Journal of Hydrology 524 (2015) 733–752

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model

K.C. Abbaspour [a,*], E. Rouholahnejad [a], S. Vaghefi [a], R. Srinivasan [b], H. Yang [a,c], B. Kløve [d]

[a] Eawag, Swiss Federal Institute of Aquatic Science and Technology, P.O. Box 611, 8600 Dübendorf, Switzerland
[b] Texas A&M University, Texas Agricultural Experimental Station, Spatial Science Lab, College Station, TX 77845, USA
[c] Department of Environmental Science, University of Basel, Switzerland. Petersplatz, 1, 6003 Basel, Switzerland
[d] Univ Oulu, Fac Technol, POB 4300, FIN-90014 Oulu, Finland

ARTICLE INFO

Article history:
Received 18 December 2013
Received in revised form 10 February 2015
Accepted 7 March 2015
Available online 18 March 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Michael Brian Butts, Associate Editor

Keywords:
Water resources
Runoff ratio
Nitrate load
SWAT-CUP
SUFI-2
Blue water

SUMMARY

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 sever 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. In this research we build and calibrate an integrated hydrological model of Europe using the Soil and Water Assessment Tool (SWAT) program. Different components of water resources are simulated and crop yield and water quality are considered at the Hydrological Response Unit (HRU) level. The water resources are quantified at subbasin level with monthly time intervals. Leaching of nitrate into groundwater is also simulated at a finer spatial level (HRU). 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. In this article we discuss issues with data availability, calibration of large-scale distributed models, and outline procedures for model calibration and uncertainty analysis. The calibrated model and results 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. The approach and methods developed are general and can be applied to any large region around the world.

1. Introduction

Higher standards of living, demographic changes, land and water use policies, and other external 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 quality in particular has come under server pressures and water levels have decreased, resulting in negative environmental impacts (Kløve et al., 2014). Rapid, and often, unpredictable changes with regard to freshwater supplies create uncertainties for water managers. At the same time, climate change adds a new level of uncertainty with regard to freshwater supplies and to the main water use sectors such as agriculture and energy, which will in turn exacerbate uncertainties regarding future demands for water. As meeting future water demands becomes more uncertain, and water scarcity is continuously increasing (Yang et al., 2003), societies become more vulnerable to a wide range of risks associated with inadequate water supply in quantity and/or quality (UN Report, 2012).

Hydrological models are important tools for planning sustainable use of water resources to meet various demands.

Abbreviations: REVAMPM.gw, treshold depth of water in the shallow aquifer required for capillary flow into root zone to occur (mm); GW_REVAP.gw, groundwater "revap" coefficient; GWQMN.gw, treshold depth of water in the shallow aquifer required for return flow to occur (mm); SHALLST_N.gw, concentration of nitrate in groundwater contribution to streamflow from subbasin (mg N l^{-1}); CN2.mgt, SCS runoff curve number for moisture condition II; FRT_SURFACE.mgt, fraction of fertilizer applied to top 10 mm of soil; SOL_AWC.sol, Available water capacity of the soil layer (mm mm^{-1}); ESCO.hru, soil evaporation compensation factor; HRU_SLP.hru, average slope steepness (m m^{-1}); OV_N.hru, Manning's "n" value for overland flow; SLSUBBSN.hru, average slope length (m); RCN.bsn, concentration of nitrogen in rainfall (mg N l^{-1}); NPERO.bsn, nitrogen percolation coefficient; CMN.bsn, rate factor for humus mineralization of active organic nitrogen; SOL_NO_3.chm, initial NO_3 concentration in the soil layer (mg kg^{-1}).

* Corresponding author at: 133 Ueberlandstr, 8600 Dübendorf, Switzerland.
E-mail address: abbaspour@eawag.ch (K.C. Abbaspour).

http://dx.doi.org/10.1016/j.jhydrol.2015.03.027
0022-1694/

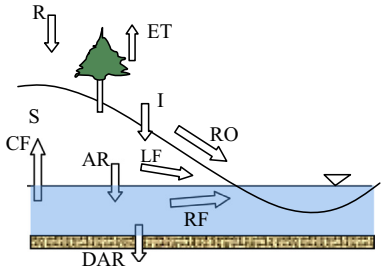


Fig. 1. Schematic illustration of the conceptual water balance model in SWAT.

Some works on the estimation of global water resources are published as early as 1970s by Lvovitch (1973), Korzun et al. (1978), and Baumgarten and Reichel (1975). The country or global-based water resource estimates are performed 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 (Yates, 1997; WMB, Vörösmarty et al., 2000; Fekete et al., 1999; Macro-PDM, Arnell, 1999; WGHM, Alcamo et al., 2003; Yang and Musiake, 2003; LPJ, Gerten et al., 2004; WASMOD-M, Widén-Nilsson et al., 2007; PCR GLOBWB, van Beek et al., 2011). 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). 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. The main shortcomings of the models mentioned above are the 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.

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. As calibration is "conditional" (i.e., conditioned on the model structure, model inputs, analyst's assumptions, calibration algorithm, calibration data, etc.) and not uniquely determined, uncertainty analysis is essential to evaluate the strength of a calibrated model.

The Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998) has demonstrated its strengths in the aspects specified above. It is an open source code with a large and growing number of model applications in various studies ranging from catchment to continental scales. In the "Hydrologic Unit Model for the United States" (HUMUS), Arnold et al. (1999) used SWAT to simulate the entire U.S.A. for river discharges at around 6000 gauging stations. This study was then extended within the national assessment of the USDA Conservation Effects Assessment Project (CEAP). Gosain et al. (2006) modeled twelve large river catchments in India with the purpose of quantifying the climate change impact on hydrology. Schuol et al. (2008) simulated hydrology of the entire Africa with SWAT in a single project and calculated water resources at a subbasin spatial resolution and monthly time intervals. Faramarzi et al. (2009) simulated hydrology and crop yield for Iran with SWAT. In a subsequent work, Faramarzi et al. (2013) used the African model to study the impact of climate change in Africa.

Table 1
Data description and sources used in the European SWAT project.

Data type	Resolution	Source
Digital Elevation (DEM)	90 m aggregated to 700 m	Shuttle Radar Topography Mission (SRTM) http://www2.jpl.nasa.gov/srtm/
Soil	5 km	FAO–UNESCO global soil map http://www.fao.org/nr/land/soils/digital-soil-map-of-the-world/en/
Landuse	– 300 m	– GlobCover European Space Agency http://due.esrin.esa.int/globcover/
	– 1000 m	– Global Landuse/Land Cover Characterization USGS http://landcover.usgs.gov/glcc/
	– 500 m	– MODIS land cover http://modis-land.gsfc.nasa.gov/
	– 300 m	– GlobCorine provided by European Space Agency http://www.esa.int/About_Us/ESRIN/Express_map_delivery_from_space
River network dataset	≅62 km^2 avg. size catchment	European catchments and Rivers network System (Ecrins) http://projects.eionet.europa.eu/ecrins
Climate	– Observed	– National Climate Data Center (NCDC), http://www.ncdc.noaa.gov/
	– 0.25° grid	– European Climate Assessment Dataset (ECAD), http://www.ecad.eu/
	– 0.5° grid	– Climate Research Unit (CRU), http://www.cru.uea.ac.uk/
	– 1° grid	– Climate Data Guide (NCAR), https://climatedataguide.ucar.edu/
River discharge	326 stations	Global Runoff Data Center (GRDC) http://www.bafg.de/GRDC/EN/Home/homepage_node.html
Nitrate loads	34 stations	ICPDR http://en.wikipedia.org/wiki/International_Commission_for_the_Protection_of_the_Danube_River
Crop yield	Wheat, maize, barley	McGill Univesity http://www.geog.mcgill.ca/landuse/pub/Data/175crops2000/NetCDF/ FAOSTA – Country-based average crop yield
Agricultural management and water resources	Planting, harvesting, fertilization-blue water	FAOSTAT http://faostat.fao.org/site/339/default.aspx – AQUASTAT, FAO http://www.fao.org/nr/water/aquastat/water_res/index.stm
Population		Eurostat http://epp.eurostat.ec.europa.eu/portal/page/portal/population/introduction
Population growth rate		World Bank http://data.worldbank.org/indicator/SP.POP.GROW
Point source pollution		Eurostat for the period of 2000–2009

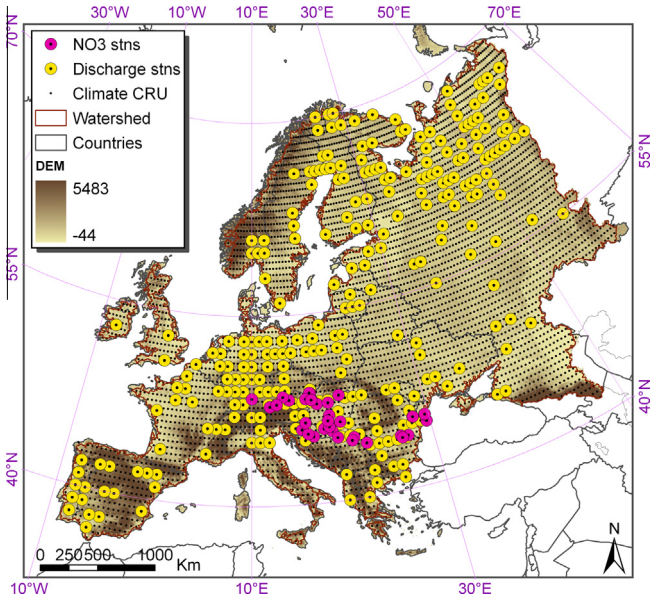


Fig. 2. Map of the modeled area showing the location of measured discharge and nitrate stations as well climate CRU grid points.

SWAT is recognized by the U.S. Environmental Protection Agency (EPA) and has been incorporated into the EPA's BASINS (Better Assessment Science Integrating Point and Non-point Sources). The larger capacity of ArcGIS (version 9.3 and higher) now allows building finer resolution large-scale models, which could be calibrated using powerful parallel processing (Rouholahnejad et al., 2012) and grid computations (Gorgan et al., 2012), allowing proper uncertainty analysis.

SWAT-CUP (Calibration and Uncertainty Procedures) (Abbaspour et al., 2007) is a standalone program developed for calibration of SWAT. The program contains five different calibration procedures and includes functionalities for validation and sensitivity analysis as well as visualization of the area of study using Bing Map. With this feature, the subbasins, simulated rivers, and outlet, rainfall, and temperature stations can be visualized on the Bing map. In the current work we used the program SUFI-2 (Abbaspour et al., 2004; Abbaspour et al., 2007) for model calibration and uncertainty analysis. For time-consuming large-scale models, SUFI-2 was found to be quite efficient (Yang et al., 2008).

Against this background, the goal of this work is to use SWAT to build a hydrological model of Europe at subbasin level and monthly time intervals. The key objectives of this work are: (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.

The results from this study provide a consistent information package on the quantity and quality of water resources on temporal and spatial dimensions, and on internal renewable water resources for individual countries across Europe. A calibrated model at this scale can be used for various analyses such as cross-boundary water analysis, assessment of nitrate loads generated in various countries; quantification of the nitrate input into various lakes, sees, and ocean; highlighting the sources of contaminant generation; and finally studying the impact of climate and landuse change on the continent's water resources.

