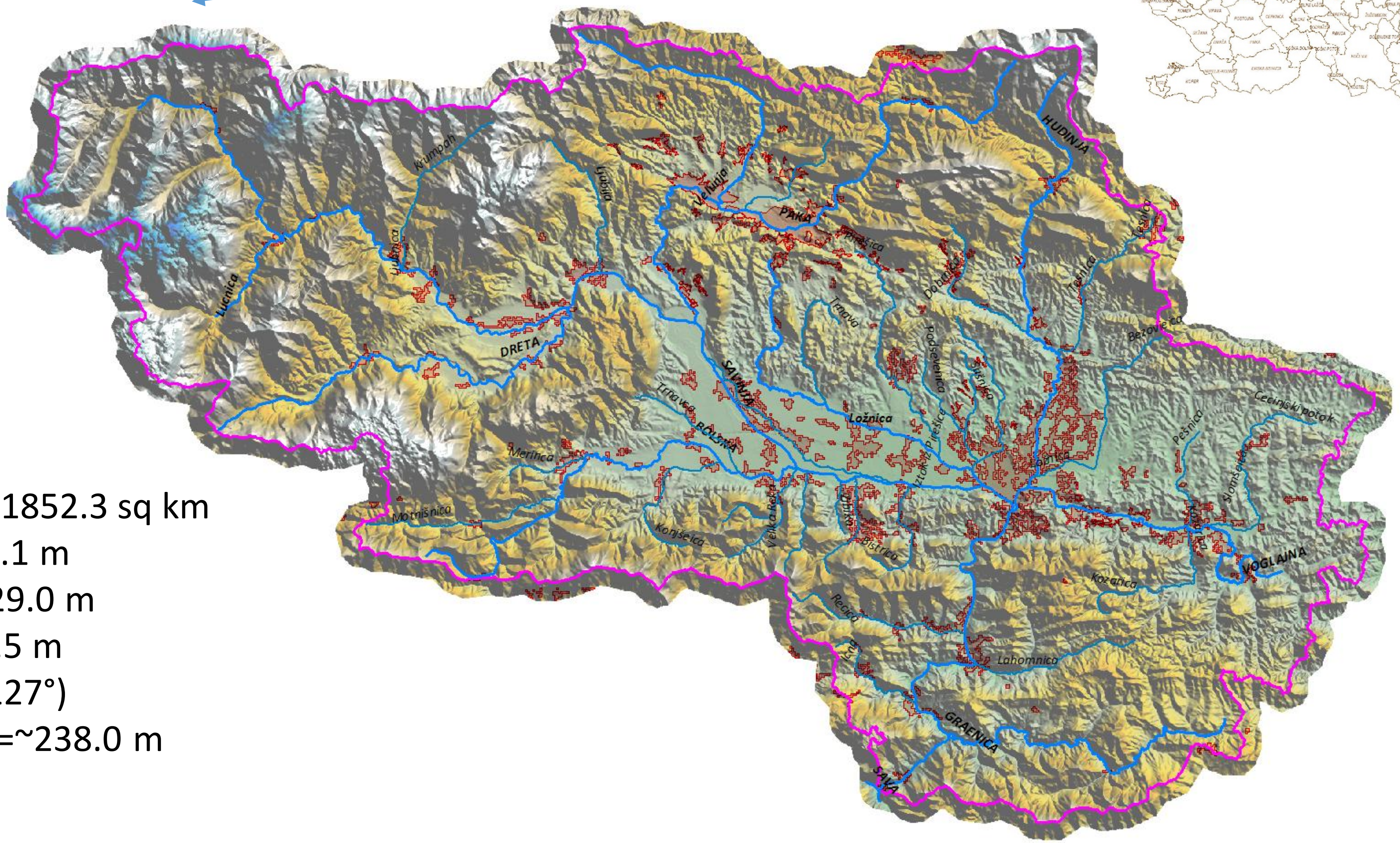
- **Calibration and Uncertainty Analysis for Complex Environmental Models**

BeoPEST – parallel processing (one manager and many workers)



