LIFE ENVIRONMENT STRYMON

Ecosystem Based Water Resources Management to Minimize Environmental Impacts from Agriculture Using State of the Art Modeling Tools in Strymonas Basin

LIFE03 ENV/GR/000217

Task 1. Strymonas Basin Integrated Surface Water – Ground Water Model

Technical Report
"Strymonas Basin Integrated Surface Water & Groundwater Model – Phase II. Calibration, Validation and Sensitivity Analysis of the Model"

THE GOULANDRIS NATURAL HISTORY MUSEUM
GREEK BIOTOPE / WETLAND CENTRE

The present work is part of the 4-years project: "*Ecosystem Based Water Resources Management to Minimize Environmental Impacts from Agriculture Using State of the Art Modeling Tools in Strymonas Basin*" (contract number LIFE03 ENV/GR/000217). The project is co-funded by the Goulandris Natural History Museum - Greek Biotope/Wetland Centre (EKBY), the Prefecture of Serres – Directorate of Land Reclamation of Serres (DEB-S), the Development Agency of Serres S.A. (ANESER S.A.) and the Local Association for the Protection of Lake Kerkini (SPALK).

This document may be cited as follows:
Halkidis, I. and D. Papadimos. 2007. Technical Report. Strymonas basin integrated surface water and groundwater model – Phase II. Calibration, validation and sensitivity analysis of the model. Greek Biotope/Wetland Centre (EKBY). Thermi, Greece. 13 p.

CONTENTS

1. Model Calibration and Impact Analysis on Water Resources from Agricultural Activities

Distributed hydrological models are structure to enable the special variation in catchment characteristics to be represented by providing data for a network of grid points. Often model applications required several thousands of grid points, each of which characterized by several parameters and variables. In this way distributed models differ fundamentally from lumped models, where a catchment is considered as one unit characterized by, typically, a few tens of parameters and variables. Thus the number of parameters and variables in a distributed model is, in principle, often two or three orders of magnitude higher than it would be for a lumped model of the same area. Obviously, this generates different requirements to lumped and distributed models with regard to parameterisation, calibration and validation procedures.

2. Model parameterisation and choice of calibration parameters

A distributed hydrological modelling system such as MIKE SHE potentially involves a large number of model parameters to be specified by the user during the model setup. Some of these parameters may be assessed from field data, e.g. geological descriptions from well-logs, pumping test analysis, maps of soil profiles, soil analysis (texture, density, retention curves), and vegetation maps. Comprehensive field data, however, are seldomly available to fully support specification of all model parameters. In addition, some model parameters are of a more conceptual nature and cannot be directly assessed from field data.

In the model parameterisation, the available field data should be used to define the spatial patterns of the parameter values to describe the most significant variations. This is often done by defining a conceptual model with appropriate parameter classes of geological units, soil types, vegetation types etc. For each class, some parameters are then assessed directly from field data while other parameters may be subject to calibration. The challenge is to formulate a relatively simple model parameterisation in order to provide a well-posed calibration problem but at the same time keep it sufficiently complex in order to capture the spatial variability of key model parameters.

Refsgaard and Storm (1996) emphasize that a rigorous parameterisation procedure is crucia1 in order to avoid methodologica1 problems in the subsequent phases of mode1 calibration and validation. In parameterisation, the spatial patterns of the parameter va1ues are defined so that a given parameter only reflects the significant and systematic variation described in the available field data, as exemplified by the practice of using representative parameter va1ues for individua1 soi1 types, vegetation types or geological layers. Thus the parameterisation process effectively reduces the number of free parameters coefficients which need to be adjusted in the subsequent calibration procedure. The following points are important to consider in the parameterisation procedure (Refsgaard and Storm, 1996).

- The parameter classes (soil types, vegetation types, climatological zones, geo-logical layers, etc.) should be selected so that it becomes easy, in an objective way, to associate parameter values. Thus the parameter va1ues in the different classes should, to the highest possible degree, be assessable from available field data.
- It should explicitly be evaluated which parameters can be assessed from field data alone and which need some kind of calibration. For the parameters subject to calibration, physical1y acceptable intervals for the parameter values should be estimated.
- The number of real calibration parameters should be kept low, both from practical and methodological points of view. This can be done, for instance, by fixing a spatia1 pattern of a parameter but allowing its absolute value to be modified through calibration.