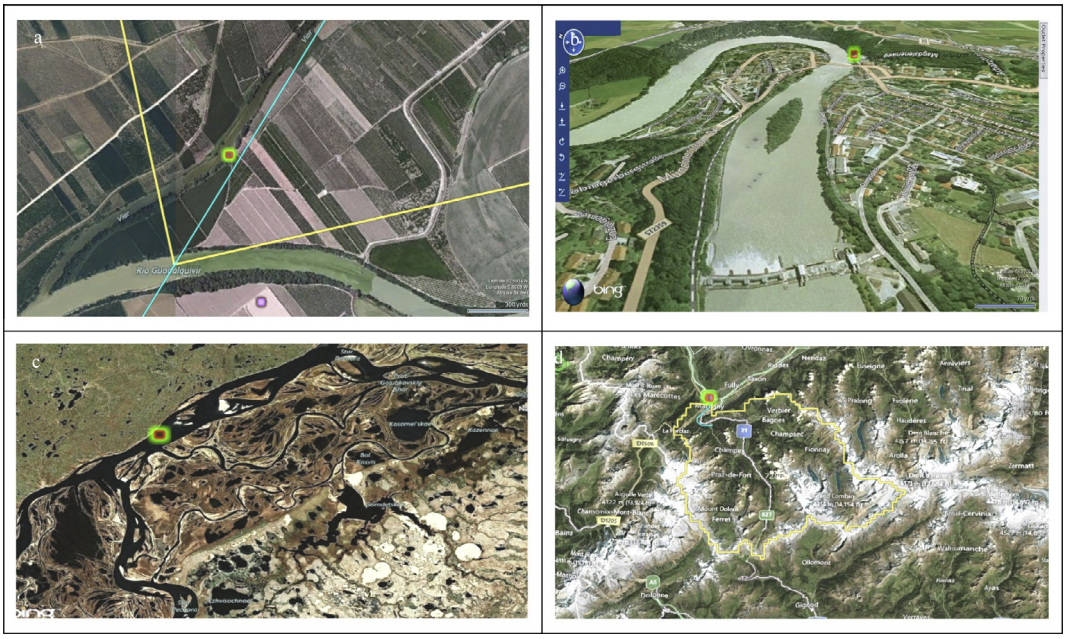


Fig. 3. Illustration of the visualization module in SWAT-CUP. This allows gathering information from the site that helps with model parameterization and parameter range identification. Figure (a) shows positioning of an outlet on Viar river instead of on the main Rio Guadalquivir in Spain; figure (b) shows position of an outlet downstream of a dam on Inn River near Munich, Germany; figure (c) shows a complex river geometry on Pechora River near Golubovo in Russia; figure (d) shows an outlet governed by glacier melt near Martigny in Switzerland. These features could explain some of the discrepancies between simulation and observed results if parameterized incorrectly. The red–green symbol indicates location of the outlet. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Material and methods

2.1. The large-scale hydrological simulator SWAT

The SWAT program is a comprehensive, semi-distributed, continuous-time, processed-based model (Arnold et al., 2012; Neitsch et al., 2005; Gassman et al., 2007). The program can be used to build models to evaluate the effects of alternative management decisions on water resources and non-point source pollution in large river basins. The hydrological component of SWAT (Fig. 1) allows explicit calculation of different water balance components, and subsequently water resources (e.g., blue and green waters) at a subbasin level.

In SWAT, a watershed is divided into multiple subbasins, which are then further subdivided into hydrologic response units (HRUs)

Table 2
Rules for parameter regionalization. ↑ indicates parameter should increase, ↓ indicates parameter should decrease. (For more detail see SWAT calibration validation literature http://swat.tamu.edu/publications/calibrationvalidation-publications/).

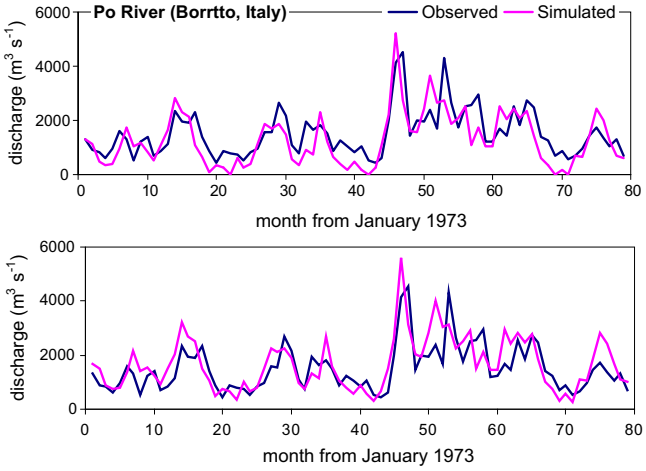
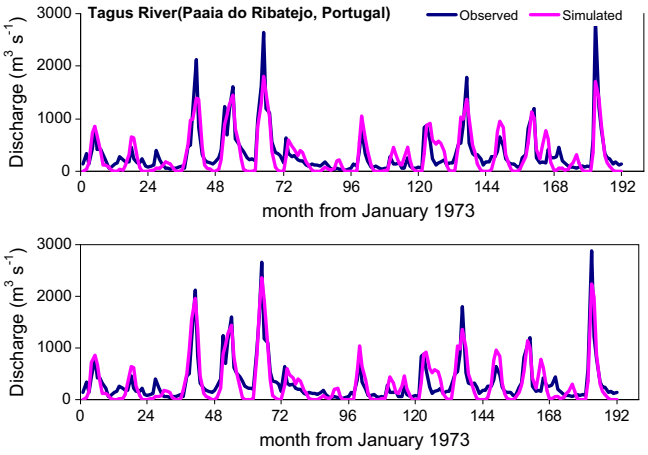
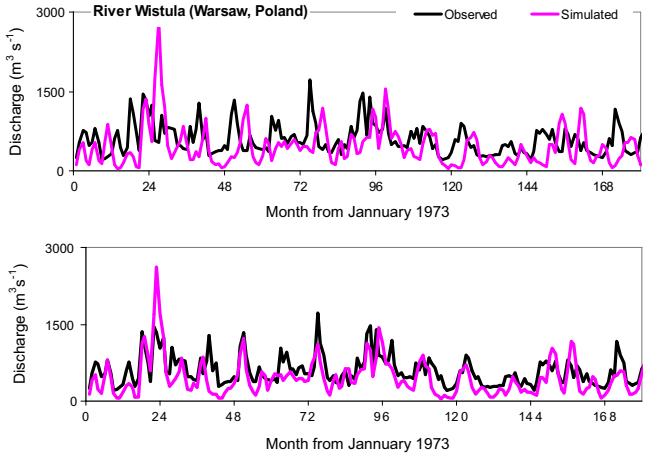
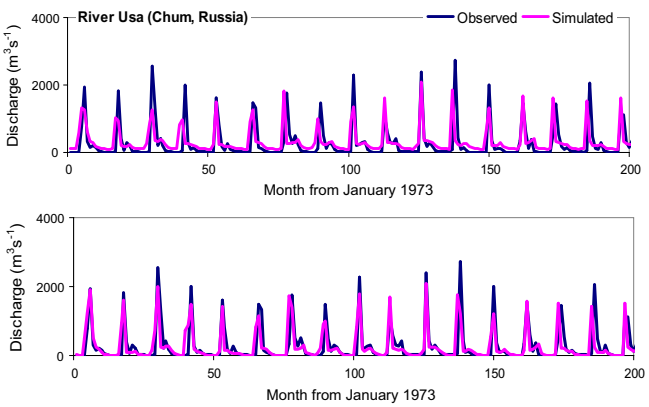
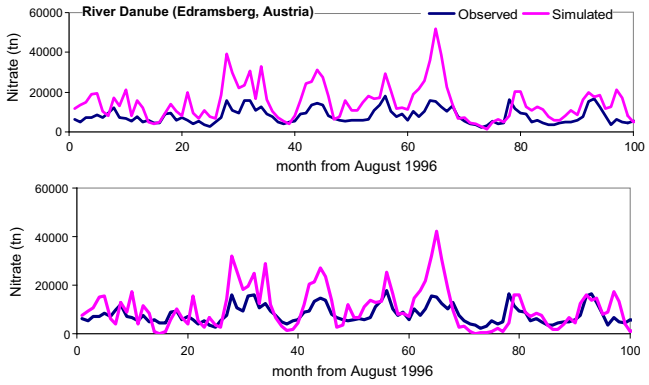
Simulated condition and significant parameters	Before (top) and after (below) parameter regionalization
– **Base flow too low** **–Evapotranspiration too high** GWQMN.gw ↓ GW_REVAP.gw ↓ REVAMPM.gw ↑	
Peak flow too low CN2.mgt ↑ SOL_AWC.sol ↓ ESCO.hru ↓	
Discharge shift to the right HRU_SLP ↑ OV_N.hru ↓ SLSUBBSN.hru ↓	

Table 2 (*continued*)

Simulated condition and significant parameters	Before (top) and after (below) parameter regionalization
– **Base flow too high** – **Peaks too low** CN2.mgt ↑ SOL_AWC.sol ↓ ESCO.hru ↓ GWQMN.gw ↑ GW_REVAP.gw ↑ REVAMPM.gw ↓	
Nitrate load too high SHALLST_N.gw ↓ RCN.bsn ↓ NPERO.bsn ↓ CMN.bsn ↓ SOL_NO3.chm ↓ FRT_SURFACE.mgt ↓ (in our case also decreasing parameter α in Eq. (1))	

that consist of unique landuse, management, topographical, and soil characteristics. Simulation of watershed hydrology is done in the land phase, which controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each subbasin, and in the routing phase, which is the movement of water, sediments, etc., through the streams of the subbasins to the outlets.

The hydrological cycle is climate driven and provides moisture and energy inputs, such as daily precipitation, maximum/minimum air temperature, solar radiation, wind speed, and relative humidity that control the water balance. Snow is computed when temperatures are below freezing, and soil temperature is computed because it impacts water movement and the decay rate of residue in the soil. Hydrologic processes simulated by SWAT include canopy storage, surface runoff, and infiltration. In the soil the processes include lateral flow from the soil, return flow from shallow aquifers, and tile drainage, which transfer water to the river; shallow aquifer recharge, and capillary rise from shallow aquifer into the root zone, and finally deep aquifer recharge, which removes water from the system. Other processes include moisture redistribution in the soil profile, and evapotranspiration. Optionally, pumping, pond storages, and reservoir operations could also be considered. The water balance for reservoirs includes inflow, outflow, rainfall on the surface, evaporation, seepage from the reservoir bottom, and diversions.

Addressing vegetation growth is essential in a hydrological model as evapotranspiration is an important component of water balance, and management operations such as irrigation (Faramarzi et al., 2009) and fertilization have a large impact on hydrology and water quality, respectively. SWAT uses a single plant growth model to simulate growth and yield of all types of land covers and differentiates between annual and perennial plants.

In addition, SWAT simulates the movement and transformation of several forms of nitrogen and phosphorus, pesticides, and sediment in the watershed. SWAT allows users to define management practices taking place in every HRU. Once the loadings of water, sediment, nutrients, and pesticides from the land phase to the main channel have been determined, the loads are routed through the streams and reservoirs within the watershed. More details on the SWAT can be found in the theoretical documentation (http://swatmodel.tamu.edu) and in Arnold et al. (1998).

2.2. Databases

The model for the continent of Europe was constructed using, for most parts, freely available data (Table 1). These data were complemented by additional sources provided by project partners on climate and agricultural management.

2.3. Model setup

The ArcSWAT 2009 interface is used to setup and parameterize the model. On the basis of DEM and the stream network, a threshold drainage area of 5000 km^2 was chosen to discretize the continent into 8592 subbasins, which were further subdivided into 60,012 HRUs based on soil, landuse, and slope. Each HRU is thought to be a uniform unit where water balance calculations are made. The entire simulation period is from 1970 to 2006. As each station has data for different years, we used about two-third of the data for calibration and the remaining for validation. The first 3 years are

Table 3
Simulated mean annual river discharges ($m^3 s^{-1}$) for a selection of European rivers based on four different climate datasets.

River	Station name	Discharge	Climate database			
		GRDC[a]	ECAD[b]	CRU[c]	NCAR[d]	Station[e]
Volga	Volgograd	8141	5570	6924	5951	7465
Danube	Ceatal Izmail	6415	3243	5244	4724	4204
Pechora	Oksino	4444	1849	2396	2330	2277
North. Dvina	Ust-Pinega	3331	1419	1597	1470	1705
Rhine	Lobith	2229	1725	2065	2110	1589
Rhone	Beaucairw	1709	1735	1808	1707	1364
Sava	Sremska M.	1563	411	1078	881	973
Po	Pontelagoscuro	1514	1108	1589	1424	1210
Dnieper	Dnieper P.P	1492	1987	2182	1811	2421
Vistula	Tczew	1042	462	715	620	846
Loire	Montjean	838	761	948	909	831
Tisza	Senta	784	445	469	448	435
Inn	Passau-Ingling	737	493	723	510	367
Elbe	Neu-Darchau	708	327	516	495	500
Garonne	Mas-d'Agenais	606	343	451	422	605
Aare	Untersiggenthal	561	405	515	528	299
Duero	Regua	545	454	573	645	626
Oder	Gozdowice	530	181	391	345	365
Angermaelven	Sollefteae	503	144	190	215	239
Luleaelven	Boden	501	291	297	328	229
Ebro	Tortosa	485	678	469	509	250
Tiber	Roma	231	128	114	115	182
Guadalquivir	Sevilla	207	219	226	244	204
Siret	Lungoci	198	36	119	115	136
Maros	Mako	173	9	95	86	85
Olt	Stoenesti	161	21	62	67	47
Szamos	Satu Mare	126	41	63	81	54
Trent	Colwick	85	67	65	59	61
Thames	Teddington	82	37	47	72	36
Root mean square error			5500	3360	4100	3900

[a] The GRDC values are the observed annual average river discharges ($m^3 s^{-1}$).
[b] ECAD, European Climate Assessment Dataset.
[c] CRU, Climate Research Unit.
[d] NCAR, Climate Data Guide.
[e] Station, Observed data from National Climate Data Center (NCDC).

used as equilibration period to mitigate the initial conditions and were excluded from the analysis. This large project was built on a laptop with a 64 bit operating system, 4 CPUs, 8 GB of RAM and 2.7 GHz processors. It took about 12 h for a single SWAT run of 37 years. Initial runs were performed with the parallel processing routine linked to SUFI-2 (Rouholahnejad et al., 2012). Final runs were performed using the server-based network cluster gSWAT (Gorgan, et al., 2012; Mihon, et al., 2013).

