

## Building an Agro-Hydrologic Model of Europe: Model Calibration Issues

**K.C. Abbaspour<sup>a</sup>, E. Rouholahnejad<sup>a</sup>, S. Vaghefi<sup>a</sup>, R. Srinivasan<sup>b</sup>, H. Yang<sup>a</sup> and B. Kløve<sup>c</sup>**

<sup>a</sup>*Eawag, Swiss Federal Institute of Aquatic Science and Technology, P.O. Box 611, 8600 Dübendorf, Switzerland*

<sup>b</sup>*Texas A&M University, Texas Agricultural Experimental Station, Spatial Science Lab, College Station, TX 77845, USA*

<sup>c</sup>*Bioforsk-Norwegian Institute for Agricultural and Environmental Research*

[Abbaspour@eawag.ch](mailto:Abbaspour@eawag.ch)

**Abstract:** A combination of driving forces are increasing pressure on local, national, and regional water supplies needed for irrigation, energy production, industrial uses, domestic purposes, and the environment. In many parts of Europe groundwater quantity, and in particular quality, have come under severe degradation and water levels have decreased resulting in negative environmental impacts. Rapid improvements in the economy of the eastern European block of countries and uncertainties with regard to freshwater availability create challenges for water managers. At the same time, climate change adds a new level of uncertainty with regard to freshwater supplies. To address the above issues, a calibrated and reliable model of the region is of paramount importance. The Soil and Water Assessment Tool (SWAT) was used to build an agro-hydrological model of Europe. We simulated water resources as well as wheat, barley, and maize yields, and nitrate concentration of the surface and percolating water at the Hydrological Response Unit (HRU) level. The use of large-scale, high-resolution water resources models enables consistent and comprehensive examination of integrated system behavior through physically-based, data-driven simulation. However, calibration of large-scale models suffers from a number of conceptual and technical issues, which we believe require a more careful consideration by the scientific community.

The issues with large-scale model calibration include: 1) parameterization, 2) non-uniqueness, 3) objective function definition, 4) use of different optimization algorithms, 5) model conditionality, 6) inadequate definition of the base model, and 7) time constraints. A software for calibration of SWAT, SWAT-CUP was developed to enable the analysts to deal or to analyze some of these issues. In SWAT-CUP, i) detailed parameterization scheme is included to enable regionalization of parameters at the smallest SWAT unit (HRU); ii) Non-uniqueness is captured by including programs that address uncertainty issues; iii) 11 different objective functions are included to highlight the effect of using different objective functions; iv) 5 different optimization methods are included to highlight the effect of using different algorithms; and v) a Windows-based parallel processing scheme is included to speed up the calibration procedure.

In light of the above issues, “conditionality” of the calibrated model becomes an important issue, which can partially be addressed by using different variables in the objective function. Finally, we emphasize the importance of performing pre-calibration processing of the hydrologic model and the use of an initially adequate model to be used for calibration.

The SWAT agro-hydrological model developed for Europe was calibrated based on a large number of river discharge outlets and measured nitrate, mainly in the Danube Basin. Results of major river discharges and river nitrate concentration were quite well simulated all over Europe and in the Danube Basin, respectively. The results produced here could provide information support to the European Water Framework Directive and lay the basis for further assessment of the impact of climate change on water availability and quality. Many applications of this model could be foreseen such as: conducting policy and impact studies, calculating cross-boundary water transfers, calculating quantities of nitrogen loads being transferred from upstream to downstream of a river, and calculating nitrogen loads entering the seas and ocean. The approach and methods developed are general and can be applied to any large region around the world.