Refsgaard in his paper (Refsgaard, 1997), subdivided a catchment area of 440 km^2 into grid squares of 500 X 500 m^2 and Vazquez and Feyen, (Vazquez and Feyen, 2003) subdivided a catchment area of 586 km^2 with a resolution of 600 X 600 m^2.
Sensitivity tests were made by Refsgaard in his paper (Refsgaard, 1997) using the model (calibrated and validated on a 500 m grid) on coarser grids: 1000, 2000 and 4000 m. The results of the four model-simulations of discharge for the entire catchment are shown that the 500 m and the 1000 m models only differ marginally.
At the project Life Strymon the model domain area of 1510 km^2 subdivided in grids

with a resolution of 400 X 400 m^2. The smaller subdivision for larger area than the combinations derived from the literature gives that there is no necessity for the grid size calibration.

The elevation of the area varies approximately from a minimum of 0 m in the southern part to a maximum of 100 m in the surrounding area. The profiles definition of the river tributaries was based on interpolation/extrapolation of a few measured profiles. Roughness coefficients were based on values from literature for rivers.
Drainage starts when the water table rises above the elevation of the drains and is proportional to the difference in level between the water table and the drainage depth. The drainage depth of 1,5 m was derived from LRA of Serres as the most appropriate for the drain area. Also, Sahoo et.al. (2005) used drainage depth of 1 m. So the drainage depth of 1,5 m decided not to be included in the calibration process. The drainage depth has more influence on the recession of the hydrograph (Refsgaard, 1997).
On the other hand the drainage time constant or drainage coefficient was calibrated. This parameter determines the velocity of the drainage and mostly influences the peak of the hydrograph (Feyen et al., 2000).

Model performance in the calibration and validation process can be evaluated both qualitatively, based on visual graphical techniques, and quantitatively, based on some statistical measures. In this study, both methods were combined but with emphasis on the statistical appreciation of the model performance. A first idea of the accuracy of the results was based on visual inspection of their graphical representation; then statistical parameters for the simulations were calculated. The statistical criteria used in the analysis are the Relative Root Mean Square Error (RRMSE), the Coefficient of Determination (CD), the Coefficient of Efficiency (EF) and the Mean Absolute Error (ABSERR) and can be depicted by the following expressions (Feyen et al. 2000).

$$\text{RRMSE} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - Q_i)^2}{n}} = \frac{1}{\bar{0}} \ldots \qquad 0 \le RRMSE$$

$$\text{ABSERR} = \frac{\sum_{i=1}^{n}|O_i - P_i|}{n} \qquad 0 \le ABSERR$$

$$CD = \frac{\sum_{i=1}^{n}(O_i - \overline{O})^2}{\sum_{i=1}^{n}\left(P_i - \overline{O}\right)^2} \qquad 0 \prec CD \leq +\infty$$

$$EF = \frac{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2 - \sum_{i=1}^{n}\left(P_i - \overline{O}\right)^2 -}{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2 -} \qquad -\infty \prec EF \leq 1$$

Where Pi is the i-th simulated value, Oi is the i-th observed $\overline{O}$ is the average of the observed values, n is the number of observations in the consider period.

The water mass balance in the Strymonas basin, except from the surface water, depends also from the evaporo-transpiration and the water exchange between the surface and ground water bodies.The established monitoring network in Strymonas basin gives the appropriate measurements in order to avoid uncertainties and set the right parameters for the surface water bodies. The lake of measurements for the groundwater and evaporo-transpiration parameters drives to the modeling techniques of sensitivity analysis, calibration and validation of those model parameters.