In general, the state of freely available river discharge measurements in terms of water quantity and quality, and crop yield data is relatively poor in Europe. A total of 326 discharge stations and only 34 stations with nitrate data were found reliable enough to be used in model calibration/validation process (Fig. 2). Majority of the measured nitrate stations were on the Danube River. Furthermore, much correction needed to be made to the river data as many broken links caused wrong flow directions, and many corrections were also made to the coordinates of the measured discharge stations. Lack of precision in the coordinates causes ArcSWAT to place the observed outlets on the wrong river, causing a major calibration problem. The visualization option of SWAT-CUP, using the Bing map, is of paramount help in detecting these problems and other abnormalities (Fig. 3).

Five elevation bands were used in the model to adjust the temperature and rainfall based on subbasin elevation variation. As detailed operational information on lakes and reservoirs were lacking, we did not use any outlets that were influenced directly by reservoirs.

Point sources of pollution were assigned to each subbasin based on the population percentage connected to sewage treatment plant (Eurostat 2000–2009). This percentage was above 80% in approximately half of the European Union countries for which data were available. As the data was not available for all countries and all years from 1970 to 2006, we made a few assumptions. For example, Moldova, Ukraine, Russia, Belarus, and Georgia were assumed to have the same rates as Bulgaria. Serbia, Bosnia, Montenegro was assumed to be in the same situation as Slovenia. Also, point source discharges 1970–2000 were assumed to be constant and equal to year 2000. In terms of treatment levels, tertiary wastewater treatment was most common in Germany, Austria and Italy where at least four in every five persons were connected to this type of wastewater treatment. In contrast, not more than 1% of the population was connected to tertiary wastewater treatment in Romania and Bulgaria. We assumed the treatment efficiency to be 80% in all cases. Furthermore, the effluent from 20% of the pollution was assumed to be untreated wastewater and directly loaded into surface waters.

The regional population was calculated based on the population map in the year 2005 (CIESIN, 2005) and extrapolated to other years based on the national population growth rate provided by World Bank. Zessner and Lindtner (2005) investigated the emissions of nutrient from municipal point sources based on detailed evaluation of data from 76 municipal wastewater treatment plants and calculation of discharge of nitrogen (N) from households into wastewater in Austria. The results of this investigation show that the average value of nitrogen load is 8.8 gN/P_e day for municipal wastewater (contributions from households and industry). Population equivalent P_e accounts for industrial releases. Nitrate loading to rivers was calculated as:

$$L_N = \alpha_i P_e I_N [(1 - S_{rate}) + (1 - T_{eff}) S_{rate}], \quad (1)$$

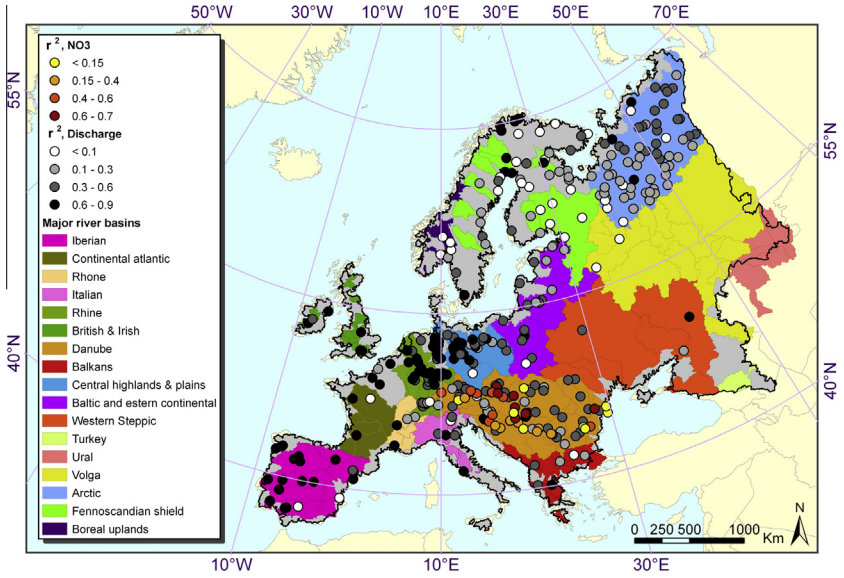


Fig. 4. Overall modeling results at observed discharge and nitrate stations. r^2 is used as an index showing the goodness of fit.

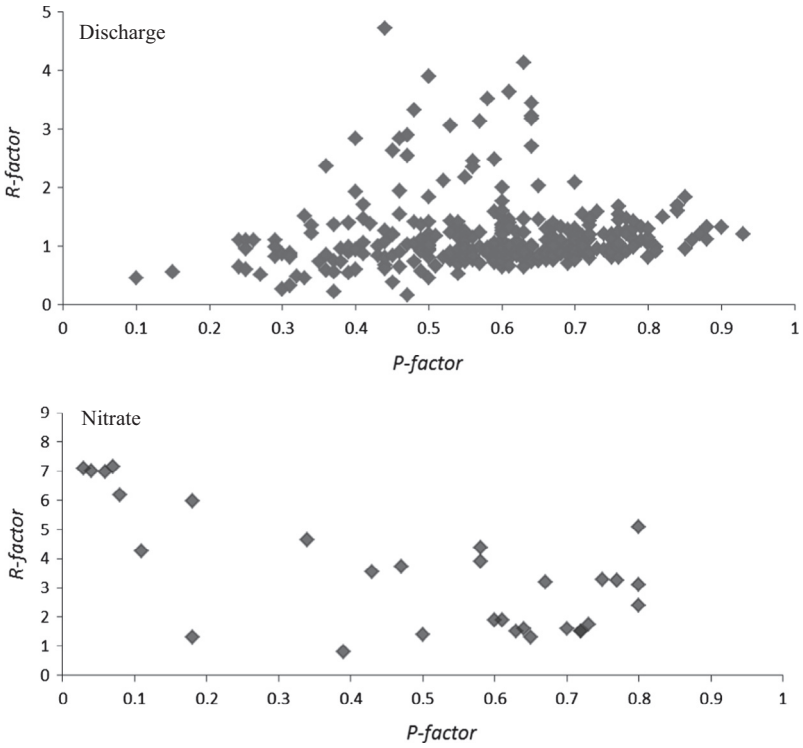


Fig. 5. *P-factor* versus *R-factor*. *P-factor* is the percentage of observed data bracketed by the 95PPU. *R-factor* is the width of the 95PPU band. A working value of >0.7 for *P-factor* and <1.5 for *R-factor* is recommended.

where L_N is the nitrate loading to river from the subbasins contributing to outlet i (g day^{-1}), α_i is a correction factor we added to adjust the quantity of input load based on the long-term average nitrate load of the river at outlet i, I_N is the average input of nitrogen from household to wastewater (g N P_e^{-1} day^{-1}), S_{rate} is the percentage of the population connected to any kind of sewage treatment, and T_{eff} is the wastewater treatment efficiency assumed to be 0.8.

Three major crops, maize, wheat and barley, were considered in this study. These crops were allocated to "agricultural land" in landuse maps and their planting areas in each subbasin proportioned to their planted areas in each country as indicated in MIRCA2000 report (Portmann et al., 2010). Irrigated and rainfed cropping areas were differentiated based on the MIRCA2000 report at a spatial resolution of 5 arc min. Crop yield for model calibration was obtained from Monfreda et al. (2008) at 5 arc minute resolution and from FAOSTAT at country level.

Twenty-five different management plans were designed across Europe based on the crop type, planting and harvest dates, winter or summer crops, irrigated or rain-fed applications. Automatic fertilization scheduling was employed based on plant nutrient deficit

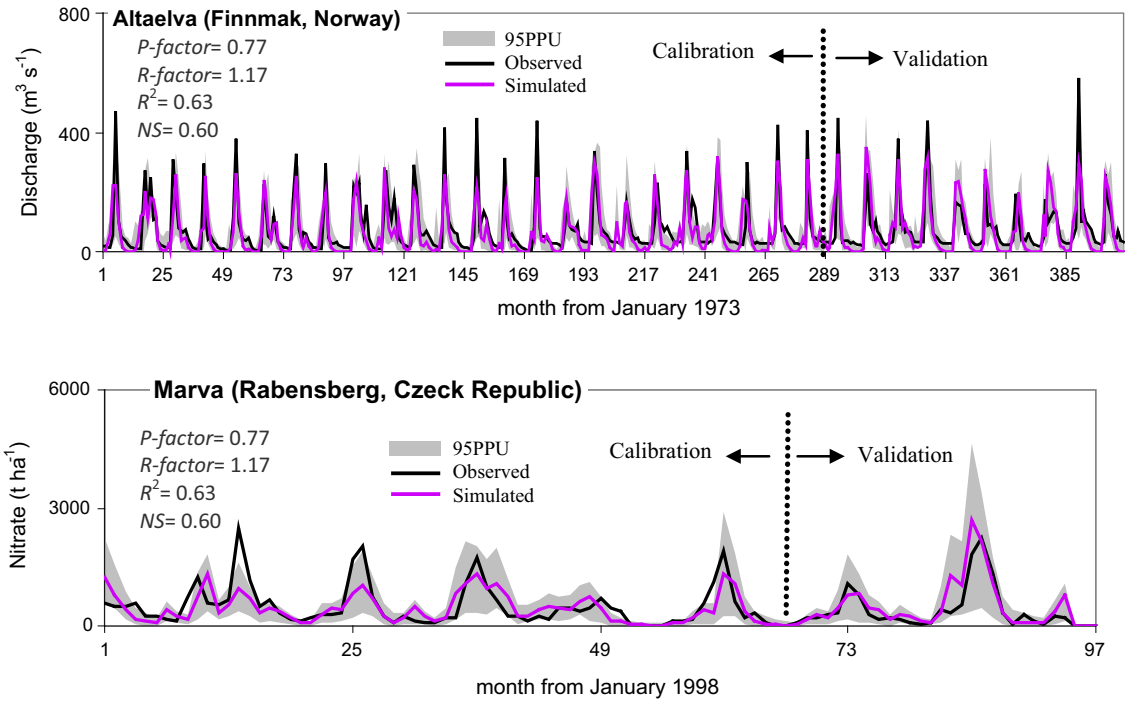


Fig. 6. Illustration of full SWAT-CUP output showing the observed, the best simulation, and the 95% prediction uncertainty (95PPU). *P-factor* is the percentage of observations points bracketed by the prediction uncertainty band, and *R-factor* is the average thickness of the band divided by the standard deviation of the observed values showing the degree of uncertainty. At each station about two-third of the data was used for calibration and the remaining for validation.

and the annual maximum application amount was set to 300 kg N ha^{-1}. Elemental nitrogen and elemental phosphors were selected as the main fertilizer. A few parameters were fitted to match the long-term (1973–2006) average country yields reported by FAO. These included: total heat units for crops to reach maturity (HEAT_UNITS); harvest index (HI_TARG) and biomass target (BIO_TARG), which allow control of biomass production by the plant every year; nitrogen stress factor (AUTO_NSTRS), which ensures that there is almost no reduction of plant growth due to nutrient stress; application efficiency (AUTO_EFF), which allows the model to apply fertilizer to meet harvest removal plus an extra amount to make up for nitrogen losses due to surface runoff/leaching; and fraction of fertilizer applied to top 10 mm of soil (AFERT_SURFACE), where the remaining fertilizer is applied to the first soil layer. Irrigation was based on plant-water-stress automatic scheduling, and withdrawn from outside sources so that the model will add water to the soil until its field capacity. We fitted water stress threshold (Auto_WSTRS), which triggers irrigation, and irrigation efficiency (IRR_EFF). Remaining crop parameters and parameters for non-crop land covers originated from SWAT default database (Neitsch et al., 2005). Parameters sensitive to crop model outputs, such as heat units, were subsequently calibrated to the local conditions.

In this study, potential evapotranspiration (PET) was estimated using the Hargreaves method (Hargreaves et al., 1985) while actual evapotranspiration (ET) was estimated based on Ritchie (Ritchie, 1972). The daily value of leaf area index was used to partition between evaporation and transpiration.