**Keywords:** *Water resources, nitrate load, SWAT, SWAT-CUP*

## 1. INTRODUCTION

Hydrological models are important tools for planning sustainable use of water resources to meet various demands. Estimates of country or global water resource are often obtained based on: i) data generalization of the world hydrological network (Shiklomanov, 2000), ii) general circulation models (GCMs) (TRIP, Oki et al. 2001; HO8, Hanasaki et al., 2013), and iii) hydrological models (WGHM, Alcamo et al., 2003; LPJ, Gerten et al., 2004; PCR GLOBWB). More accurate estimations, in terms of the hydrological processes, are based on the global hydrological models mentioned above, which are all raster models with a spatial resolution of 0.5° (55.7 km at the equator) and driven by monthly climatic variables. Probably the most sophisticated of these models is WGHM (Alcamo et al., 2003) that combines a hydrological model with a water use model and calculates surface runoff and groundwater recharge based on a daily water balance of soil and canopy. The global model is calibrated against observed discharge at 724 gauging stations spread globally by adjusting the runoff coefficient and, in case this was not sufficient, by applying up to two correction factors, especially in snow-dominated and semiarid or arid regions. Global runoff estimates performed with existing global climate models, e.g., Nijssen et al. (2001) and Oki et al. (2001), among others, suffer from low accuracy due to their low spatial resolution, poor representation of soil water processes, and, in most cases, lack of calibration against measured discharge (Döll et al., 2003). Similarly, most hydrologic models also have shortcomings with respect to weak hydrology, calibration and validated against long-term annual discharge, application of correction factors to the modeled discharges leading to an inconsistent water balance, and lack of quantification of model prediction uncertainty, which could be quite large in distributed models.

In this work, we use the program Soil and Water Assessment Tool (SWAT) to build an agro-hydrological model of Europe. SWAT is a comprehensive, semi-distributed, continuous-time, processed-based program used to build models to evaluate the effects of alternative management decisions on water resources quantity and quality in large river basins. The hydrological component of SWAT allows explicit calculation of different water balance components, and subsequently water resources at the hydrologic response unit (HRU), which is a unique combination of soil, landuse, and slope. The program is widely used to simulate the impact of landuse and climate changes on water resources and crop yield, among others.

The current modeling philosophy requires that models are transparently described; and that calibration, validation, sensitivity, and uncertainty analysis are routinely performed as part of modeling work. However, calibration of large-scale distributed models are besets with many problems including: 1) parameterization, 2) non-uniqueness, 3) objective function definition, 4) use of different optimization algorithms, 5) model conditionality, 6) inadequate definition of the base model, and 7) time constraints. For these reasons many researchers avoid model calibration and uncertainty analysis, a decision that results in unreliable model output. We recognize the shortcomings in calibration methods and uncertainty analysis and believe a closer look at the problems and issues in calibration/uncertainty analyses is called for by the scientific community.

SWAT-CUP (SWAT Calibration and Uncertainty Procedures) (Abbaspour, et al., 2007) is a standalone program developed for calibration of SWAT. The program contains five different calibration procedures: Parasol (Van Griensven and Meixner, 2006), Glue (Beven, 1992), and MCMC (Vrugt et al., 2003), PSO (Eberhart, and Kennedy, 1995), and SUFI-2 (Abbaspour, et al., 2007). It includes functionalities for validation and sensitivity analysis as well as visualization of the area of study using the Bing Map. With Bing Map features such as subbasins, simulated rivers, and outlet, rainfall, and temperature stations can be visualized. In the current work we used the program SUFI-2 for model calibration and uncertainty analysis.

In the following section we discuss the above calibration issues and describe the SWAT-CUP tools that enable these issues to be investigated.

## 2. CALIBRATION ISSUES

1) *Parameterization*. Parameterization is perhaps the most difficult and neglected aspect of calibration. Also referred to as regionalization of parameters, parameterization refers to the degree of parameter lumping in watershed-scale models. As an example, often a unit of soil map covers different landuses and many soil hydrologic parameters such as hydraulic conductivity, bulk density, etc. are greatly influenced by landuse/landcover. Treating soil hydrologic parameters as separate for different landcovers may results in a large number of parameters, while lumping them may cause a loss of heterogeneity. Choice of parameterization will affect the final parameter set and consequently hydrologic characterization of the basin under study.

In SWAT-CUP, parameterization is allowed to take place at the HRU level. The syntax for specifying a SWAT parameter is as follows:

```
x_<parname>.<ext>_<hydrogrp>_<soltext>_<landuse>_<subbsn>_<slope>
```

where  $x$  is one of:  $v$  (value change),  $a$  (value added), and  $r$  (relative change),  $parname$  is a SWAT parameter name,  $ext$  is SWAT file extension,  $hydrogrp$  is the hydrologic group of soil,  $soltext$  is soil texture,  $landuse$  is landuse type,  $subbsn$  is the subbasin number, and  $slope$  is the HRU slope. In this formulation the user can change a parameter based on any or a combination of characteristics, hence having the ability to apply different degrees of parameter lumping and studying the effect of different parameterization on model calibration.