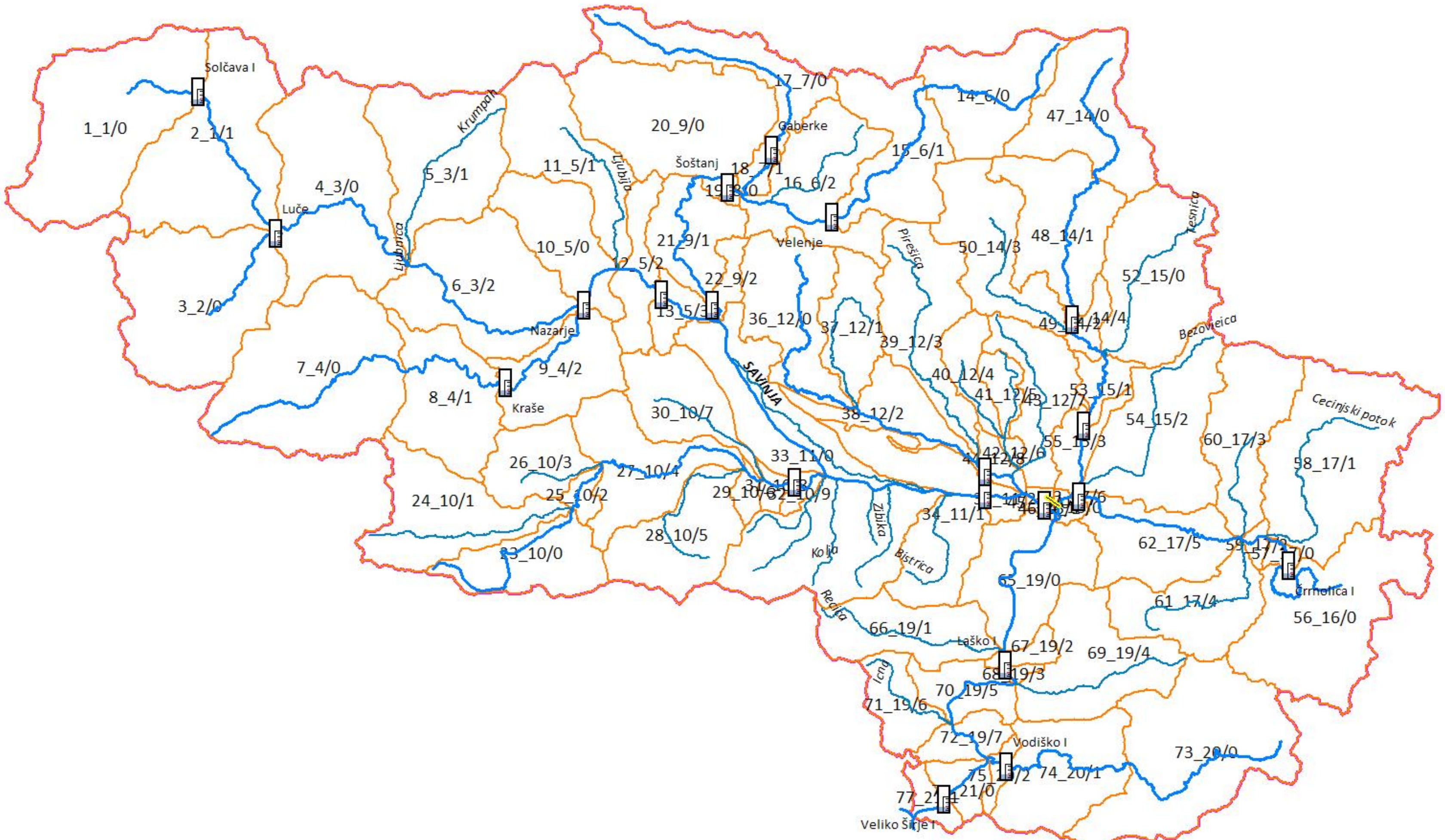
Savinja Catchment Case





- Enclosed Area of 1852.3 sq km
- min_Elev_m=190.1 m
- max_Elev_m=2429.0 m
- avg_Elev_m=604.5 m
- avg_Aspect=SE (127°)
- Older_Celje_elev=~238.0 m