2.4. Calibration, parameterization, and uncertainty analysis

The SUFI-2 algorithm (Abbaspour et al., 2004, 2007) in the SWAT-CUP software package (Abbaspour, 2011) was used for model calibration, validation, sensitivity, and uncertainty analysis. This algorithm maps all uncertainties (parameter, conceptual model, input, etc.) on the parameters (expressed as uniform distributions or ranges) and tries to capture most of the measured data within the 95% prediction uncertainty (95PPU) of the model in an iterative process. The 95PPU is calculated at the 2.5% and 97.5% levels of the cumulative distribution of an output variable obtained through Latin hypercube sampling. For the goodness of fit, as we are comparing two bands (the 95PPU for model simulation and the band representing measured data plus its error), the first author coined two indices referred to as "*P-factor*" and "*R-factor*" (Abbaspour et al., 2004). The *P-factor* is the fraction of measured data (plus its error) bracketed by the 95PPU band and varies from 0 to 1, where 1 indicates 100% bracketing of the measured data within model prediction uncertainty (i.e., a perfect model simulation considering the uncertainty). The quantity (1 – *P-factor*) could hence be referred to as the model error. For discharge, we recommend a value of >0.7 or 0.75 to be adequate. This of course depends on the scale of the project and adequacy of the input and calibrating data. The *R-factor* on the other hand is the ratio of the average width of the 95PPU band and the standard deviation of the measured variable. A value of <1.5, again depending on the situation, would be desirable for this index (Abbaspour et al., 2004, 2007). These two indices are used to judge the strength of the calibration and validation. A larger *P-factor* can be achieved at the expense of a larger *R-factor*. Hence, often a balance must be reached between the two. In the final iteration, where acceptable values of *R-factor* and *P-factor* are reached, the parameter ranges are taken as the calibrated parameters. SUFI-2 allows usage of ten different objective functions such as r^2, Nash-Sutcliff (*NS*), and mean square error (*MSE*). In this study we used br^2 for discharge and nitrate loads. The efficiency criterion was defined as (Krause et al., 2005):

$$\phi = \begin{cases} |b|r^2 & \text{for} \quad |b| \le 1 \\ |b|^{-1}r^2 & \text{for} \quad |b| > 1 \end{cases}, \tag{2}$$

where r^2 is the coefficient of determination and b is the slope of the regression line between the measured and simulated variables. The

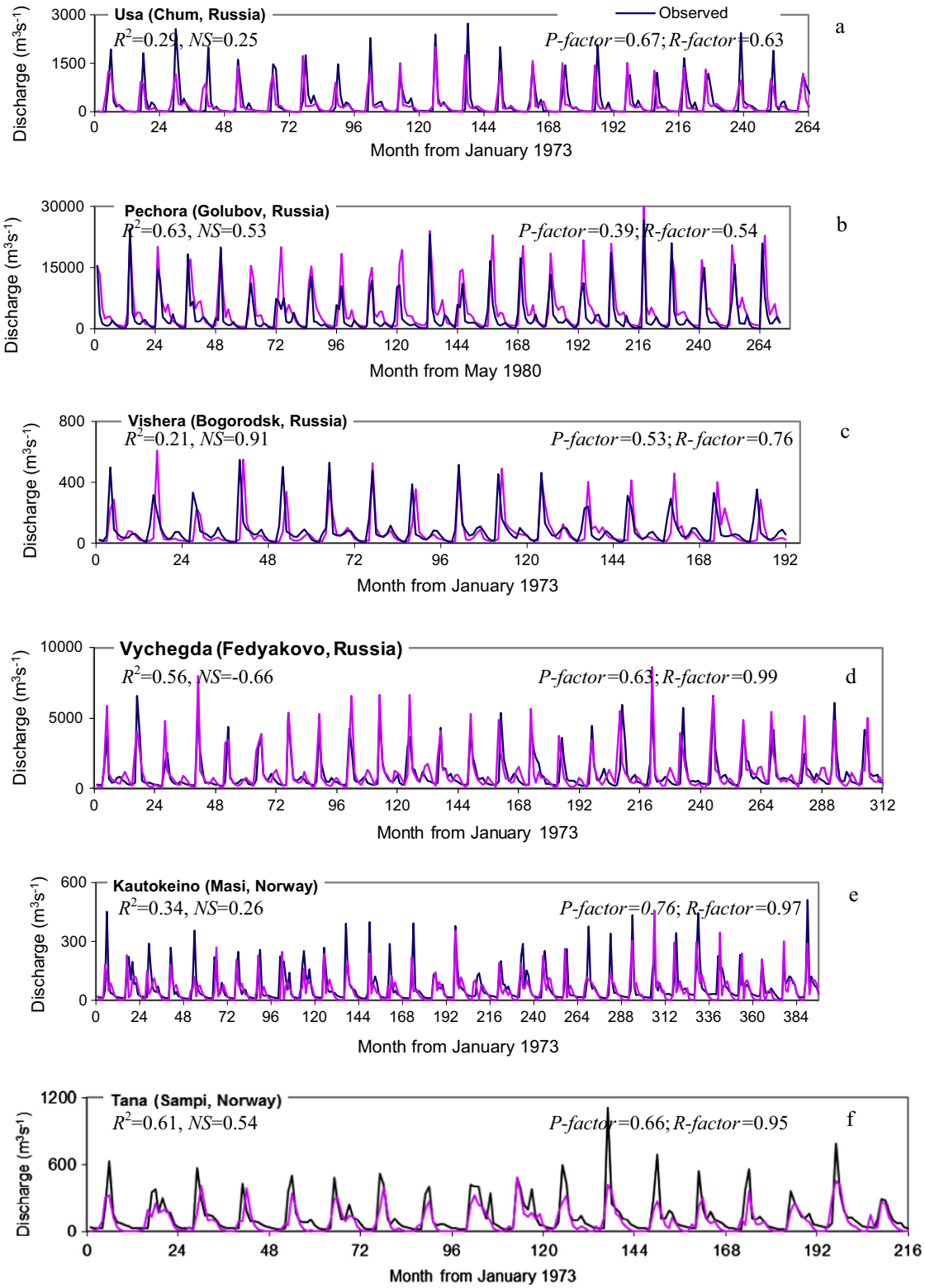


Fig. 7. Illustration of model results for some major rivers across Europe. The figures covers both calibration and validation periods. The statistics r^2, *NS*, and α are for both periods, *P-factor* and *R-factor* are for validation period only.

objective function containing two variables (e.g., river discharge and nitrate load) at multiple sites was formulated as:

$$\Theta = \frac{1}{w_{v1}+w_{v2}}\left[\frac{w_{v1}}{\sum_{i=1}^{n_1} w_i}\sum_{i=1}^{n_1} w_i\phi_i + \frac{w_{v2}}{\sum_{i=1}^{n_2} w_i}\sum_{i=1}^{n_2} w_i\phi_i\right], \quad (3)$$

where w_{v1} and w_{v2} are weights of the two variables, n_1 and n_2 are the number of discharge and nitrate stations, respectively, and w_i's are the weights of variables at each station. The function ϕ, and consequently Θ vary between 0 and 1. In this form, the objective function, unlike, for example, *NS*, which may range from $-\infty$ to 1, is not

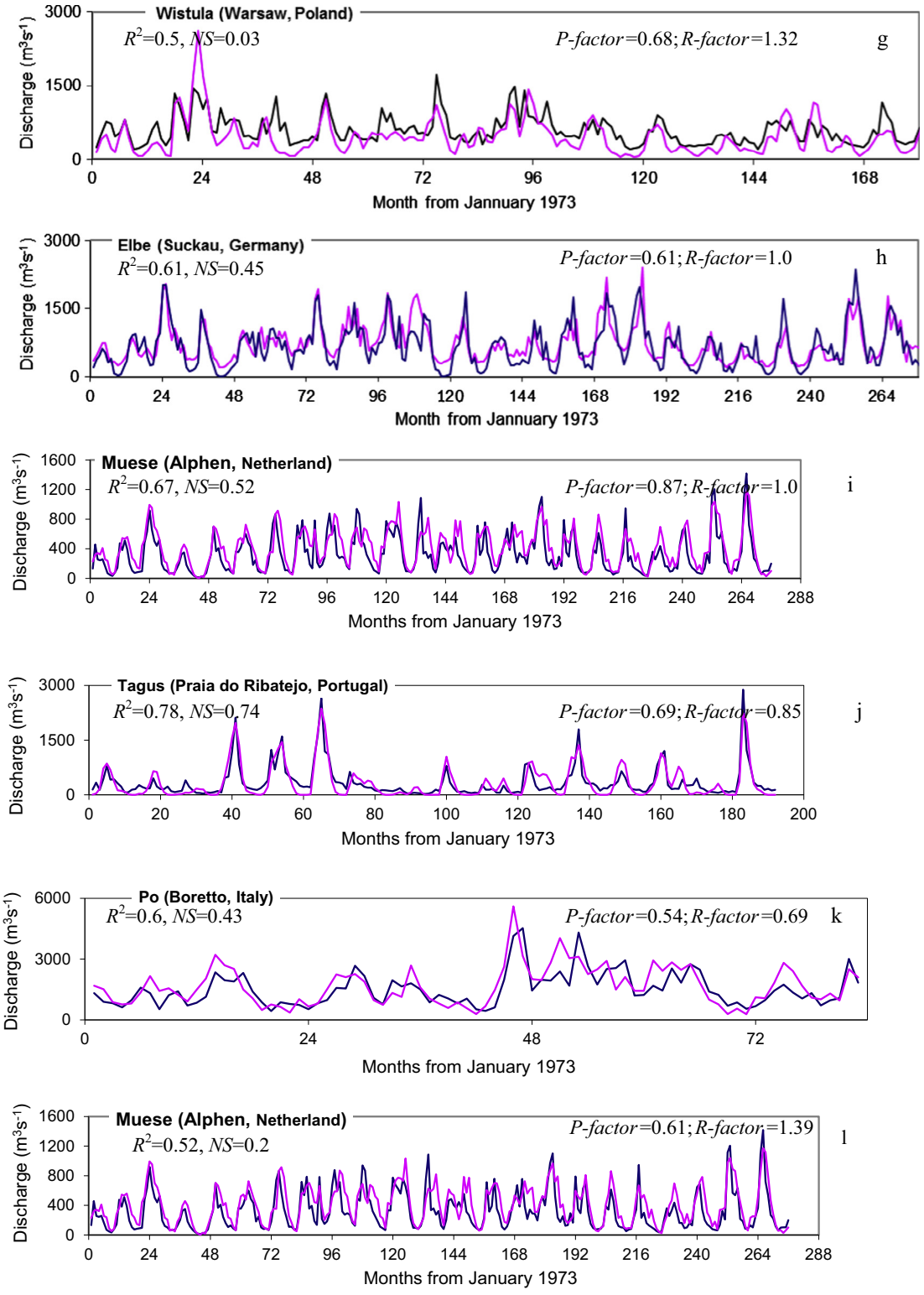


Fig. 7 (*continued*)

dominated by any one or a few very badly simulated stations. Weights in Eq. (3) become critical if an objective function such as mean square error is used, but because of using br^2 they do not make a significant difference to model calibration results. For this reason we set all weights to 1. For crop yield, which was calibrated after an initial calibration of the model for discharge and nitrate, we used mean square errors for each crop as the objective function:

$$\phi_3 = \frac{1}{n_3}\sum_{i=1}^{n3}(Y_i^o - Y_i^s)^2, \quad (4)$$

where n_3 is the number of sites with wheat, barley, and corn yield data, Y^o (t ha^{-1}) is the observed yield, and Y^s (t ha^{-1}) is the simulated yield.