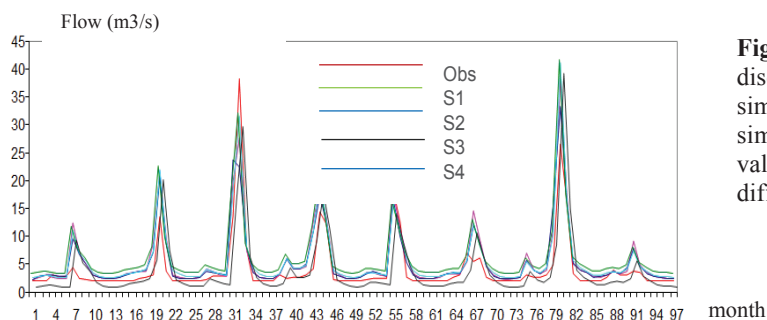
2) *Non-uniqueness*. Another issue with calibration of large-scale models is the non-uniqueness in the calibrated parameters. If there is one set of good parameters to be found, then there are many sets that would provide statistically similar results. In other words, if there is a single model that fits the measurements there will be many of them (Abbaspour et al., 2007). Our goal in calibration is, therefore, to characterize the set of models, mainly through assigning distributions (uncertainties) to the parameters that fit the data and satisfy our presumptions as well as other prior information. To make the parameter inferences quantitative, one must consider: the error in the input data; the error in the measured variables used for calibration; and the error in the conceptual mode. For this reason calibration and uncertainty analysis are intimately linked and must be done simultaneously.

3) *Objective function definition*. This issue is the most disappointing aspect of calibration because after calibration with a certain objective function, one realizes that with a different choice of objective function a different set of parameters would have been achieved with the same calibration results. To analyze the differences between different objective functions, SWAT-CUP provides 11 different functions. The differences between different objective functions on parameter sets could be quite striking. Table 1 shows four sets of SWAT parameters obtained by four different objective functions. The parameters sets are significantly different from each other describing different watersheds from a hydrological point of view, while calibrated flows are similar and each parameter set is favored by a different objective function (Table 1, Figure 2).

**Table 1.** Final calibrated model parameters with different objective functions followed by a Table of objective function values and graphs of calibrated rivers discharges.

	CN2	ALPHA BF	REVAPMN	GW REVAP	RCHRG DP	SOL AWC	SOL K	SOL BD
S1	-0.206	0.25	83.46	0.059	0.095	-0.195	0.572	-0.135
S2	0.110	0.59	66.06	0.056	0.191	-0.256	0.456	-0.191
S3	0.338	0.35	86.14	0.071	0.194	-0.047	0.300	0.086
S4	0.297	0.50	32.87	0.073	0.290	-0.165	0.297	-0.075

Parameter set	RMSE	R2	NS	bR2
S1	24.90	0.79	0.67	0.75
S2	29.70	0.85	0.67	0.42
S3	23.50	0.84	0.83	0.58
S4	21.10	0.83	0.80	0.62



**Figure 2.** Observed discharge and four sets of simulated discharges with similar objective function values but significantly different parameters

4) *Optimization algorithms*. Yang et al. (2008) compared Parasol, Glue, MCMC, and SUFI-2 in calibrating Chaohe Basin in China. Each optimization algorithm has a different way of looking for optimum parameter and quantifying the prediction uncertainty. Similar to objective function differences, they also found all algorithms produced more or less similar calibration/validation results. However, the final parameter uncertainties and the best parameters (parameter set that maximizes or minimizes the objective function) were quite different.