The main surface water body which exchange water with the groundwater body of the Strymonas basin is Lake Kerkini. So the basic parameter for calibration was the leakage coefficient of the lake Kerkini. For this parameter used the try and error method. The values of leakage coefficient that were tested are $8*10^{-8}$, $1*10^{-7}$, $3*10^{-7}$, $4*10^{-7}$ and $8*10^{-7}$.

The outflow from Lakes Kerkini’s dam was simulated in order to keep the water level of the lake as it measured.

The amount of leakage discharge for the values $8*10^{-7}$ and $4*10^{-7}$ was large enough and the water level of the lake was lower than the measured, as shown at the figures 5. On the other head the leakage coefficient values off $8*10^{-8}$, $1*10^{-7}$ and $3*10^{-7}$ following very steeped way the water level of the lake as shown at the figures 5 and the statistic calculations that MIKE SHE produce.

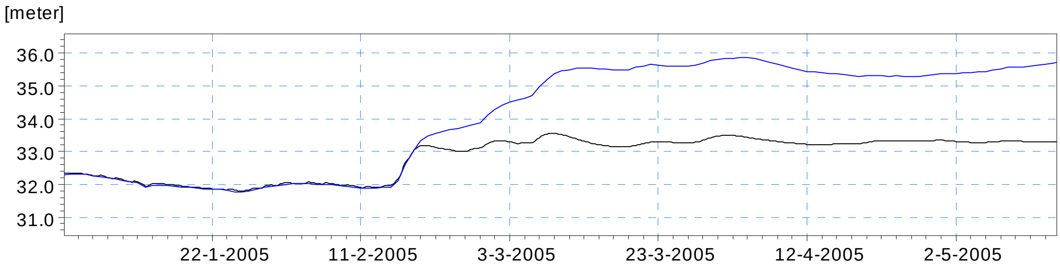

Figure 1. Application of the model with Leakage Coefficient = $8*10^{-6}$ (black line), observed time series (blue line).

Statistical measures of the application with Leakage Coefficient = $8*10^{-6}$

ME = 1.22567

MAE = 1.24836

RMSE = 1.61399

STDres = 1.05009

R (Correlation)=0.957414

R2 (Nash Sutcliffe) = 0.00874

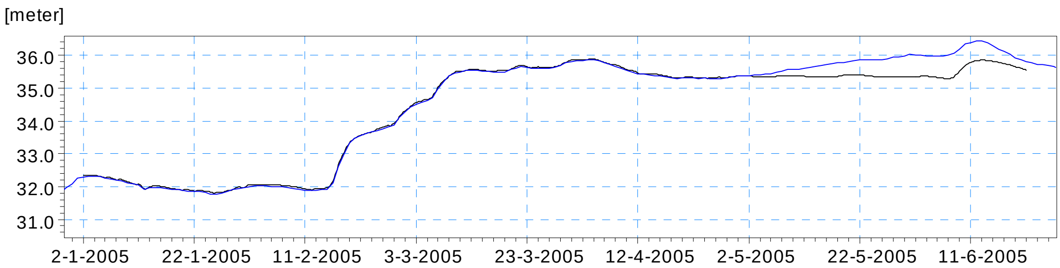

Figure 2. Application of the model with Leakage Coefficient = $4*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with Leakage Coefficient = $4*10^{-7}$

ME = 0.106571

MAE = 0.149095

RMSE = 0.259739

STDres = 0.236869

R (Correlation)=0.991283

R2 (Nash Sutcliffe) = 0.974887

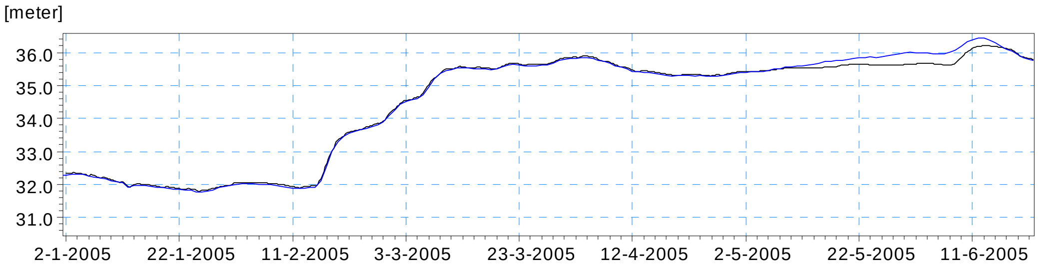

Figure 3. Application of the model with Leakage Coefficient = $3*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with Leakage Coefficient = $3*10^{-7}$