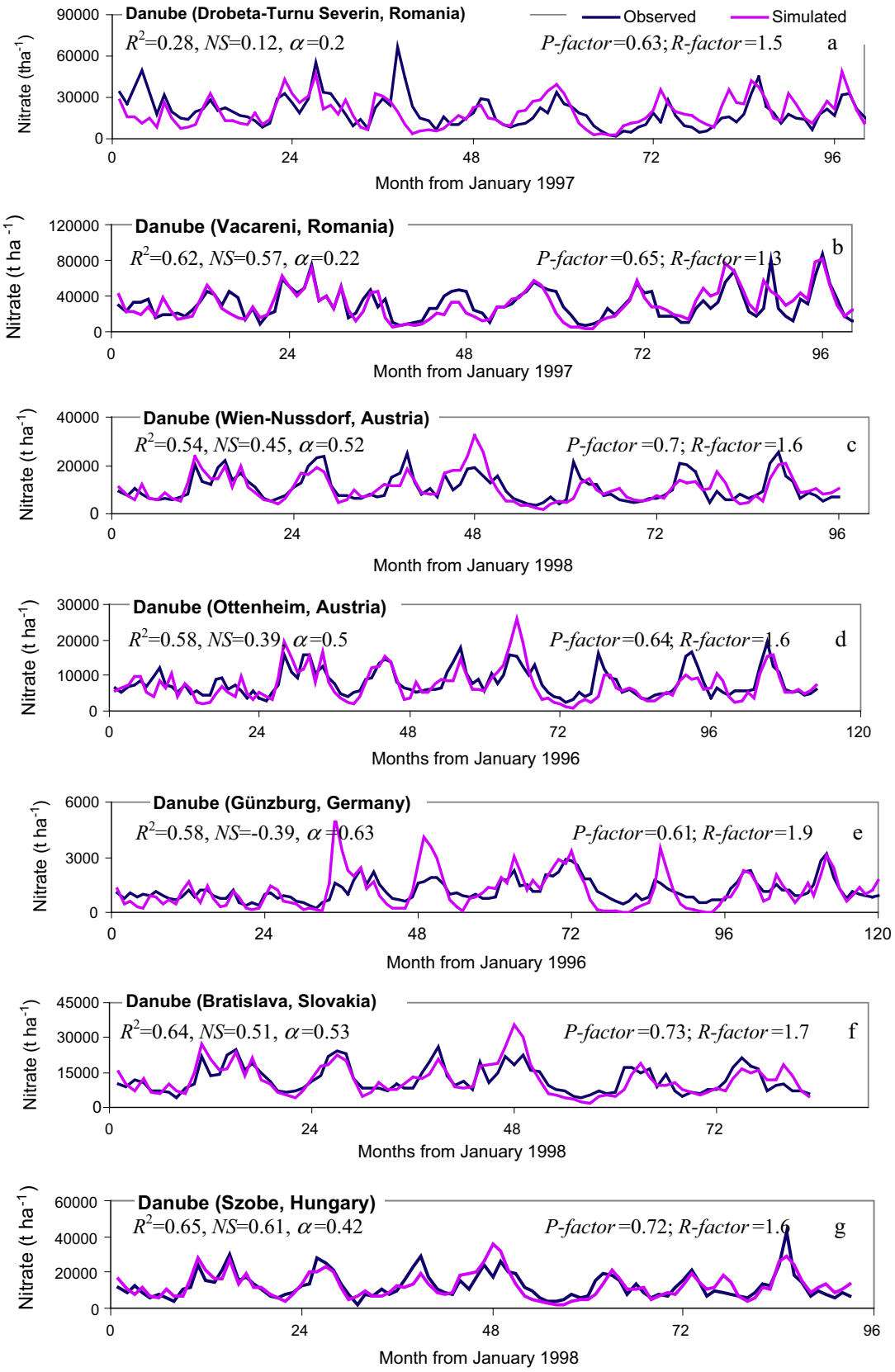


Fig. 8. Examples of model simulation results for nitrate loads on different locations on the Danube River and on other rivers, which are mostly tributaries of Danube. The Figures covers both calibration and validation periods. The statistics r^2, *NS*, and α are for both periods, *P-factor* and *R-factor* are for validation period only.

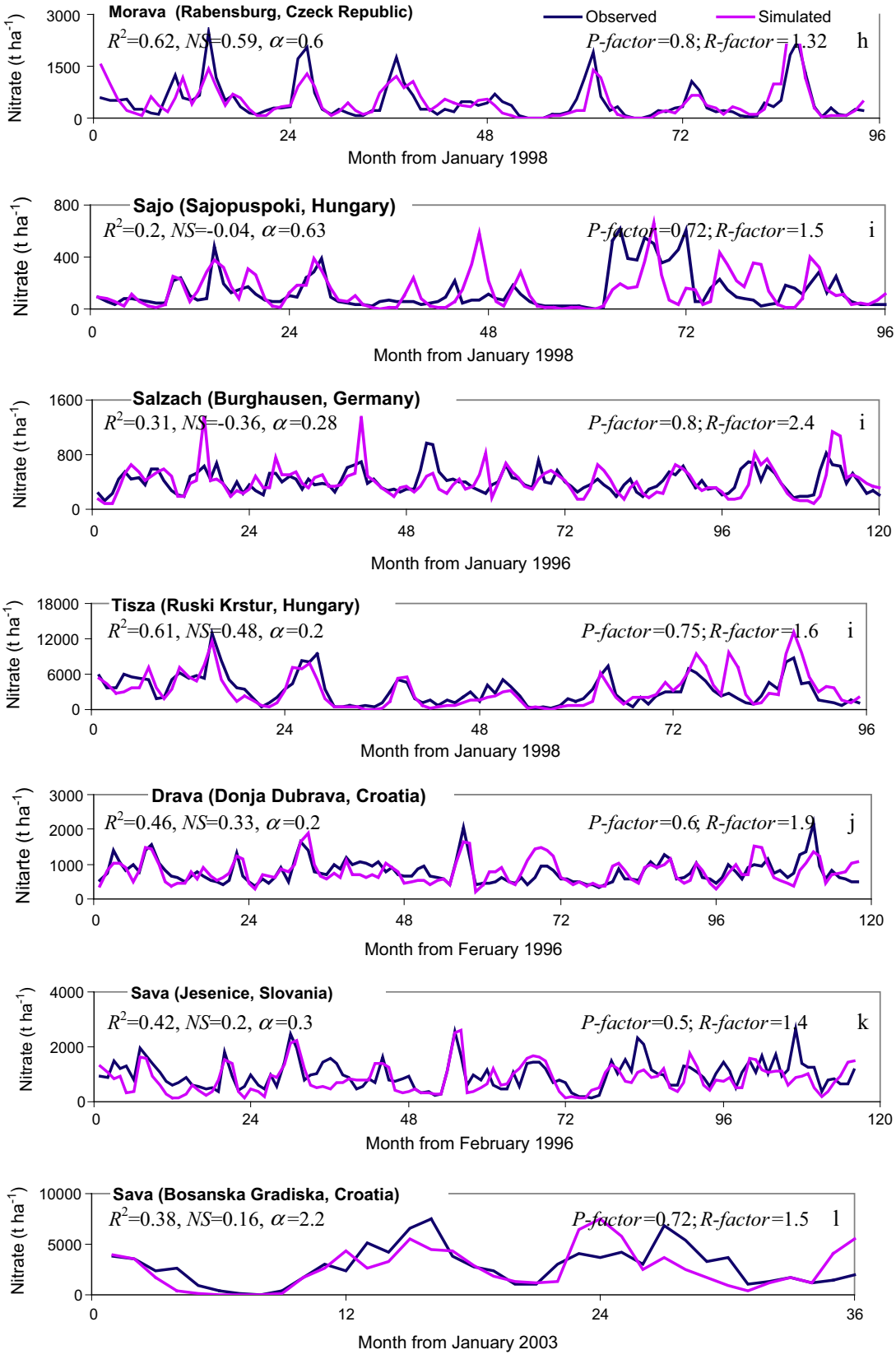


Fig. 8 (*continued*)

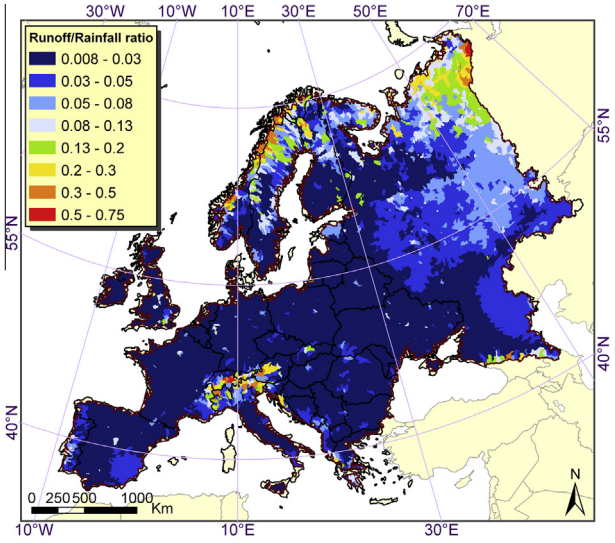


Fig. 9. Map of runoff ratio (runoff/rainfall) providing an indication of erosion potential.

2.4.1. Calibration protocol for large-scale distributed models

To calibrate the model we used the following general approach:

(1) Build the model with ArcSWAT using the best parameter estimates based on the available data, literature, and analyst's expertise. There is always more than one data set (e.g., soil, landuse, climate, etc.) available for a region. For Europe we were in possession of four different landuse maps, two different soil maps, and four different sets of climate data (Table 1). Hence, initially several models were built and ran without any calibration (referred as the default model) with different databases. The model results were compared with observations (discharge and nitrate) and the best overall performing database was selected for further analysis. It should be noted that the performance of the default model should not be too drastically different from the measurement. If so, often calibration can be of little help.

(2) Use the best default model to calibrate. Based on the performance of the default model at each outlet station, relevant parameters in the upstream subbasins are parameterized using the guidelines summarized in Table 2. This procedure results in regionalization of the parameters.

(3) Based on parameters identified in step 2 and one-at-a-time sensitivity analysis, initial ranges are assigned to parameters of significance. Experience and hydrological knowledge of the analyst is also of great asset in defining parameter ranges. In addition to the initial ranges, user-defined absolute parameter ranges are also defined for every SWAT parameter in SWAT-CUP where parameters are not allowed to be outside of this range.

(4) Once the model is parameterized and the ranges are assigned, the model is run some 300–1000 times, depending on the number of parameters, speed of the model execution, and the system capabilities. Great time saving could be achieved by using the parallel processing option of SWAT-CUP (Rouholahnejad et al., 2012).

(5) After all simulations are completed; the post processing option in SWAT-CUP calculates the objective function and the 95PPU for all observed variables in the objective function. New parameter ranges are suggested by the program for another iteration, which modifies the previous ranges focusing on the best parameter set of the current iteration (Abbaspour et al., 2004, 2007).

(6) The suggested new parameter ranges could be modifies by the user using Table 2 and one-at-a-time sensitivity analysis again. Another iteration is then performed. The procedure continues until satisfactory results are reached (in terms of the *P-factor* and *R-factor*) or no further improvements are seen in the objective function. Normally, three to five iterations are sufficient for satisfactory results. More detailed information could be found in Abbaspour et al. (2004, 2007) and Rouholahnejad et al. (2012).

3. Results

3.1. Calibration/validation of discharge

In the preliminary analyses we tested different landuse and climate databases and selected Modis landuse map and CRU's 0.5° global climate database. The selection was based on a comparison of measured discharges of major rivers against default model simulations (Table 3). We chose the CRU climate database because it had the smallest overall error. The difference from different landuses however was not very significant. To compare the model results here with other past or future works, we should keep in mind the inputs used to develop this model as using, for example, a different climate database, would result in different outputs.

The overall performance of the model in terms of r^2 has quiet satisfactory results for the European model (Fig. 4). 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 may results in some individual outlets to be poorly simulated. A more local simulation with a smarter algorithm that could account for nested outlets would certainly result in a better overall simulation. A graph of *P-factor* versus *R-factor* (Fig. 5) shows that a large number of the 326 outlets fall in desirable regions of the two indices. A few outlets were located downstream of dams. These outlets would of course not be well simulated by SWAT in any way as their flow regimes are entirely controlled by the operation of the dam or reservoir. Hence, we plotted these factors after removing those outlets, which had quite large *R-factor*. For nitrate, the uncertainties, as indicated by the *R-factor* are generally larger. As SUFI-2 is iterative, each iteration results in a reduction of parameter uncertainties causing a narrower 95PPU band, which subsequently results in a smaller *P-factor*. Hence, a balance must be struck between the two indices.

In general, model uncertainties are due to: (i) conceptual simplifications (e.g., SCS curve number method for flow partitioning), (ii) processes occurring in the watershed but not included in the program (e.g., wind erosion, wetland processes), (iii) processes that are included in the program, but their occurrences in the watershed are unknown to the modeler or unaccountable because of data limitation (e.g., dams and reservoirs, water transfers, farm management affecting water quality), and (iv) input data quality. In large watershed applications one expects to have all these forms of uncertainties, which explains some of the large prediction errors. In this study, outlets on rivers with small flows showed particularly large uncertainties.

In the following we describe some details of the important river basins of Europe (colored[1] in Fig. 4). The characteristics of major rivers and the model performances are assessed for both discharge and

[1] For interpretation of color in Fig. 4, the reader is referred to the web version of this article.