5) *Model conditionality*. A further issue with calibrated models is their “conditionality”. Meaning, a calibrated model is conditioned on a number of assumptions and will usually not apply to situations outside of those boundaries. Specifically, if a model is calibrated on river discharge, it should not be expected to correctly predict water quality, evapotranspiration, soil moisture, etc. Or, if a model is calibrated at the outlet of a watershed, it should not be expected to provide correct values at subbasins inside the watershed. To partially

overcome the conditionality problem, more variables (sediment loads, water quality variables, crop yield, etc.) should be included in the objective function at several points in a watershed.

6) *Inadequate initial model definition.* Another major problem with calibration of large-scale models is the use of an initially inadequate model. Common problems are wrong river pathways, wrong location of measured river discharge stations, inadequate accounting of hydrological processes such as point sources, snowmelt, wetlands, dam, reservoirs, etc. For this reason we recommend a thorough pre-calibration analysis of the model (Faramarzi et al., 2015).

7) *Time constraints.* Finally, time is of critical element in model calibration and uncertainty analyses, which call for a large number of model runs. Each model run may take a long time, therefore, often calibration and uncertainty analysis are inadequately performed. A Window's-based parallel processing routine is implemented in SWAT-CUP for SUFI-2 program, which substantially increases calibration speed (Rouholahnejad, 2013).

### 3. BUILDING AND CALIBRATING A EUROPEAN-SCALE SWAT MODEL

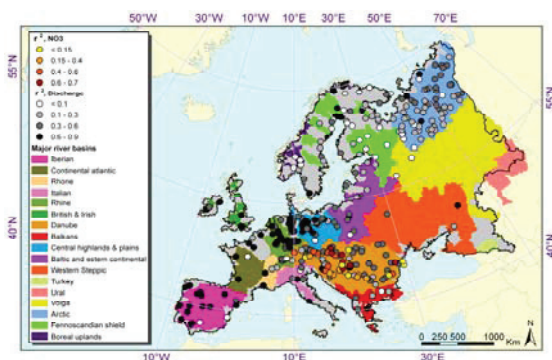
The goal of this work was to use SWAT to build a hydrological model of Europe at subbasin level and monthly time intervals. The key objectives of this work were: i) to incorporate agricultural management and crop yield to the hydrological model for a more accurate calculation of evapotranspiration, ii) to add water quality to the model by adding point sources and diffuse sources of nitrogen to investigate the nitrate leaching into the groundwater, iii) to quantify spatial and temporal variations in water availability at European Continent scale. In particular, components such as blue water (water yield plus deep aquifer recharge), green water flow (evapotranspiration), green water storage (soil moisture), and nitrate concentration of groundwater recharge are quantified. The model estimations at the subbasin level are then aggregated to country and river basin levels for comparison with other studies. Detailed description of model building and input data are reported in Abbaspour et al., (2015) and are not repeated here.

Data used to build the European model was mostly obtained from the Internet and included: 90 m DEM from (SRTM) aggregated to 700 m (<http://www2.jpl.nasa.gov/srtm/>); 5-km FAO-UNESCO global soil map (<http://www.fao.org/nr/land/soils/digital-soil-map-of-the-world/en/>); 500-m MODIS land cover (<http://modis-land.gsfc.nasa.gov/>); 62-km (average catchment size) European catchments and Rivers network System (Ecrins) (<http://projects.eionet.europa.eu/ecrins/>); 0.5° Climate Research Unit (CRU) gridded climate data (<http://www.cru.uea.ac.uk/>); 326 discharge station from Global Runoff Data Center (GRDC) ([http://www.bafg.de/GRDC/EN/Home/homepage\\_node.html](http://www.bafg.de/GRDC/EN/Home/homepage_node.html)); 34 measured nitrate data from ICPDR ([http://en.wikipedia.org/wiki/International\\_Commission\\_for\\_the\\_Protection\\_of\\_the\\_Danube\\_River](http://en.wikipedia.org/wiki/International_Commission_for_the_Protection_of_the_Danube_River)); planting, harvest, and fertilization data from FAOSTAT (<http://faostat.fao.org/site/339/default.aspx>); blue water flow from AQUASTAT, FAO ([http://www.fao.org/nr/water/aquastat/water\\_res/index.stm](http://www.fao.org/nr/water/aquastat/water_res/index.stm)); population data from Eurostat (<http://epp.eurostat.ec.europa.eu/portal/page/portal/population/introduction>); population growth rate from World Bank (<http://data.worldbank.org/indicator/SP.POP.GROW>); and point source pollution Eurostat for the period of 2000 to 2009.