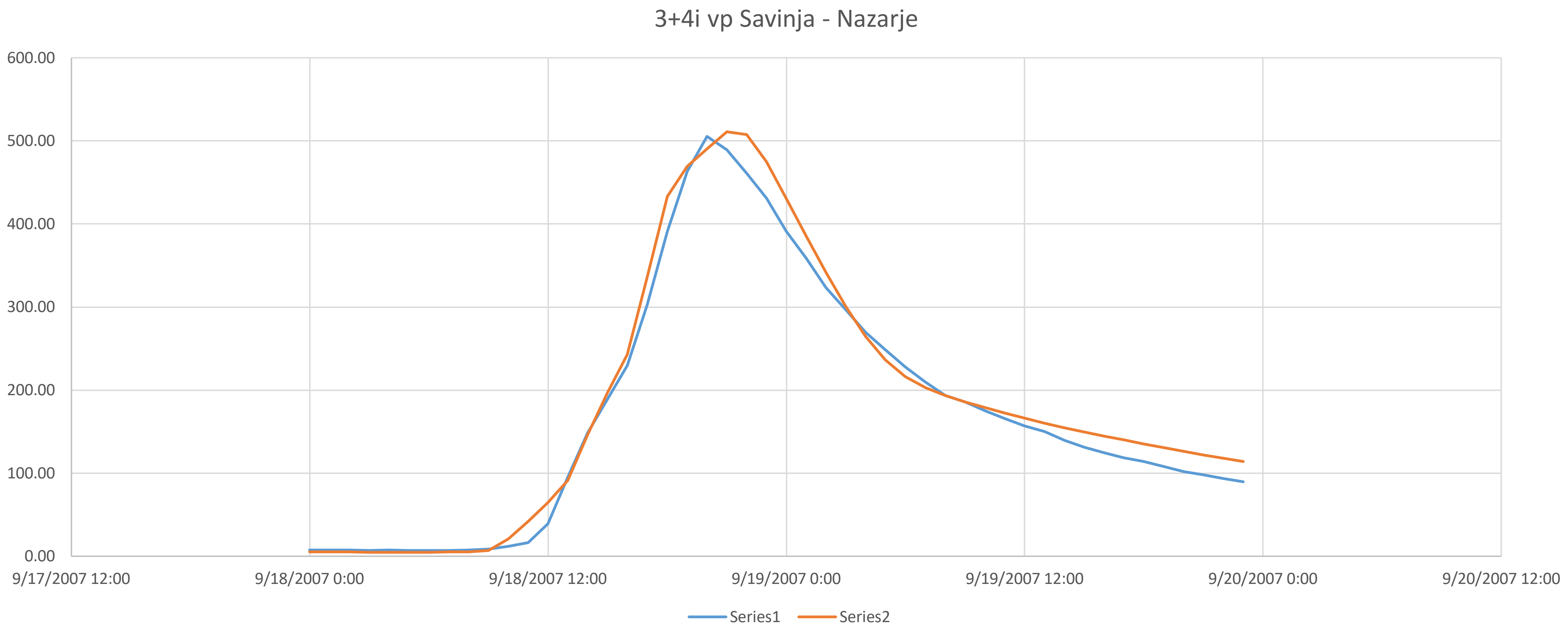
Savinja 77 sub-catchments; II. model



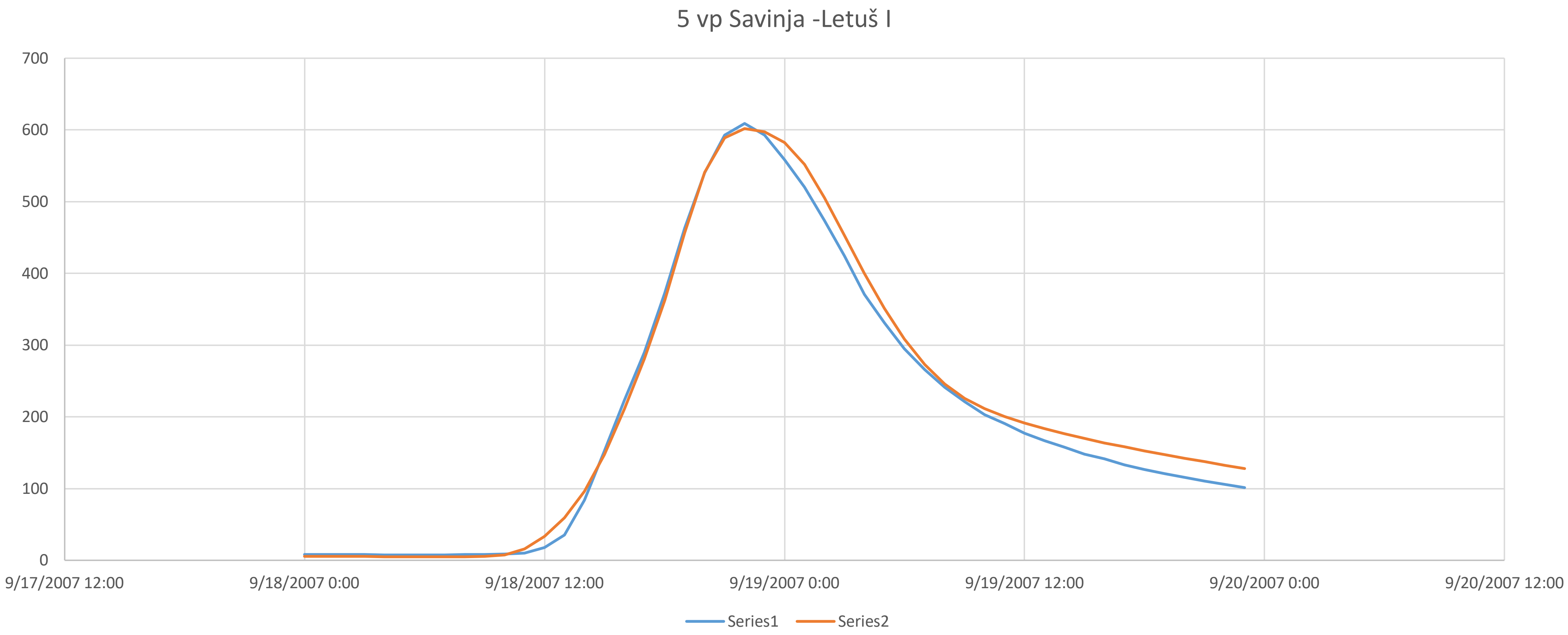
Goodnes of Fit for calibration period - year 2007

Average model efficiency of Savinja River to Sava outflow for whole calibration period 2007	**0.952**
Average model efficiency for 5-days **flood waves** 18.-22.09.2007	0.988