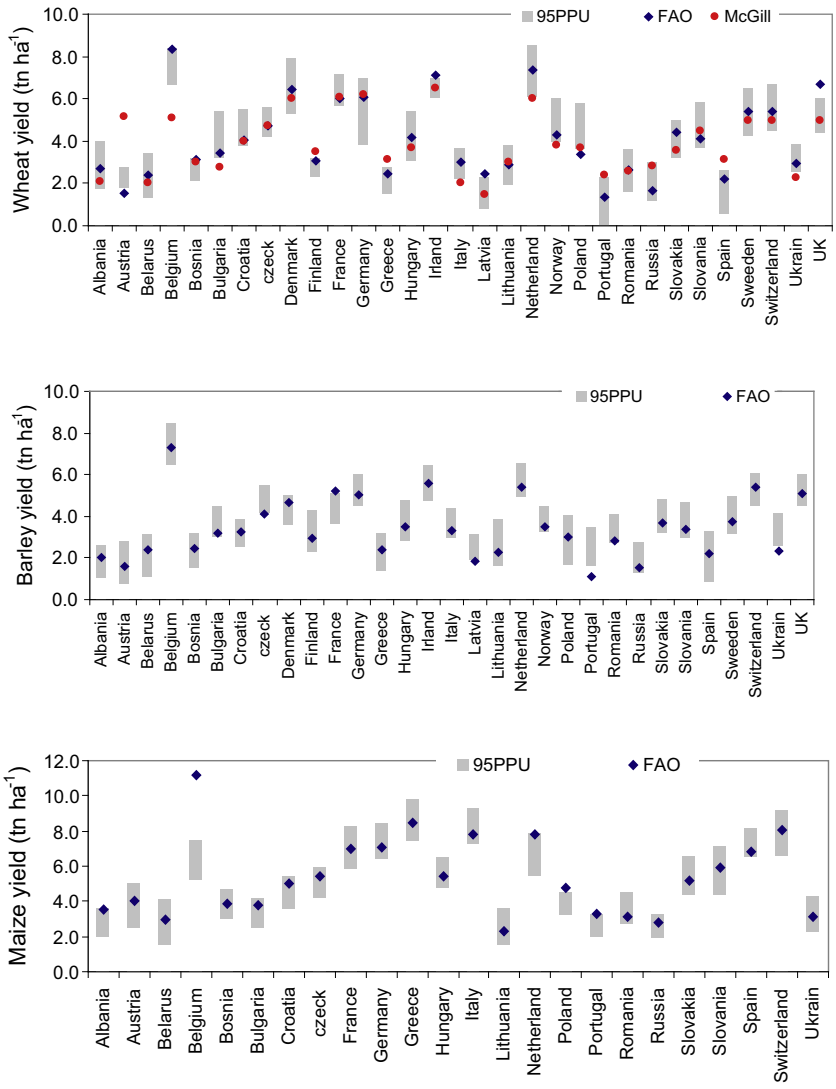


Fig. 10. Comparison of the simulated long-term (1973–2006) average crop yield with those reported by FAO and McGill University for wheat.

nitrate. SWAT-CUP produces output results at each station as 95PPU as well as showing the best fit (e.g., the simulation run with the best objective function value) (Fig. 6), but for simplicity and clarity of presentation we only show the calibration/validation results for the best simulation as a continuous graph (Figs. 7 and 8) and report the overall statistics.

The river discharges in the northeast corner of Europe in Russia are governed by snow melts. They are frozen in winter and active in the spring and summer. This is evident from the regular peaks in the spring and near zero discharges in winter months. Examples from rivers Usa (length = 565 km, Basin area = 93,600 km^2), Pechora (1800 km, 322,000 km^2), and Vishera (415 km, 31,200 km^2) (Fig. 7a–c) are characterized with freezing around October–November, and staying frozen until the spring thaw in April–May. This region was calibrated using among other parameters SNOEB (initial snow water content), SFTMP (snowfall temperature), SMTMP (snow melt base temperature), SMFMX (melt factor on June 21st), and SMFMN (melt factor on December 21st), as well as by increasing the CRU rainfall throughout the year by about 20%. Increasing the rainfall was important to catch the peaks of spring and summer discharges. SWAT accumulates the rainfall during fall and winter as snow, and then melts the snow in spring and summer. The larger discrepancy in the recession dynamics between observed and simulated discharges in the Pechora River (Fig. 7b) is due to the complex branching of this main river before the point of measurement (Fig. 3c). Given the complexity of the system, it is surprising that the model can still simulate the discharge at this location with such accuracy (r^2 = 0.63, *NSE* = 0.53) albeit with small *P-factor* (0.39).

Further west still in the European part of Russia, is the Vychegda River (1100 km, 120,000 km^2) (Fig. 7d), a tributary to the Northern Dvina, with a large average discharge rate of over 1000 m^3 s^{-1}. Simulation of this navigable river, although visually very good, suffers from a small lag in the simulation, which could not be corrected by the parameters listed in Table 2. This could have been caused by a small shift in the CRU rainfall governing this outlet in SWAT.

In the far north of Europe is Kautokeino River (Fig. 7e) in Norway (240 km, 7400 km^2). This river drains into Alta river and then to the Norwegian Sea and is fed by a complex system of lakes, peatlands, and tributaries. The result shows relatively good simulation of the river discharge dynamics (r^2 = 0.34).

The Tana River (Fig. 7f) in northern Scandinavia (361 km, 16,000 km^2) flows in pristine condition through Norway and the Lapland region of Finland. This river is very well simulated (r^2 = 0.61, *NS* = 0.54) with a rather large *P-factor* (0.66).

Closer to the heart of Europe, the Vistula River (Fig. 7g) is the longest river in Poland (1047 km, 194,000 km^2). The simulation although captures the overall dynamics of the flow (r^2 = 0.5) has a small *NS* (0.03) indicating mismatches mostly in timing. This could be expected given the diversity of water uses and management in the region.

Further west is the Elbe River (Fig. 7h), which is the fourth largest in Europe (1094 km, 148,000 km^2). It rises in northern Czech Republic before entering Germany and flowing into the North Sea. This large river is surprisingly well simulated (r^2 = 0.61, *NS* = 0.45) given the intense management along the course of the river.

Muese (Fig. 7i) is a major European river, beginning in France and flowing through Belgium and The Netherlands and draining into the North Sea (925 km, 35,000 km^2). With an r^2 = 0.67 and *NS* = 0.52, this river is quite well simulated in the European model.

Tagus (Fig. 7j) is the longest river on the Iberian Peninsula (1038 km, 80,100 km^2). Although this river is highly utilized with several dams along its course, the model simulation at Praia do Ribatejo is rather excellent with r^2 = 0.78 and *NS* = 0.74.

Going through south of Europe is the river Po (680 km, 74,000 km^2) (Fig. 7k). Po crosses Italy, Switzerland, and France and has a large average discharge of about 1500 m^3 s^{-1}. This river is also quite well simulated with r^2 = 0.60 and *NS* = 0.43.

Finally, Danube River (Fig. 7l) in Central Europe is the continent's second largest river (After the Volga) and is classified as an international waterway (2860 km, 817,000 km^2). Simulations of discharges along the river Danube were quite good except when the outlet was immediately below a dam or a reservoir. The simulation at the location near Tulcea is quite good r^2 = 0.52 and *NS* = 0.2 albeit with a larger uncertainty (*R-factor* = 1.39).

3.2. Calibration/validation of nitrate

Calibration of nitrate in the European model reflects mostly measurements at the Danube watershed as we only had nitrate data from this region. In the SWAT model, the main sources of nitrate in the rivers are from the soil organic matter, fertilizer input (diffuse source), discharges from water treatment plants, and direct sewage inputs into the rivers (point sources). The point source loadings are made available to the rivers on equal daily intervals; hence do not contribute much to the dynamics of nitrate concentration in the rivers. The dynamics is mostly governed by the fate and transport of the fertilizers in the soil, decomposition of organic matter, and the climate. A summary of calibration (and validation) results in different sections of the Danube and other rivers (Fig. 8) in the Danube Basin show excellent accounting of the dynamics of nitrate transport. The quantity of nitrate was adjusted through the factor α in (Eq. (1)) for each outlet.

The results of nitrate load in Danube River show consistently good simulation at different locations (Fig. 8a–g), indicating a good accounting of agricultural input in the model. It is interesting to note the variation in nitrate loadings in different sections of the same river as it crosses different countries with the largest nitrate load observed in Romania (Fig. 8a, b) and Hungary (Fig. 8g) and the smallest ones in Germany (Fig. 8c).

In other rivers (Fig. 8h–n), most of which are also tributaries of Danube we also see very good simulations. The Morava River (350 km, 26,600 km^2) is a blackwater river in Central Europe (Fig. 8h). Morava flows into the Danube with an average discharge rate of 120 m^3 s^{-1} and brings much nitrate into Danube. The simulation statistics show very good accounting of nitrate load in this river.

The Sajo is a river (Fig. 8i) in Slovakia and Hungary (230 km, 12,700). About 110 km of the river is in Slovakia and eventually flows into the River Tisza. Due to unusually large nitrate peaks in the summer of 2000, the simulation statistics are quite modest at r^2 = 0.2 and NS = −0.04 with a large reduction in point source inputs of α = 0.6.

The Salzach (Fig. 8j) is a river in Austria and Germany (225 km, 2600 km^2). It is a tributary of the Inn. The river has relatively low nitrate load as it drains several alpine pastures at its headstreams. Except for a couple of simulated peaks in the fall of 1996 and 1997, the observation is quite well reproduced by the simulation.

The Tisza (Fig. 8k) is one of the main rivers of Central Europe (970 km, 156,000 km^2). It rises in Ukraine, and flows roughly along the Romanian border before entering Hungary and then flowing into Danube. With a mean annual discharge of about 800 m^3 s^{-1}, the Tisza contributes the largest nitrate load to the Danube. Tisza is quite well simulated at r^2 = 0.61 and *NS* = 0.48 with a small adjustment to Eq. (1) in terms of the α coefficient.

Drava (Fig. 8l) is a tributary of the Danube in Central Europe (707 km, 11,800 km^2). It begins in Italy and flows east through Austria, Slovenia, Croatia and Hungary, before it joins the Danube. The nitrate load is relatively small and is quite well simulated (r^2 = 0.46, *NS* = 0.33, α = 0.2).

The Sava River (Fig. 8m, n) is also a tributary of Danube in Southeast Europe (990 km, 97,700 km^2). Sava is the third longest tributary of the Danube, as well as the greatest by volume of water. The river increases in nitrate load as it crosses Croatia (Fig. 8n).

Because of lack of data, other pollutants such as phosphor and sediment were not modeled. However, a plot of rainfall runoff ratio (Fig. 9) provides an indication of potential erosion and non-point source pollution (e.g., to some extent phosphor as it moves attached to sediment) carried by surface runoff.

3.3. Calibration/validation of crop yield

The last variable used to calibrated the European model was the yields of wheat, barley, and maize crops, as these cover the largest harvested areas of Europe. Simulating yield is in general a difficult task because of the limitations in the input data, especially the management data such as planting and harvesting time, fertilizer and irrigation water inputs, losses to pests and droughts, etc. As explained before, we specified 25 different management scenarios to cover the spatial heterogeneity of management across Europe. We calibrated crop parameters to achieve a reasonable long-term (1973–2006) average yield corresponding to data reported by FAO. The simulations (Fig. 10) indicate that in most countries the long-term average yield, depicted as 95PPU, contains the FAO reported data. This result indicates that on a long-term average basis, factors such as soil moisture, aquifer recharge, and actual evapotranspiration are relatively well simulated; hence increasing our confidence on the model performance. For wheat, we also plotted the data from McGill estimates, which shows both the uncertainty of the reported data and also the reasonability of the 95PPU prediction. The reported model uncertainty reflects uncertainty in hydrologic parameters as well as the climate variation during the simulation period.