ME = 0.00770456

MAE = 0.0423322

RMSE = 0.0584749

STDres = 0.0579651

R (Correlation)=0.999568

R2 (Nash Sutcliffe) = 0.998809

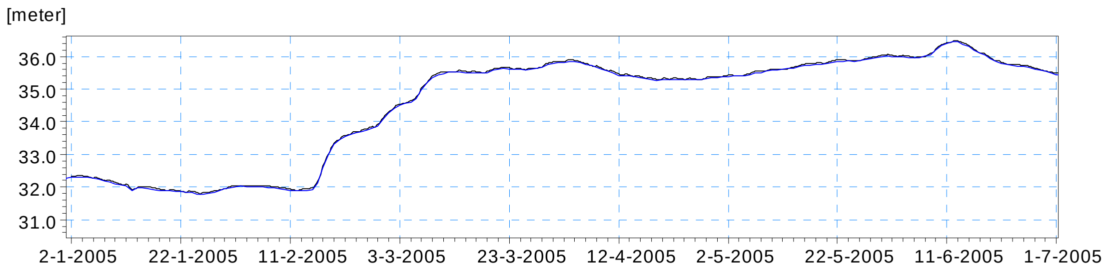

Figure 4. Application of the model with Leakage Coefficient = $8*10^{-8}$ (black line), observed time series (blue line).

Statistical measures of the application with Leakage Coefficient = $8*10^{-8}$

ME = 0.0029944

MAE = 0.030518

RMSE = 0.034401

STDres = 0.016940

R (Correlation)=0.99996

R2 (Nash Sutcliffe) = 0.99959

Very important information that the local authorities obtain was that after twentieth of June summer period there is no outflow discharge from the dam in order to keep the water in the lake for irrigation purpose. This tip helped to reject the leakage coefficient values of $8*10^{-8}$and $1*10^{-7}$ because with this values outflow discharge occurred downstream of Lake Kerkini's Dam as shown at the figure 5.

So the most appropriate value for the leakage coefficient in the lake Kerkini is the value of $3*10^{-7}$.