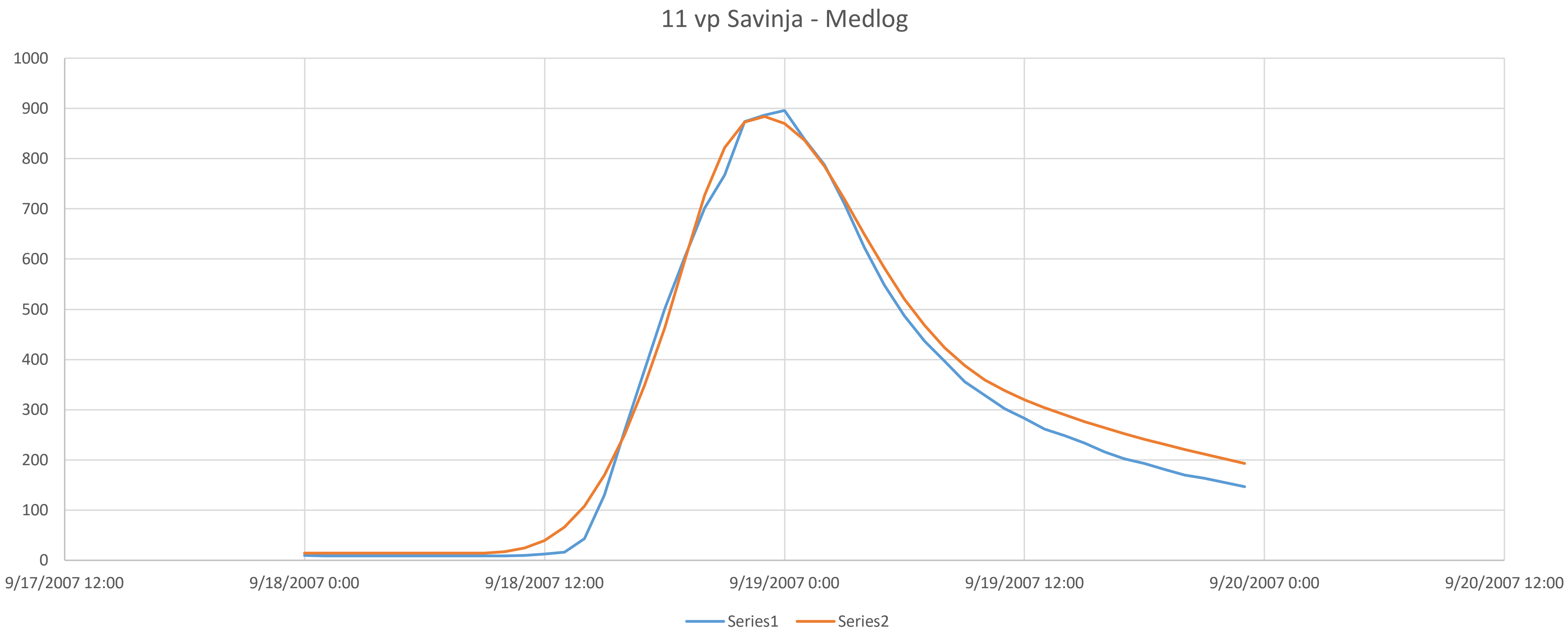
Savinja - vp Nazarje



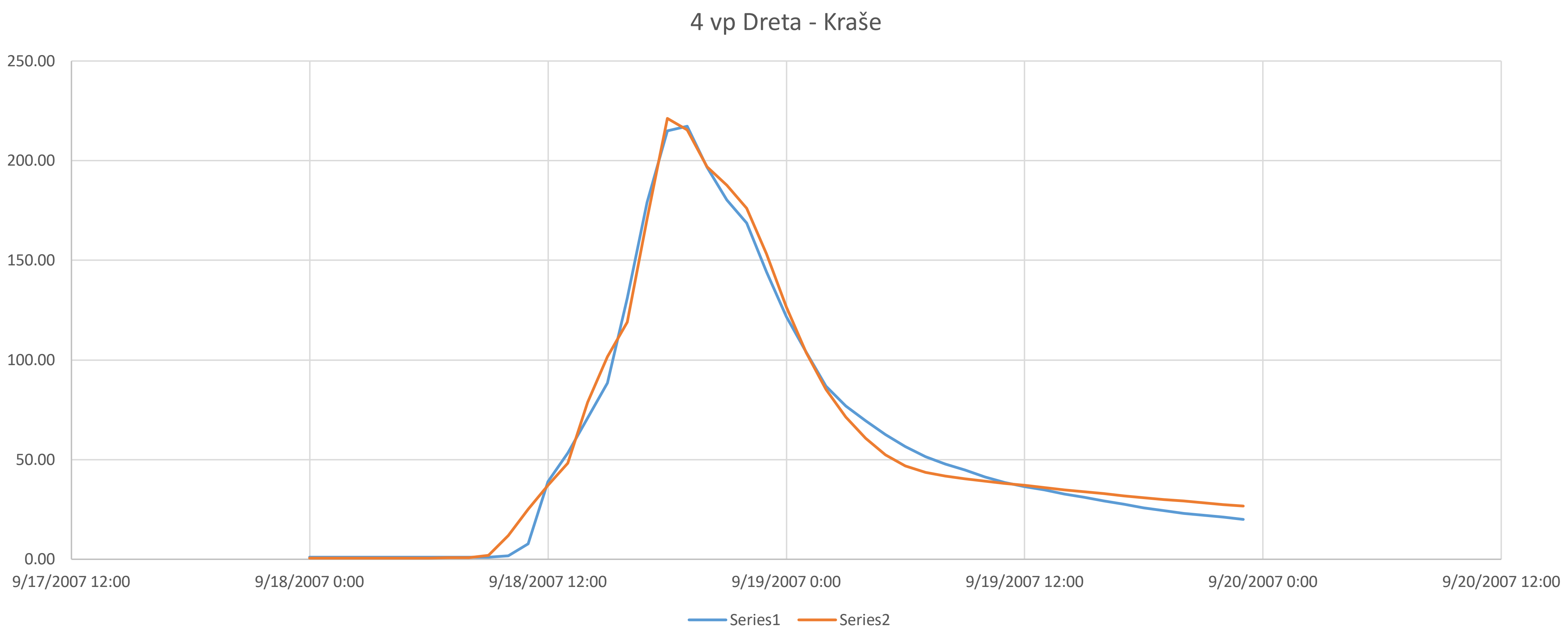
Savinja - vp Letuš 1



Savinja – vp Medlog



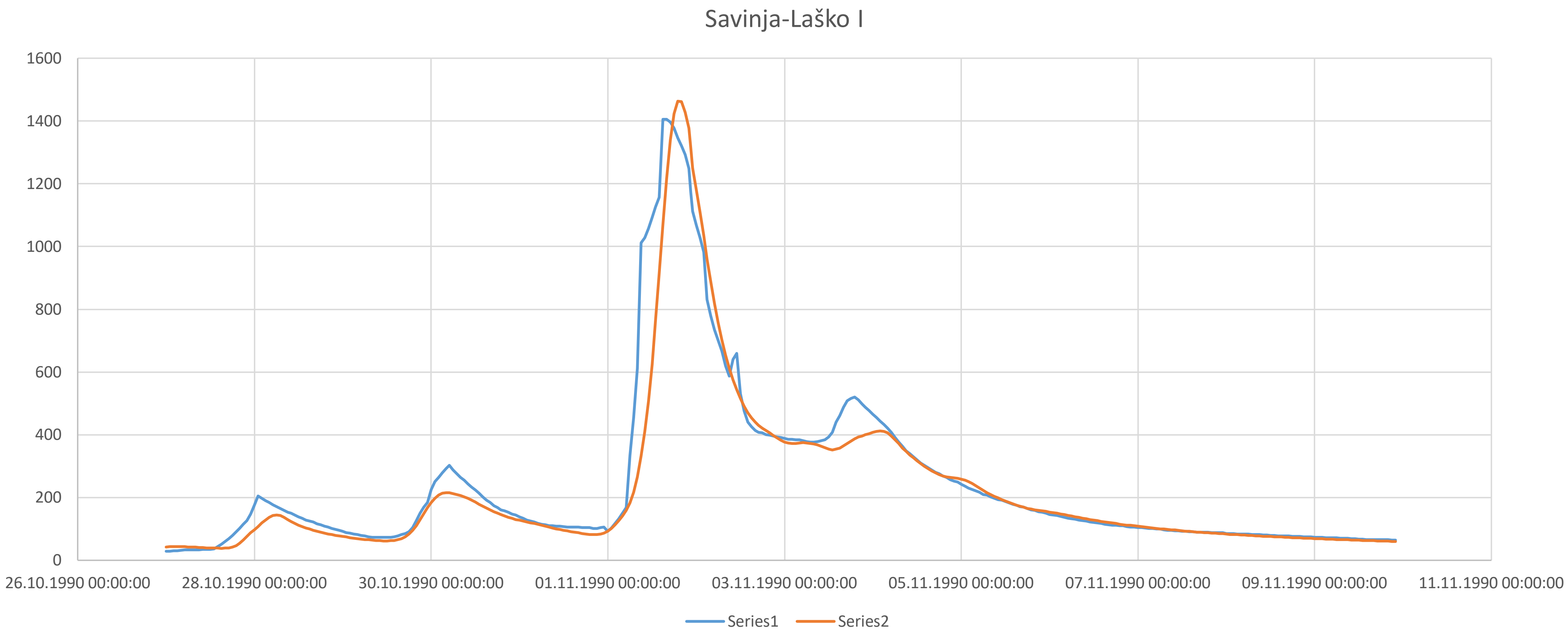