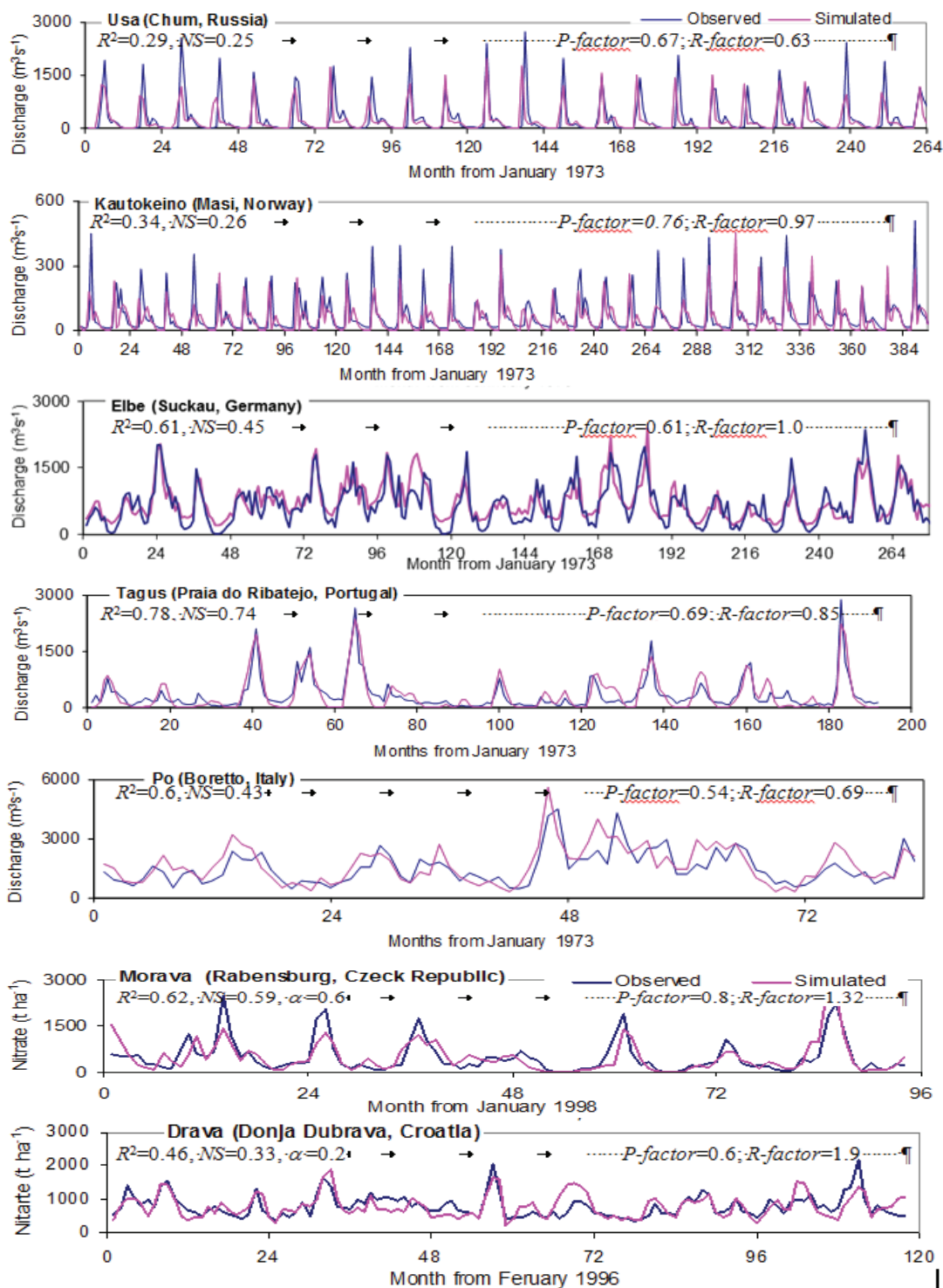
## 4. RESULTS

### 3.1. Calibration/Validation of discharge

The overall performance of the model in terms of  $R^2$  has quite satisfactory results (Fig. 3). It should be kept in mind that all 326 discharge outlets and 34 nitrate outlets were parameterized and optimized simultaneously. This results in an overall good simulation and not the best solution at every individual outlet. Calibrated results of a few important river basins of Europe are illustrated in Fig. 4. For more details we refer the reader to Abbaspour et al., (2015). The characteristics of major rivers and the model performances are assessed for both discharge and nitrate.