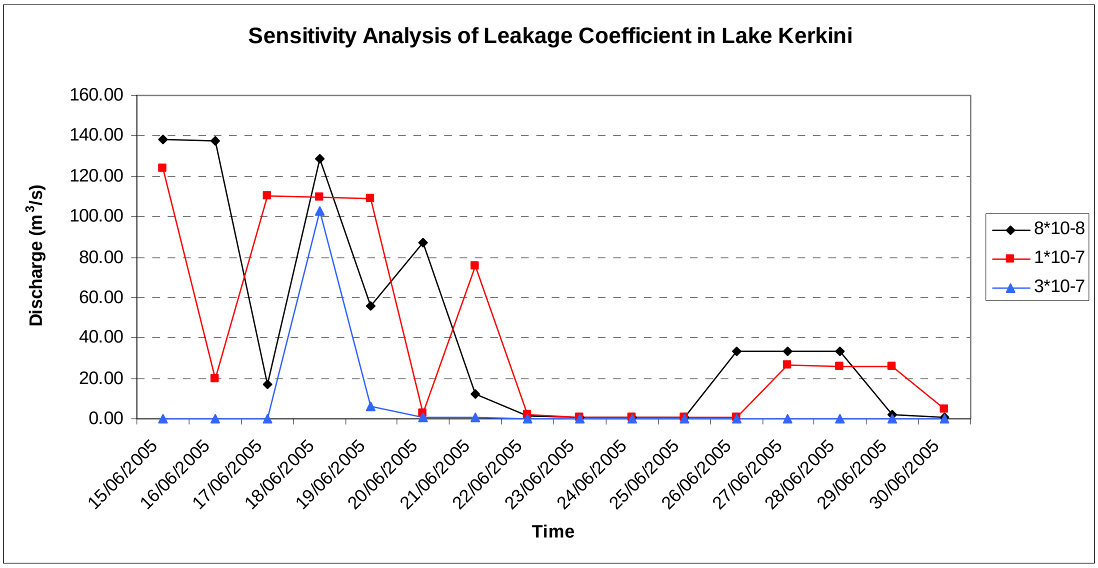

Figure 5. Discharge occurred downstream of Lake Kerkini's Dam

Model calibration by sensitivity analysis of the drainage coefficient for the cultivated areas.

An other parameter that has calibrated is the drainage time constant or drainage coefficient. This parameter determines the velocity of the drainage and mostly influences the peak of the hydrograph as mentioned by Feyen (Feyen et al., 2000) and Vazquez (Vazquez et al., 2002).

The values of drainage coefficient that were tested are $7*10^{-8}$, $1*10^{-7}$, $3*10^{-7}$, $5*10^{-7}$ and $7*10^{-7}$. The most important period for the calibration is when the irrigation starts. So there are two statistical periods one for the irrigation period and one for the hole period.

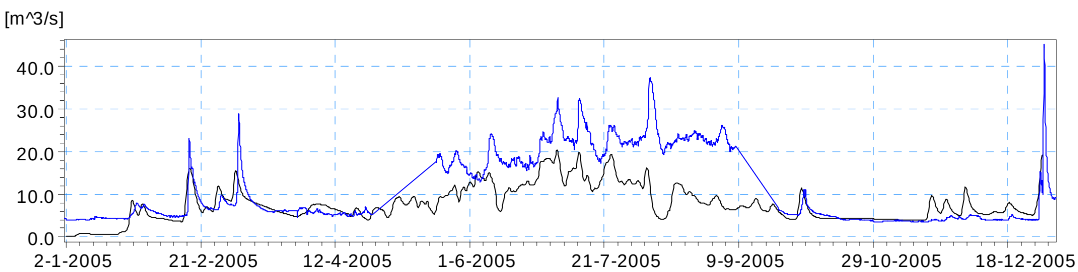

Figure 6. Application of the model with drainage Coefficient = $7*10^{-8}$ (black line), observed time series (blue line).

Statistical measures of the application with drainage Coefficient = $7*10^{-8}$

23/5/2005 - 30/7/2005	
Correlation coeficient R2	**0.505**
Max. positive difference	10.624 m^3/s
Max. negative difference	-28.512 m^3/s
Volume observed	357554428.4 M^3
Volume modelled	244630070.8 M^3
Volume error	**-31.582** %
Peak observed value	45.284 m^3/s
Peak modelled value	20.339 m^3/s
Peak error	-55.085 %

1/1/2005 - 31/12/2005	
Correlation coeficient R2	**0.538**
Max. positive difference	9.158 m^3/s
Max. negative difference	-28.512 m^3/s
Volume observed	3.51E+08 M^3
Volume modelled	2.37E+08 M^3
Volume error	**-32.557** %
Peak observed value	37.28 m^3/s
Peak modelled value	20.339 m^3/s
Peak error	-45.441 %

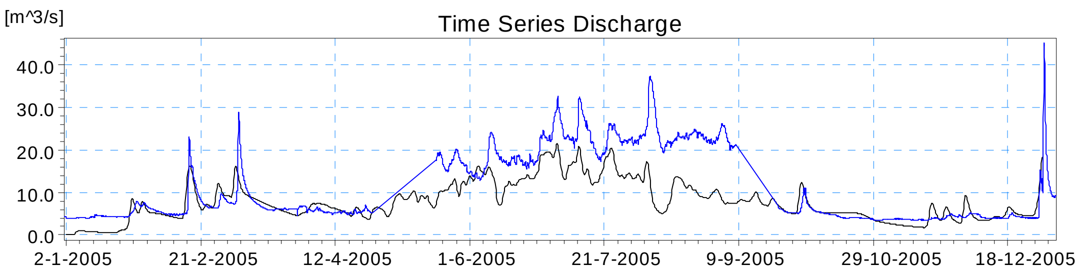

Figure 7. Application of the model with drainage Coefficient = $1*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with drainage Coefficient = $1*10^{-7}$