Dreta – vp Kraše



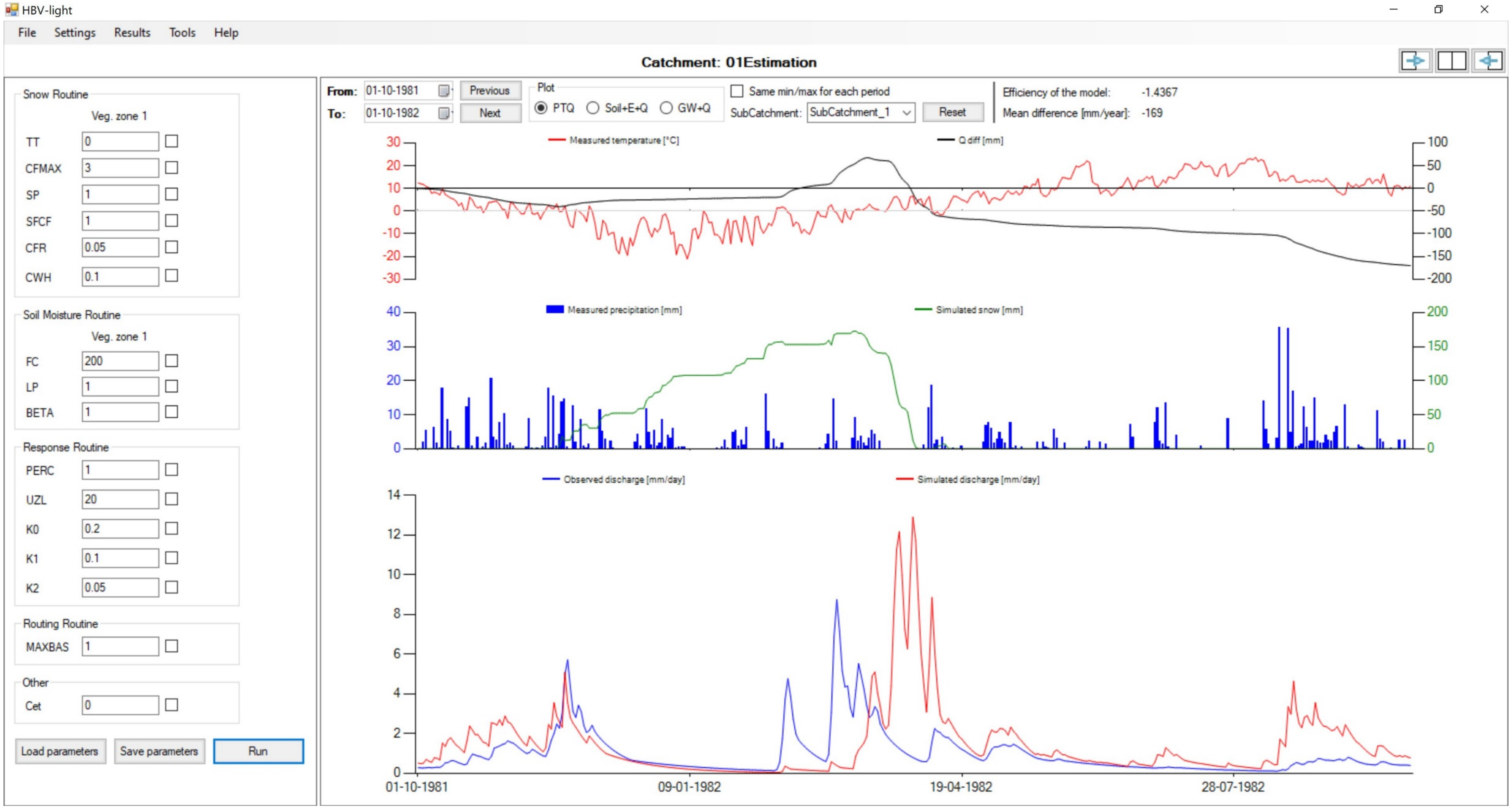
Goodnes of Fit for validation period - year 1990

WS2	WS2_Name	NS (1.10-14.11.1990)
1	Savinja do VP Solčava I	0.85
8	Dreta do VP Kraše	0.90
38	Ložnica do VP Levec I	0.94
45	Savinja do VP Celje II - brv	0.97
53	Hudinja do VP Škofja Vas	0.8
62	Voglajna do VP Celje II	0.8
67	Savinja do VP Laško	0.97
76	Savinja do VP Veliko Širje I	0.84