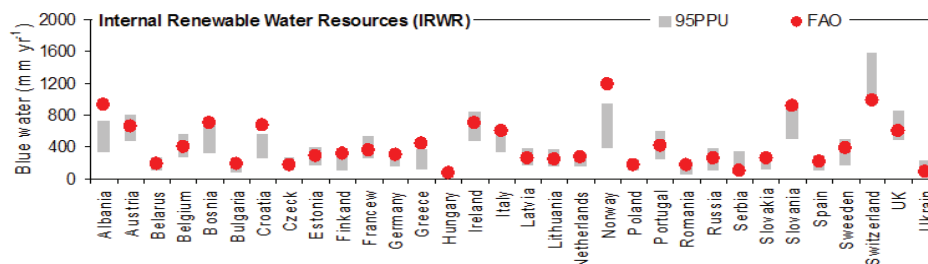
**Figure 3.** (left) Overall modeling results at observed discharge and nitrate stations.  $R^2$  is used as an index showing the goodness of fit.



**Figure 4.** Illustration of model discharge and nitrate results for some major rivers across Europe. The Figures covers both calibration and validation periods. The statistics  $R^2$ ,  $NS$ , and  $\alpha$  are for the entire periods,  $P$ -factor (% observed data bracketed by the 95% prediction uncertainty, 95ppu) and  $R$ -factor (difference of upper and lower 95ppu divided by the standard deviation of the measured data) are for validation period only.

#### 4. DISCUSSION

Calibration and uncertainty analysis are intimately linked as model calibration is conditional on the type and length of data used for calibration, the objective function definition, the hydrologic model, the optimization routine, and all other model assumptions. Therefore comparison of outputs from different hydrologic models is often not possible because of the conditionality principle. However some comparison is necessary to show the relevancy of the results obtained. Here, we compared the modeled blue water results with values reported by FAO (AQUASTAT, 2013) (Fig. 5). Data from FAO were selected not to serve as “true” values but because they are very commonly used by governmental and scientific organizations. The comparison, shows that most of the FAO values fall within our model prediction uncertainty. No such comparison was made for soil moisture and actual evapotranspiration as we did not find comparable data in the literature.



**Figure 5.** Comparison of the country-based blue water flow results with those reported by FAO (AQUASTAT, 2013).

#### 5. SUMMARY AND CONCLUSION

In this study a subbasin-scale hydrologic model of Europe was built using the well-established SWAT program. The model was calibrated for a large number of river discharge stations, nitrate loads of rivers, and yields of wheat, maize, and barley. The program SUFI-2 in SWAT-CUP package was used for calibration/uncertainty analysis, validation, and sensitivity analysis. Only readily available data were used for model setup as well as calibration and validation. The final model results for the freshwater availability components, blue water flow, green water flow, and green water storage are presented. Particular attention was paid to quantify and display the 95% prediction uncertainty of the outputs, which turned out to be quite large in some cases. A protocol for calibration of large-scale models is also presented (see Abbaspour et al., 2015). Many of the difficulties and limitations within this continental modeling study were data related and resulted from, among others, (1) limited nitrate data and discharge stations with varying time series lengths, (2) limited knowledge of the attributes and management of reservoirs, (3) lack of data on soil moisture and/or deep aquifer recharge making a proper calibration/validation of these components impossible, (4) lack of knowledge of spatially detailed agricultural management operations, and (5) conceptual model assumptions and simplifications.