23/5/2005 - 30/7/2005	
Correlation coeficient R2	**0.365**
Max. positive difference	2.979 m^3/s
Max. negative difference	-13.573 m^3/s
Volume observed	119448473.2 M^3
Volume modelled	85072134.46 M^3
Volume error	**-28.779** %
Peak observed value	32.687 m^3/s
Peak modelled value	21.461 m^3/s
Peak error	-34.345 %

1/1/2005 - 31/12/2005	
Correlation coeficient R2	**0.627**
Max. positive difference	8.709 m^3/s
Max. negative difference	-27.543 m^3/s
Volume observed	3.51E+08 M^3
Volume modelled	2.43E+08 M^3
Volume error	**-30.599** %
Peak observed value	37.28 m^3/s
Peak modelled value	21.461 m^3/s
Peak error	-42.434 %

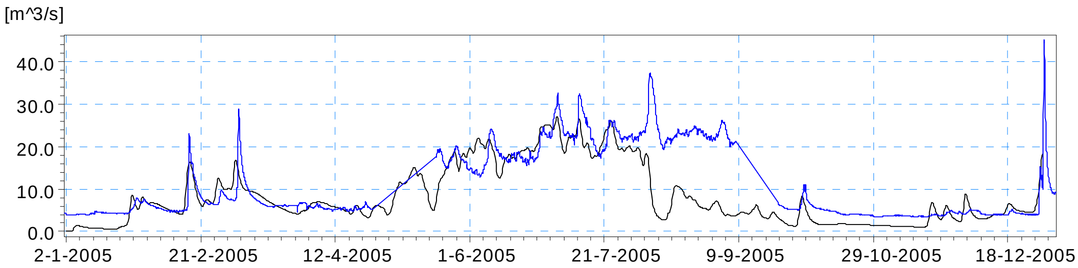

Figure 8. Application of the model with drainage Coefficient = $3*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with drainage Coefficient = $3*10^{-7}$

23/5/2005 - 30/7/2005	
Correlation coeficient R2	**0.345**
Max. positive difference	8.765 m^3/s
Max. negative difference	-8.05 m^3/s
Volume observed	119448473.2 M^3
Volume modelled	118079027.1 M^3
Volume error	**-1.146** %
Peak observed value	32.687 m^3/s
Peak modelled value	27.14 m^3/s
Peak error	-16.969 %

1/1/2005 - 31/12/2005	
Correlation coeficient R2	**0.531**
Max. positive difference	9.155 m^3/s
Max. negative difference	-28.881 m^3/s
Volume observed	3.51E+08 M^3
Volume modelled	2.53E+08 M^3
Volume error	**-27.939** %
Peak observed value	37.28 m^3/s
Peak modelled value	27.14 m^3/s
Peak error	-27.199 %

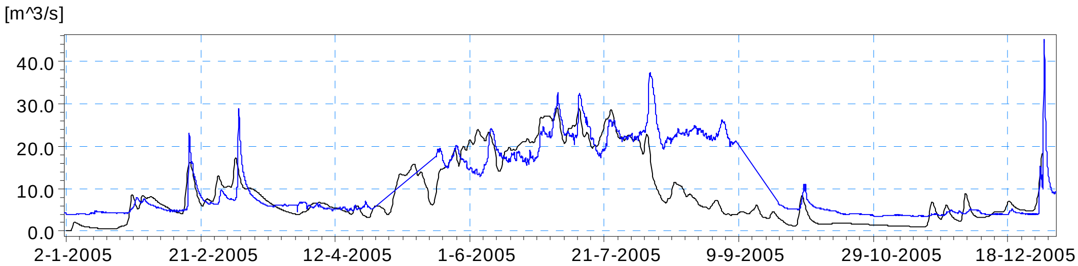

Figure 9. Application of the model with drainage Coefficient = $5*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with drainage Coefficient = $5*10^{-7}$