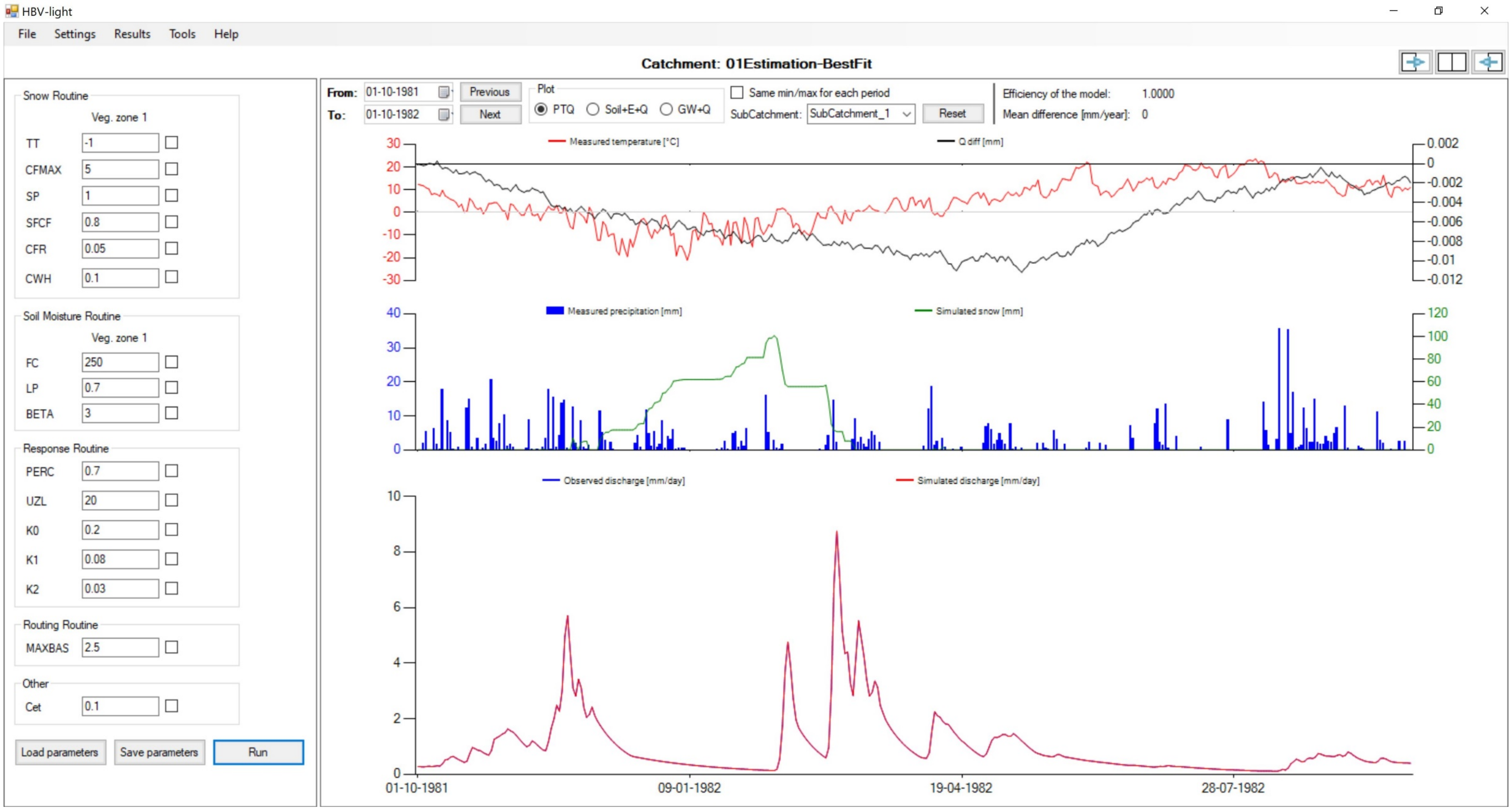
Savinja – vp Laško



Uncalibrated – eliminated measurements noise example



Calibrated


HBV-light
File Settings Results Tools Help
Catchment: 01Estimation-BestFit
Snow Routine
Veg. zone 1
TT -1
CFMAX 5
SP 1
SFCF 0.8
CFR 0.05
CWH 0.1
Soil Moisture Routine
Veg. zone 1
FC 250
LP 0.7
BETA 3
Response Routine
PERC 0.7
UZL 20
K0 0.2
K1 0.08
K2 0.03
Routing Routine
MAXBAS 2.5
Other
Cet 0.1
Load parameters
Save parameters
Run
From: 01-10-1981
To: 01-10-1982
Previous
Next
Plot
PTQ
Soil+E+Q
GW+Q
Same min/max for each period
SubCatchment: SubCatchment_1
Reset
Efficiency of the model: 1.0000
Mean difference [mm/year]: 0
Measured temperature [°C]
Q diff [mm]
Measured precipitation [mm]
Simulated snow [mm]
Observed discharge [mm/day]
Simulated discharge [mm/day]
01-10-1981
09-01-1982
19-04-1982
28-07-1982