Overall, this study provides significant insights into continental freshwater availability and water quality at a subbasin level with a monthly time step. This information is very useful for developing an overview of the actual water resources status and helps to spot regions where an in-depth analysis may be necessary. Many applications of this model could be foreseen such as: conducting policy and impact studies, using the model for climate and landuse change studies, calculating cross-boundary water transfers, calculating quantities of nitrogen loads being transferred from upstream to downstream of a river, and calculating nitrogen loads entering the seas and ocean.

Finally, it is shown that given the available technology on model building and calibration tools, and the availability of freely available data it is possible to build a continental model at high spatial and temporal resolution. Better data availability of course, would help to make model predictions more accurate and uncertainties smaller.

#### REFERENCES

- Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R., 2007. Modelling hydrology and water quality in the pre-Alpine/Alpine Thur watershed using SWAT. *J. Hydrol.* 333, 413-430.
- Abbaspour, K. C. E. Rouholahnejad, S. Vaghefi, R. Srinivasan, B. Klöve. 2015. Modelling hydrology and water quality of the European Continent at a subbasin scale: calibration of a high-resolution large-scale SWAT model. *Journal of Hydrology*, 524: 733-752.

- <http://www.sciencedirect.com/science/article/pii/S0022169415001985>
- Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. *J. Sci. Hydrol.* 48 (3), 317–337.
- Beven, K. and Binley, A., 1992. The Future of Distributed Models - Model Calibration and Uncertainty Prediction. *Hydrological Processes*, 6(3): 279-298.
- Döll, P., Kaspar, F., and Lehner, B., 2003. A global hydrological model for deriving water availability indicators: Model tuning and validation, *J. Hydrol.* 270(1–2), 105– 134.
- Eberhart, R. C., and Kennedy, J. 1995. A new optimizer using particle swarm theory. *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan, 39-43. Piscataway, NJ: IEEE Service Center
- Faramarzi, M., Srinivasan, R., Irvani, M., Bladon, K.D., Abbaspour, K.C., Zehnder, A., Goss G., 2015. Setting up a hydrological model of Alberta: Data discrimination analyses prior to calibration. *Environmental Modelling and Software* 74:48-65...
- Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., Sitch, S., 2004. Terrestrial vegetation and water balance - Hydrological evaluation of a dynamic global vegetation model, *J. Hydrol.* 286, 249– 270.
- Hanasaki, N., Fujimori, S., Yamamoto, T., Yoshikawa, S., Masaki, Y., Hijioka, Y., Kainuma, M., Kanamori, Y., Masui, T., Takahashi, K., Kanae, S., 2013. A global water scarcity assessment under shared socio-economic pathways – Part 2: Water availability and scarcity. *Hydrol. Earth Syst. Sci.* 17, 2393–2413.
- Nijssen, B., O'Donnell, G.M., Lettenmaier, D.P., Lohmann, D., Wood, E.F., 2001. Predicting the discharge of global rivers, *J. Clim.*, 14(15), 3307– 3323.
- Oki, T., Agata, Y., Kanae, S., Saruhashi, T., Yang, D.W., Musiak, K., 2001. Global assessment of current water resources using total runoff integrating pathways. *Hydrol. Sci. J.: J. Sci. Hydrol.* 46 (6), 983–995.
- Rouholahnejad, E., Abbaspour, K.C., Vejdani M., Srinivasan R., Schulin R., Lehmann A., 2012. Parallelization framework for calibration of hydrological models, *Environmental Modelling Software*, 31, 28-36.
- Shiklomanov, I.A., 2000. Appraisal and assessment of world water resources. *Water Int.* 25 (1), 11–32.
- Van Griensven, A., Meixner, T., 2006. Methods to quantify and identify the sources of uncertainty for river basin water quality models. *Water Science and Technology*, 53(1): 51-59.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian. 2003. A shuffled Complex Evolution Metropolis Algorithm for Estimating Posterior Distribution of Watershed Model Parameters, in *Calibration of Watershed Models*, ed. Q. Duan, S. Sorooshian, H. V. Gupta, A. N. Rousseau, and R. Turcotte, AGU Washington DC, DOI: 10.1029/006WS07.
- Yang, J., Reichert P., Abbaspour K.C., Yang H., 2008. Comparing uncertainty analysis techniques for a SWAT application to Chaohe Basin in China. *Journal of Hydrology*. 358, 1-2