3.4. Quantification of water resources and water balance

Water resources are often quantified in terms of blue water flow, which is water yield plus deep aquifer recharge (FAO, 2003). Following Falkenmark and Rockström (2006), green water storage (soil moisture) has been now widely recognized as a crucially important component of water resources. Using the calibrated model, the long-term (1973–2006) average internal blue water resources, green water storage and green water flow (ET) are mapped for the continent (Fig. 11). The results show a wide range of spatial (Fig. 11a) and temporal (Fig. 11b) variation of internal blue water resources. The latter is depicted by the coefficient of

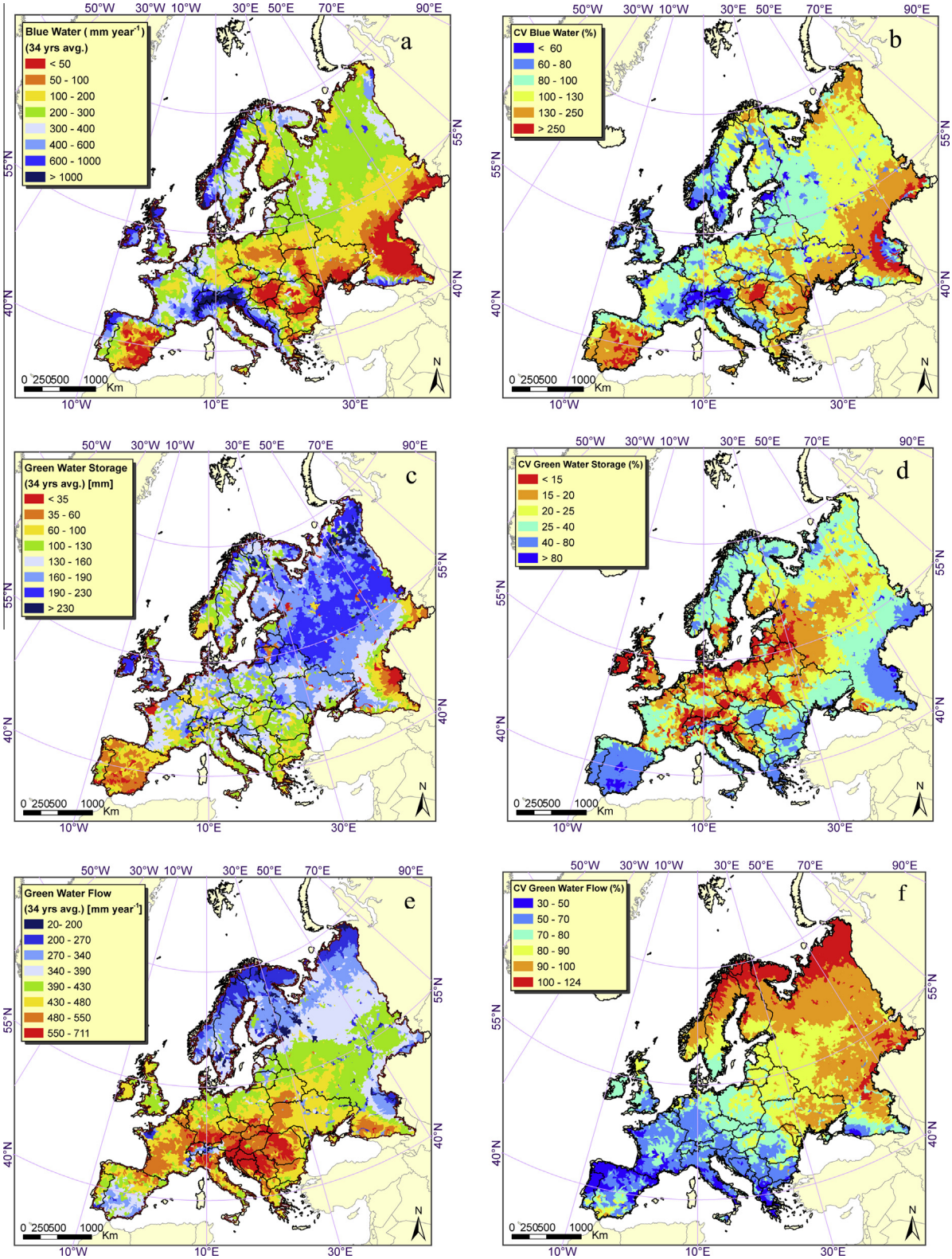


Fig. 11. Illustration of the (a) blue water resources, (c) green water storage, and (e) green water flow across Europe along with their coefficients of variations (CV). CV shows the temporal variation over the 34 years of simulations (1973–2006). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

variation for the 34-year simulation. Also shown are the long-term distributions of soil moisture (green water storage) (Fig. 11c) and actual evapotranspiration (green water flow) (Fig. 11e) along with their coefficients of variation. Areas with larger soil moisture and smaller coefficient of variation have higher potential for development of rainfed (green) agriculture.

To provide an overview of water endowments of each country, we calculated the amount of internal blue water resources and

Table 4
Average annual precipitation (1973–2006) and the 95PPU ranges for the components of freshwater availability in the European countries.

Country	Area (km^2)	Precipitation (km^3 year^{-1})	Blue water flow (km^3 year^{-1})	Greenwater storage (km^3)	Green water flow (km^3 year^{-1})
Albania	28,750	29	9–21	17–91	10–12
Austria	83,870	95	39–67	64–267	36–42
Belarus	207,600	128	20–58	211–429	80–92
Belgium	30,530	27	8–17	24–84	13–15
Bosnia	51,210	54	16–35	45–113	27–32
Bulgaria	111,000	68	7–27	72–272	44–53
Croatia	56,590	58	14–32	42–71	23–27
Czeck	78,870	51	8–21	67–151	32–37
Denmark	43,090	34	3–6	9–22	3–4
Estonia	45,230	30	8–18	25–65	9–11
Finland	338,420	186	33–117	330–602	85–100
France	549,190	465	144–301	398–842	192–227
Germany	357,130	254	56–128	283–699	139–160
Greece	131,960	88	16–49	48–138	27–35
Hungary	93,030	55	1–11	69–158	40–47
Ireland	70,280	76	33–60	47–85	15–17
Italy	301,340	294	101–203	166–419	84–102
Latvia	64,480	42	11–25	59–111	20–23
Lithuania	63,500	42	10–24	58–105	22–26
Luxamburg	2586	3	1–2	1–7	1–2
Macedonia	25,713	16	2–7	16–45	10–12
Moldova	33,846	18	2–7	33–63	13–16
Montenegro	13,812	16	6–13	7–15	5–6
Netherlands	41,549	33	7–13	19–46	9–10
Norway	323,790	307	126–306	172–386	52–63
Poland	312,680	187	29–77	265–579	115–132
Portugal	92,090	75	22–55	39–147	25–32
Romania	238,390	154	15–53	184–465	106–124
Russia	3,669,247	2133	347–1400	4114–6433	1226–1433
Serbia	88,360	67	10–30	67–125	41–49
Slovakia	49,040	37	6–16	36–76	23–27
Slovania	20,270	27	10–19	19–33	11–13
Spain	505,600	293	50–150	200–961	158–199
Sweeden	450,300	295	76–222	364–526	107–125
Switzerland	41,280	69	38–65	31–83	15–18
Ukrain	603,550	346	42–135	541–540	212–245
UK	243,610	257	119–209	208–433	82–98

green water storage at the country level in total and in per capita terms (Table 4). It should be noted that the modeled green water storage and flow were solely calibrated indirectly through crop yield and infiltrated water as there were no direct observations.

4. Discussion

4.1. Conditionality of the calibrated models

It should be pointed out from the onset that calibration of watershed models is subjective and no automatic calibration algorithm can replace the knowledge of the analyst vis-à-vis watershed hydrology and calibration issues. Therefore, calibration and uncertainty analysis are intimately linked and no calibration results should be presented without a quantification of the degree of uncertainty in the model prediction. It has been shown before (Arnold et al., 2012; Stisen et al., 2011; Henriksen et al., 2003; Abbaspour et al., 1997, 1999; Yang et al., 2008) that 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 the results of calibrated models along with their uncertainties are conditioned on the assumptions explicit or implicit in the hydrological model and model calibration. It is, therefore, up to the analyst to quantify model prediction uncertainty and to communicate it to model users, and the responsibility of the model

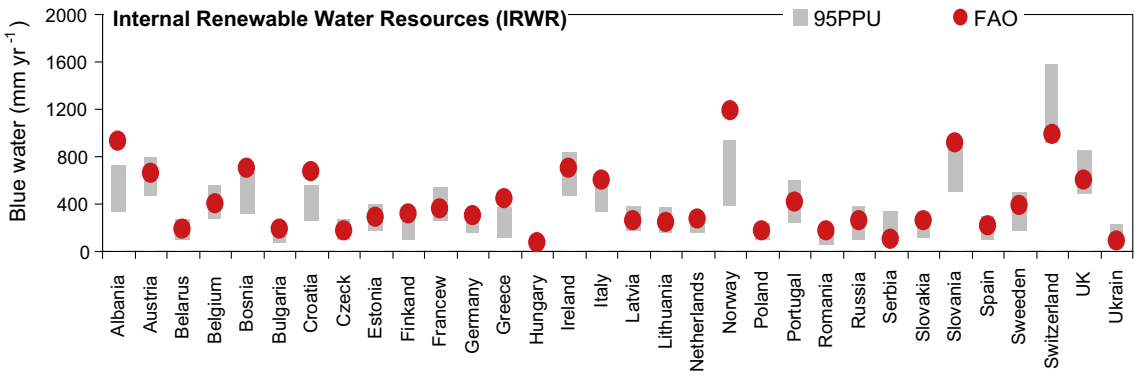


Fig. 12. Comparison of the country-based blue water flow results with those reported by FAO (AQUASTAT, 2013). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

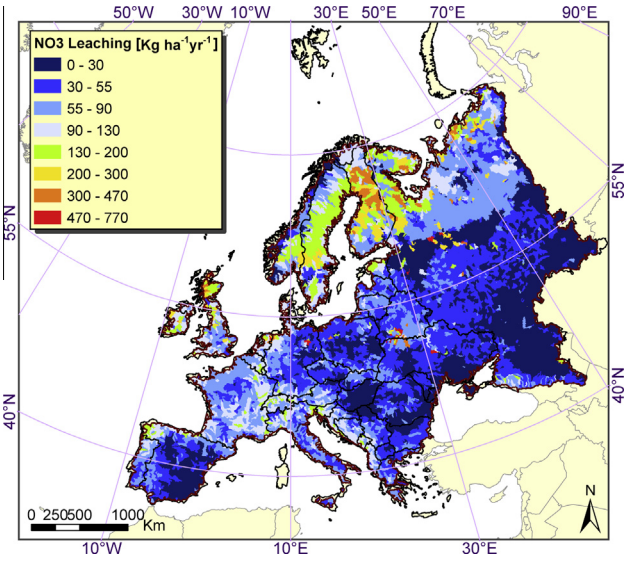


Fig. 13. Map of nitrate leaching from the bottom of the root zone. The image mostly reflects nitrate leaching from the organic constituent of the soil and underestimates the leaching of nitrogen due to excess fertilizer application and application of manure.

users to properly apply those uncertainties by using the model within the conditions of the model set up and calibration assumptions. Analysts are often not in the position of policy making. Their role is to provide the policy makers with "high-value information", which are model outputs necessary for making decisions. Hence, model uncertainty as well as model results could play an important role in policy analyses and should both be reported.

4.2. Comparison with FAO and other sources

Although a direct one-to-one comparison of our modeling results is not possible with reported literature due to the different time periods and study-specific assumptions, nonetheless we compared the modeled blue water results with values reported by FAO (AQUASTAT, 2013) (Fig. 12). Data from FAO were selected not to serve as "true" values but because they are very commonly used by government and scientific organizations. The comparison, shows that most of the FAO values fall within the model prediction uncertainty, hence increasing confidence in model results as well as, in this case, the groundwater recharge, which is a component of the blue water. No such comparison was made for soil moisture and actual evapotranspiration as we did not find comparable data in the literature.

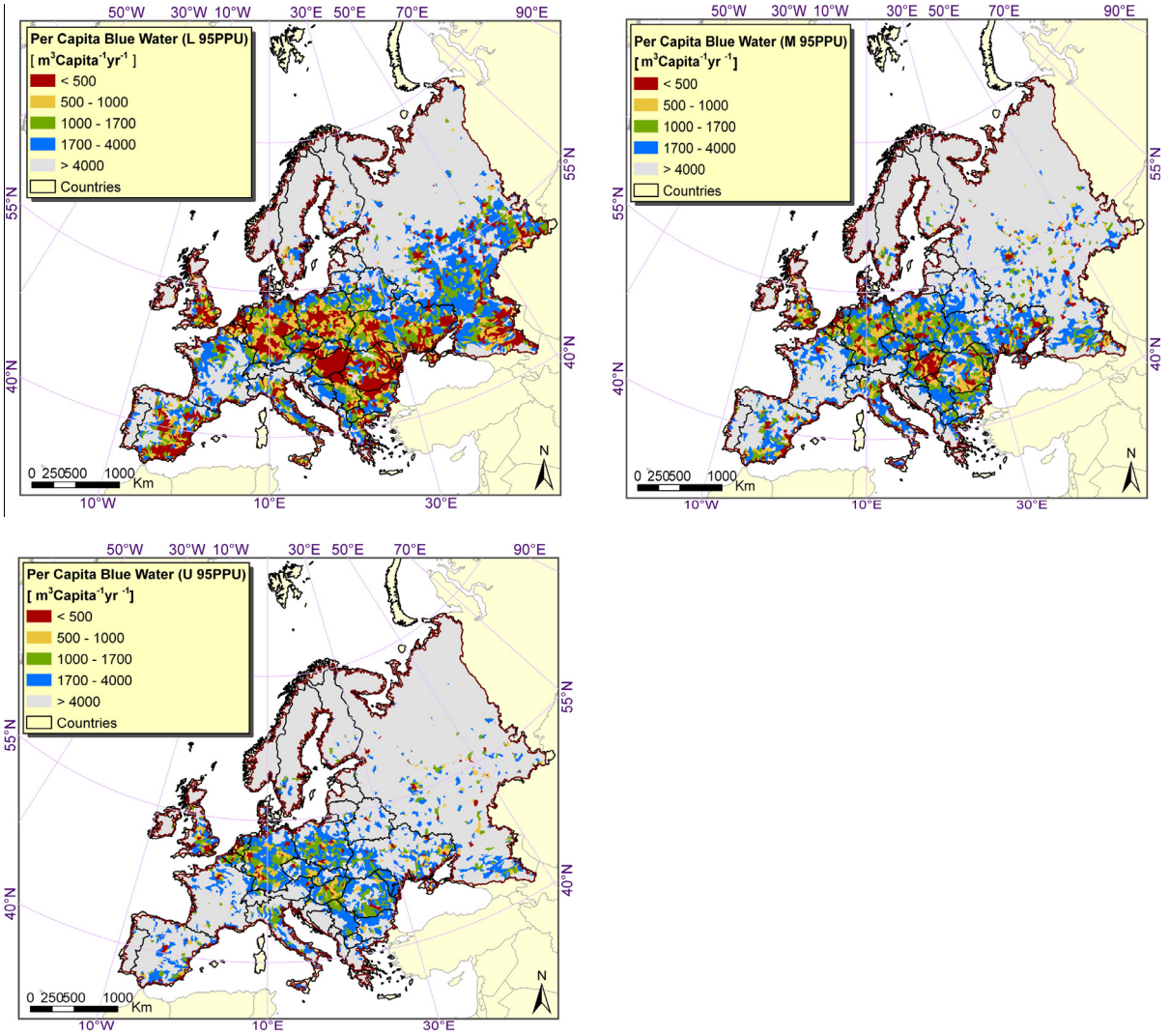


Fig. 14. Water scarcity map across Europe based on the Falkenmark's threshold of 1700 m^3 $capita^{-1}$ $year^{-1}$. The difference between the lower and the upper bounds of model prediction uncertainty highlight the importance of calculating and reporting model prediction uncertainty.