23/5/2005 - 30/7/2005	
Correlation coeficient R2	**0.859**
Max. positive difference	10.631 m^3/s
Max. negative difference	-14.699 m^3/s
Volume observed	211655523.8 M^3
Volume modelled	205677164.1 M^3
Volume error	**-2.825** %
Peak observed value	32.687 m^3/s
Peak modelled value	29.11 m^3/s
Peak error	-10.944 %

1/1/2005 - 31/12/2005	
Correlation coeficient R2	**0.588**
Max. positive difference	10.631 m^3/s
Max. negative difference	-22.721 m^3/s
Volume observed	3.51E+08 M^3
Volume modelled	2.74E+08 M^3
Volume error	**-21.978** %
Peak observed value	37.28 m^3/s
Peak modelled value	29.11 m^3/s
Peak error	-21.916 %

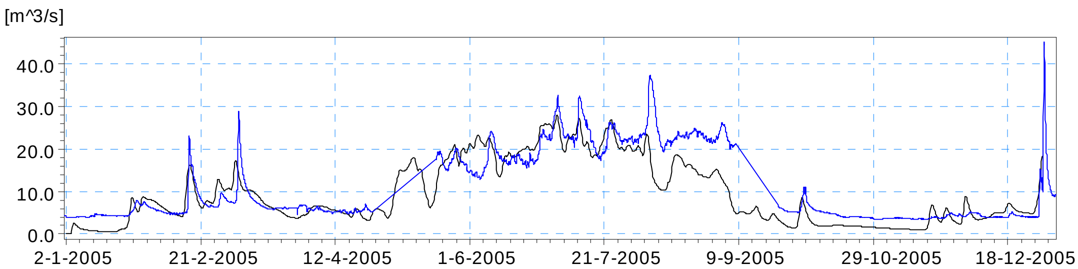

Figure 10. Application of the model with drainage Coefficient = $7*10^{-7}$ (black line), observed time series (blue line).

Statistical measures of the application with drainage Coefficient = $7*10^{-7}$

23/5/2005 - 30/7/2005	
Correlation coeficient R2	**0.302**
Max. positive difference	9.984 m^3/s
Max. negative difference	-7.17 m^3/s
Volume observed	119448473.2 M^3
Volume modelled	123954327.7 M^3
Volume error	**3.772** %
Peak observed value	32.687 m^3/s
Peak modelled value	27.92 m^3/s
Peak error	-14.582 %

1/1/2005 - 31/12/2005	
Correlation coeficient R2	**0.753**
Max. positive difference	9.984 m^3/s
Max. negative difference	-21.89 m^3/s
Volume observed	3.51E+08 M^3
Volume modelled	2.89E+08 M^3
Volume error	**-17.713** %
Peak observed value	37.28 m^3/s
Peak modelled value	27.92 m^3/s
Peak error	-25.106 %

The most appropriate value for the drainage Coefficient = $3*10^{-7}$ because gives the lowest volume error =-1.146% and because the most important for the simulation is to calculate the water mass balance.

References

Feyen, L., R. Vazquez, K. Christianes, O. Sels and J. Feyen. 2000. Application of a distributed physically-based hydrological model to a medium size catrchment. Hydrology & Earth System Sciences, 4(1): p.47-63.

Refsgaard, J.C., B. Storm. 1996. Construction, calibration and validation of hydrological models. In Abbott, M.B., Refsgaard, J.C. (Eds.), Distributed Hydrological Modelling. Kluwer Academic, pp.41-54.

Refsgaard, J.C. 1997. Parameterisation, calibration and validation of distributed hydrological models. Journal of Hydrology, 198: 69-97.

Sooroshian, S., and V.K. Gupta. 1995. Model Calibration. In: Computer Models of Watershed Hydrology, V.P. Singh, (Ed). Water Resources Publications, Colorado, USA, 23-68.

Vazquez, R.F., and J.Feyen. 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. Journal of Hydrology, 270:309-327.