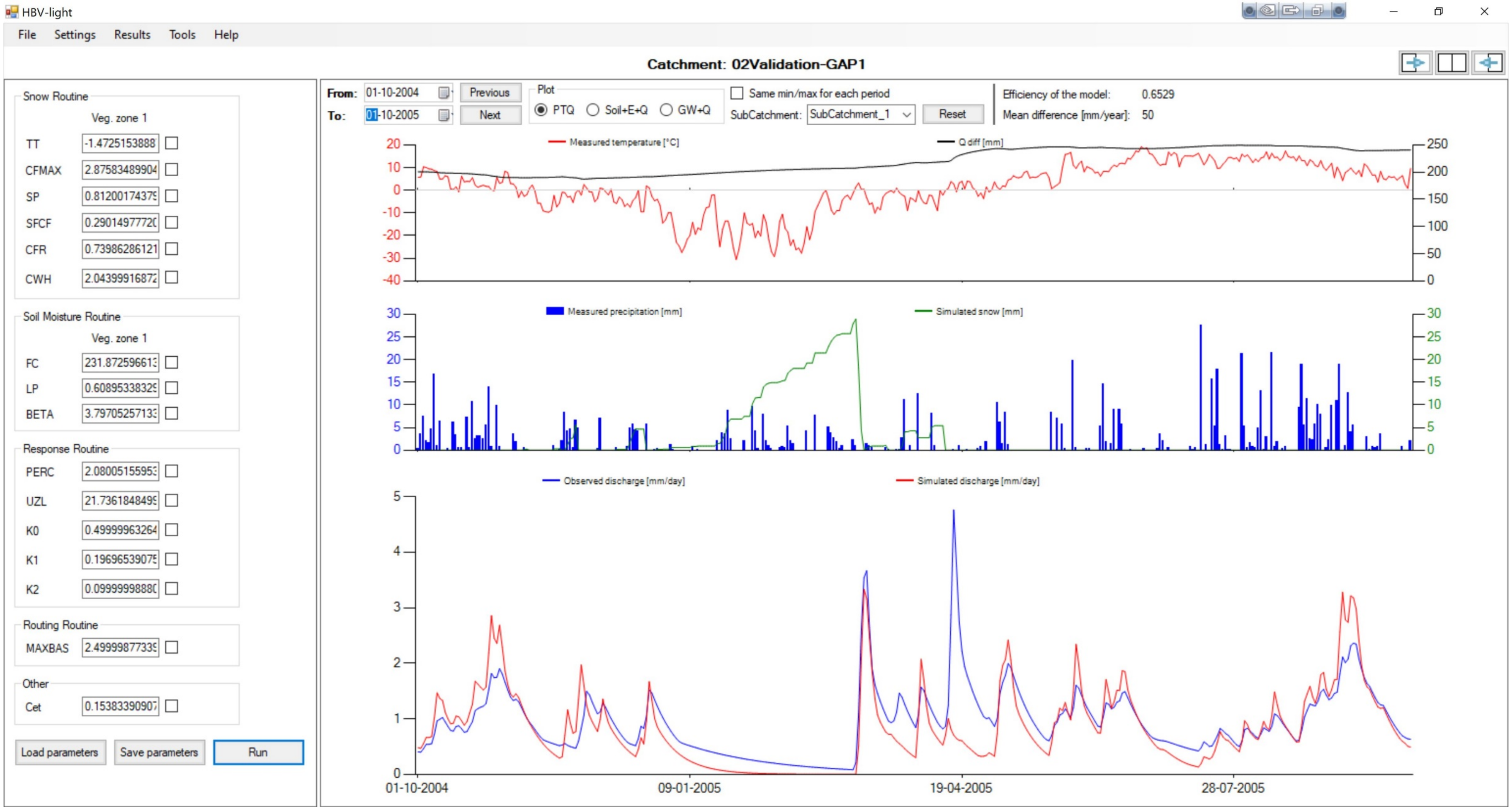
Efficiency

ESTIMATION	1.10.1981-31.12.1991	10 years period		
PEST_type	Uncalib	Perfect fit	BEOPEST	GAP
Coeff. of determination	0.20221114	0.99999986	0.99999985	0.98427475
Model efficiency	-1.43670698	0.99999986	0.99999985	0.98003139
Kling-Gupta efficiency	0.01548047	0.99999026	0.99998011	0.93638967
Efficiency for log(Q)	-0.40112962	0.99999888	0.99999887	0.94828924
Flow weighted efficiency	-0.26136485	0.99999997	0.99999996	0.98843674
Mean difference	-168.90532793	0.00282528	0.00382001	18.47635686
Efficiency for peak flows	-2.61414165	0.99999998	0.99999997	0.98586698
Volume Error	0.42853933	0.99999044	0.99998708	0.93748858
VALIDATION	1.10.2001-31.12.2011	10 years period		
PEST_type	Uncalib	Perfect fit	BEOPEST	GAP
Coeff. of determination	0.13979933	0.99999981	0.99999979	0.94799054
Model efficiency	-1.43126502	0.99999979	0.99999979	0.94790032
Kling-Gupta efficiency	0.01823037	0.99986979	0.99992483	0.95712343
Efficiency for log(Q)	-1.77575874	0.99999750	0.99999748	0.97555861
Flow weighted efficiency	-0.14028966	0.99999985	0.99999984	0.96343725
Mean difference	-121.24852609	-0.00259352	0.00252695	-2.30084655
Efficiency for peak flows	-1.99882597	0.99999952	0.99999954	0.93617965
Volume Error	0.49998999	0.99998930	0.99998958	0.99051167

Proper parameter estimation to best fit

		Estimation values			Absolute deviation from perfect fit [%]		
Parameter	**Perfect fit**	**BEOPEST**	**GAP1**	**GAP2**	**BEOPEST**	**GAP1**	**GAP2**
01perc	0.7	0.699983	2.08005156	0.75404974	0.00	197.15	7.72
01uzl	20	20.0007	21.73618485	19.3250461	0.00	8.68	3.37
01k0	0.2	0.2	0.499999633	0.24617666	0.00	150.00	23.09
01k1	0.08	0.08	0.196965391	0.09690948	0.00	146.21	21.14
01k2	0.03	0.03	0.099999989	0.03421317	0.00	233.33	14.04
01mxbs	2.5	2.499691	2.499998773	2.4388594	0.01	0.00	2.45
01cet	0.1	0.099997	0.153833909	0.13857043	0.00	53.83	38.57
01tt1	-1	-0.9998	-1.47251539	-1.18543094	0.02	47.25	18.54
01cfm1	5	5	2.875834899	4.93432181	0.00	42.48	1.31
01sp1	1	1	0.812001744	0.76142898	0.00	18.80	23.86
01sfc1	0.8	0.799951	0.290149777	0.69505231	0.01	63.73	13.12
01cfr1	0.05	0.049959	0.739862861	3.50621959	0.08	1379.73	**6912.44**
01cwh1	0.1	0.09997	2.043999169	0.08574851	0.03	1944.00	14.25
01fc1	250	250.015	231.8725966	239.820326	0.01	7.25	4.07
01lp1	0.7	0.700081	0.608953383	0.64380523	0.01	13.01	8.03
01bet1	3	3.000487	3.797052571	3.0969191	0.02	26.57	3.23

GAP validation



Conclusions

- using of PEST tools - Gauss-Marquardt-Levenberg algorithm with support included:
  - SVD – Singular value decomposition and
  - Tikhonov regularization

  we can successfully estimate thousands of parameters with minimum error variance to objective function.

- using of PEST tools we significantly decrease calibration noise