4.3. Nitrogen leaching

As observed nitrates were only available for the Danube Basin, components of nitrate prediction outside this region should be viewed as uncalibrated. In our European model, nitrate leaching beyond the root zone (Fig. 13) is mostly governed by soil organic mater content. For this reason Scandinavian countries show a larger leaching below the root zone than other regions in Europe. The global FAO soil database has two soil layers of 30 and 70 cm deep. Nitrate leaching map shows the (kg N ha^{-1} $year^{-1}$) leached below the 100 cm rooting depth. Fertilizer in the European model was applied as needed by crops; therefore, the amount of fertilizer leached below the rooting depth is quite underestimated in our results, especially for areas where more-than-needed fertilizer is applied. Eurostat's regional environmental statistics shows there is a surplus of nitrogen in most regions within the EU. Very high amounts of surplus are found in the North-West of Europe, mainly the Netherlands, indicating areas, where intensive agriculture may create severe and well known nutrient pollution problems (Fassio et al., 2005). As detail agricultural management information was not available, application of manure and excess fertilizer was not considered in our European model. However, using the current model different scenarios and their effects could by analyzed.

4.4. Water endowments in Europe

While there exists a large number of water scarcity indicators, one of the most widely used and accepted is the water stress threshold, defined by Falkenmark and Widstrand (1992) as 1700 m^3 $capita^{-1}$ $year^{-1}$ (Fig. 14). There are large differences between the lower and upper bonds of water stressed regions because of the model uncertainty. A value between 1700 and 4000 is considered as just adequate (Revenga et al., 2000). This threshold only considers the blue water availability of a country, while ignoring the green water availably. As a rule of thumb, the indicated threshold may be useful for general communication, but it can cause some confusion in viewing water endowments of European countries. For example, the Netherlands has a very low volume of per capital blue water resources. However, in reality, no one would regard the Netherland as a water scarce country. This is mainly because the country has a large amount of green water which provides a major water source to its agriculture.

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 clearly 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 with reference to parameters that are significant to major discrepancies often encountered in calibration. Many of the difficulties and limitations within this continental modeling study were data related and resulted from, among others, (1) limited and unevenly distributed nitrate data and discharge stations with varying time series lengths, (2) limited globally available knowledge of the attributes and especially the management of the reservoirs, and (3) lack of data on soil moisture and/or deep aquifer percolation, which made a desirable calibration/validation of these components impossible, (4) lack of knowledge of agricultural management operations, and (4) 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. As shown, the inherent uncertainties need to be considered, before general conclusions are drawn. 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.

Acknowledgment

The study was supported by the European Community 7th Framework Project GENESIS (226536) on groundwater systems.

References

Abbaspour, K.C., 2011. Swat-Cup2: SWAT Calibration and Uncertainty Programs Manual Version 2, Department of Systems Analysis, Integrated Assessment and Modelling (SIAM), Eawag. Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland. 106 p.

Abbaspour, K.C., van Genuchten, M.Th., Schulin, R., Schläppi, E., 1997. A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. Water Resour. Res. 33 (8), 1879–1892.

Abbaspour, K.C., Sonnleitner, M., Schulin, R., 1999. Uncertainty in estimation of soil hydraulic parameters by inverse modeling: example lysimeter experiments. Soil Sci. Soc. Am. J. 63, 501–509.

Abbaspour, K.C., Johnson, C.A., van Genuchten, M.Th., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. J. Vadouse Zone 3 (4), 1340–1352.

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.

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.

AQUASTST, FAO, 2013. <http://www.fao.org/nr/water/aquastat/water_res/index.stm>.

Arnell, N.W., 1999. Climate change and global water resources. Glob. Environ. Change Hum. Policy Dimens. 9, S31–S49.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment. Part I: Model development. J. Am. Water Resour. Assoc. 34 (1), 73–89.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Allen, P.M., 1999. Continental scale simulation of the hydrologic balance. J. Am. Water Resour. Assoc. 35 (5), 1037–1051.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., van Harmel, R.D., Van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M.K., 2012. SWAT: model use, calibration, and validation. Trans. ASABE 55 (4), 1491–1508.

Baumgartner, A., Reichel, E., 1975. The World Water Balance. Elsevier, New York, 182 pp.

CIESIN, Center for International Earth Science Information Network, 2005. Columbia University; United Nations Food and Agriculture Programme (FAO); and Centro Internacional de Agricultura Tropical (CIAT) (2005), Gridded Population of the World: Future Estimates (GPWFE). Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at <http://sedac.ciesin.columbia.edu/gpw>.

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. J. Hydrol. 270 (1–2), 105–134.

Falkenmark, M., Rockström, J., 2006. The new blue and green water paradigm: Breaking new ground for water resources planning and management. J. Water Resour. Plan. Manage. 132 (3), 129–132.

Falkenmark, M, Widstrand, C., 1992. Population and water resources: a delicate balance. Population Bulletin, Population Reference Bureau.

Faramarzi, M., Abbaspour, K.C., Schulin, R., Yang, H., 2009. Modeling blue and green water availability in Iran. Hydrol. Proc. 23 (3), 486–501.
Faramarzi, M., Abbaspour, K.C., Vaghefi, S.A., Farzaneh, M.R., Zehnder, A.J.B., Yang, H., 2013. Modelling impacts of climate change on freshwater availability in Africa. J. Hydrol. 250, 1–14.
Fassio, A., Giupponi, C., Hiederer, R., Simota, C., 2005. A decision support tool for simulating the effects of alternative policies affecting water resources: an application at the European scale. J. Hydrol. 304, 462–476.
Fekete, B.M., Vörösmarty, C.J., Grabs, W., 1999. Global composite runoff fields of observed river discharge and simulated water balances, Report No. 22, Global Runoff Data Centre, Koblenz, Germany.
Gassman, P.W., Reyes, M., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future directions. Trans. ASABE 50 (4), 1211–1250.
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.
Gorgan, D., Bacu, V., Mihon, D., Rodila, D., Abbaspour, K.C., Rouholahnejad, E., 2012. Grid based calibration of SWAT hydrological models. Nat. Hazards Earth Syst. Sci. 12, 2411–2423.
Gosain, A.K., Rao, S., Basuray, D., 2006. Climate change impact assessment on hydrology of Indian river basins. Curr. Sci. 90 (3), 346–353.
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. Hydrol. Earth Syst. Sci. 17, 2393–2413.
Hargreaves, G.L., Hargreaves, G.H., Riley, J.P., 1985. Agricultural benefits for Senegal River Basin. J. Irrig. Drain. Eng. 111, 113–124.
Henriksen, H.J., Troldborg, L., Nyegaard, P., Sonnenborg, T.O., Refsgaard, J.C., Madsen, B., 2003. Methodology for construction, calibration and validation of a national hydrological model for Denmark. J. Hydrol. 280, 52–71.
Kløve, B., Ala-Aho, P., Bertrand, G., Gurdak, J.J., Kupfersberger, H., Kvœrner, J., Muotka, T., Mykrä, H., Preda, E., Rossi, P., Bertacchi Uvo, C., Velasco, E., Wachniew, P., Pulido-Velázquez, M., 2014. Climate change impacts on groundwater and dependent ecosystems. J. Hydrol. 518, 250–266.
Korzun, V.I., Sokolow, A.A., Budyko, M.I., Voskresensky, K.P., Kalinin, G.P., Konoplyanstev, A.A., Korotkevich, E.S., Kuzin, P.S., L'vovich, M.I. (Eds.), 1978. World Water Balance and Water Resources of the Earth. UNESCO, Paris.
Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. J. Adv. Geosci. 5, 89–97.
Lvovitch, M.I., 1973. The global water balance. Trans. Am. Geophys. Union 54, 28–42.
Mihon, D., Bacu, V., Rodila, D., Stefanut, T., Abbaspour, K., Rouholahnejad, E., Gorgan, D., 2013. Grid based hydrologic model calibration and execution. Adv. Intell. Syst. Comput. 187, 279–293.
Monfreda, C., Ramankutty, N., Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. J. Global Biogeochemical Cycles 22, GB1022. http://dx.doi.org/10.1029/2007GB002947.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2005. Soil and Water Assessment Tool. Theoretical Documentation. Version 2005. Springer, Berlin.
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., Musiake, K., 2001. Global assessment of current water resources using total runoff integrating pathways. Hydrol. Sci. J.: J. Sci. Hydrol. 46 (6), 983–995.
Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000 global monthly irrigated and rainfed crop areas around the year 2000: a new high resolution data set for agricultural and hydrological modeling. J. Global Biogeochem. Cycles 24, GB1011. http://dx.doi.org/10.1029/2008GB003435.
Revenga, C., Brunner, J., Henniger, N., Kassem, K., Payner, R., 2000. Pilot Analysis of Global Ecosystems, Freshwater Systems. World Resources Institute, Washington, DC.
Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. Water Resour. Res. 8, 1204–1213.
Rouholahnejad, E., Abbaspour, K.C., Vejdani, M., Srinivasan, R., Schulin, R., Lehmann, A., 2012. Parallelization framework for calibration of hydrological models. Environ. Modell. Software 31, 28–36.
Schuol, J., Abbaspour, K.C., Srinivasan, R., Yang, H., 2008. Modelling blue and green water availability in Africa at monthly intervals and subbasin level. Water Resour. Res. 44, W07406. http://dx.doi.org/10.1029/2007WR006609.
Shiklomanov, I.A., 2000. Appraisal and assessment of world water resources. Water Int. 25 (1), 11–32.
Stisen, S., Sonnenborg, T.O., Hojberg, A.L., Troldborg, L., Refsgaard, J.C., 2011. Evaluation of climate input biases and water balance issues using a coupled surface–subsurface model. Vadose Zone J. 10 (1), 37–53.
UN Report, 2012. Managing Water Under Uncertainty and Risk. The United Nations World Water Development Report 4, vol. 1. UNESCO Publishing.
van Beek, L.P.H., Wada, Y., Bierkens, M.F.P., 2011. Global monthly water stress: 1. Water balance and water availability. Water Resour. Res. 47, W07517. http://dx.doi.org/10.1029/2010WR009791.
Vörösmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., 2000. Global water resources: vulnerability from climate change and population growth. Science 289, 284–288.
Widén-Nilsson, E., Halldin, S., Xu, C.-Y., 2007. Global water-balance modelling with WASMOD-M: parameter estimation and regionalisation. J. Hydrol. 340, 105–118.
Yang, D.W., Musiake, K., 2003. A continental scale hydrological model using the distributed approach and its application to Asia. Hydrol. Process. 17 (14), 2855–2869.
Yang, H., Reichert, P., Abbaspour, K.C., Zehnder, A.J.B., 2003. A water resources threshold and its implications for food security. Environ. Sci. Technol. 37 (14), 3048–3054.
Yang, J., Abbaspour, K.C., Reichert, P., Yang, H., 2008. Comparing uncertainty analysis techniques for a SWAT application to Chaohe Basin in China. J. Hydrol. 358, 1–23.
Yates, D.N., 1997. Approaches to continental scale runoff for integrated assessment models. J. Hydrol. 201 (1–4), 289–310.
Zessner, M., Lindtner, S., 2005. Estimations of municipal point source pollution in the context of river basin management. J. Water Sci. Technol. 52 (9), 175–182.