Seventh Framework Programme
Theme 6
Environment

Project: 603864 – HELIX

Full project title:
High-**E**nd c**L**imate **I**mpacts and e**X**tremes

Deliverable: 7.3

Evaluation of regional impact models against observed trends or recent extreme events

Version 1.0

Original Due date of deliverable: 30/4/15

Actual date of submission: 30/4/15

Title	Evaluation of regional impact models against observed trends or recent extreme events
Creator	WP7
Editor	R Betts
Brief Description	Model evaluation is a crucial and indispensable part of the impact modelling chain. The present report summarizes the performance of a set of impact models against observed trends, baseline states and/or extreme events over Europe. LISFLOOD and JULES models were examined for their ability to realistically represent major flood and drought events, respectively, while LISCOAST is used to model the sea level and storm surge events. ORCHIDEE-crop is used to simulate wheat, maize and soybean yields.
Publisher	WP7
Contributors	Lorenzo Alfieri, Ioannis K. Tsanis, Lamprini Papadimitriou, Aristeidis Koutroulis, Manolis Grillakis, Luc Feyen, Michalis Vousdoukas, Evangelos Voukouvalas, Xuhui Wang, Richard Betts.
Format	Report
Creation date	9/2/2015
Version number	2.5
Version date	28/4/2015
Last modified by	R Betts
Rights	
Audience	☐ internal ☒ public ☐ restricted, access granted to: EU Commission
Action requested	Submit
Deadline for approval	

Evaluation of regional impact models against observed trends or recent extreme events

April 30

2015

Contents

1. OVERVIEW / SUMMARY

Projection of impacts under a different future climate poses a new modeling challenge. Model evaluation is a crucial and indispensable part of the impact modelling chain. This is especially true when investigating impacts beyond current climate conditions and even more when referring to high end – extreme impacts and by nature rare events, beyond the recent observed variability. Although model credibility is not always properly assessed (Refsgaard et al. 2013; Wagener et al. 2010) a number of studies examine the capability of simulating river flows (Falloon et al. 2011; Haddeland et al. 2011), historical floods (Gosling & Arnell 2011), droughts (Forzieri et al. 2014), agriculture (Iglesias et al. 2012) and sea level fluctuation (Grinsted et al. 2010).

During the present study we used observed/re-analysis data to drive regional versions of selected biophysical and techno-economic impacts models (LISFLOOD, JULES, LISCOAST) and compare against observed trends or recent extreme events. We also performed a preliminary analysis of the new model ORCHIDEE-crop using regional climate model (RCM) outputs to drive the model. This document is structured in four sections, in addition to this introduction.

Section 2 describes the main features of the regional setup of the LISFLOOD model employed to examine floods across Europe. Spatial settings and forcing variables are described in detail as well as model performance at pan-European scale, using a large number of gauging stations, based on specific efficiency criteria. Model credibility was also thoroughly tested against recent extreme events (floods in Central Europe in June 2013) described in the HELIX Deliverable D7.1.

Section 3 describes the ability of realistic reproduction of drought events using JULES land surface model. A short model description is followed by the description of the definition of drought used in this analysis. Low flows validation was performed at pan-European scale based on the ISLSCP II UNH-GRDC monthly runoff dataset for the 1986-1995 period. Two extreme events included in the drought atlases of major events of the 2nd half of the 20th century developed within WATCH (FP6) were also examined at regional and basin scale.

Section 4 includes information on the sea level modeling using LISCOAST coastal model. A brief description of the optimum numerical model setup is followed by the validation results over a 6.5 years period against observations from 110 stations of the JRC Sea Level Database, across Europe.

Section 5 present a preliminary analysis of simulations of wheat, maize and soybean by the new ORCHIDEE-crop model, driven by RCM simulations and compared with an existing dataset of yields for the period 1982-2004.

Parallel work in HELIX (D7.2) is using the validated impact models to assess impacts and implications of uncertainties in global patterns of climate change at 4°C using EURO-CORDEX RCM projections. Future work in HELIX (D7.4 and D7.5) will be based on the results of the impact models, aiming to an improved assessment of the impacts of climate change at 4°C based on new high-resolution RCM projection.

2. LISFLOOD HYDROLOGICAL MODEL

2.1 LISFLOOD

The hydrological simulation of large river basins forced by a variety of meteorological input and of climate projections have become feasible through the revolutionary increase in computing and network power over the past few decades and the advanced Geographical Information System (GIS) analysis tools that are now available. Traditionally, GIS was associated with static data such as topography, forest areas, urban areas or infrastructure, occasionally satellite images, often in a rather static way. The incorporation of dynamic data sets such as remotely sensed data (vegetation cover, snow cover), gridded meteorological data (precipitation, temperature, evapo-transpiration, either observed or forecast), weather station observations is becoming increasingly used also in a GIS environment to form the basis for effective hydrological monitoring on a variety of spatial and temporal scales.

The hydrological model hereby described is called Lisflood. It was developed at the Institute for Environment and Sustainability (IES) of the European Commission - Joint Research Centre (JRC) and its main features are described by van der Knijff et al. (2010). Lisflood is a hybrid between a conceptual and physically based rainfall-runoff model combined with a routing module in the river channel. It has been specifically designed for large river catchments (De Roo et al. 2001) but has also been applied to smaller watersheds (e.g., Alfieri et al. 2012; Younis et al. 2008). Recent works using different model versions include the comparison of various satellite precipitation input over Africa (Thiemig et al. 2013), a global flood forecasting and early warning system (Alfieri et al. 2013) and the simulation of climate scenarios over Europe on multi-decadal time scales for floods (Rojas et al. 2012; Rojas et al. 2013) and droughts (Forzieri et al. 2014) impact assessment under a changing climate.

Whenever possible parameters in Lisflood are linked to physical properties, e.g. soil or land use properties. Although for all model parameters default values are proposed, a number of calibration works have been performed on different model domains for better model performance (e.g., see Alfieri et al. 2014). Analysis of model parameter uncertainty and its impact on discharges simulated by the Lisflood model is presented in Feyen et al. (2007). The JRC stores all maps required to set up Lisflood at the Global and European scale, with grid resolution as fine as 0.1 degrees and 1 km, respectively. In the context of HELIX, in the WP7, hydrological simulations with Lisflood are foreseen to be run at 5 km grid resolution over Europe. This appears as a good tradeoff between computational feasibility and high enough level of detail to reproduce accurate streamflow statistics for a wide range of river basins. In addition, such model grid is of a similar magnitude to the EURO-CORDEX downscaled climate scenarios EUR-11, with approximately 12.5 km input resolution. Available static maps for Europe at 5 km resolution include maps related to topography (i.e., digital elevation model, local drain direction, slope gradient, elevation range), land use (i.e., land use classes, forest fraction, fraction of urban area), soil (i.e., soil texture classes, soil depth), and channel geometry (i.e., channel gradient, Manning's roughness, bankfull channel depth, channel length, bottom width and side slope). Physically-based input parameters and variables necessary for the model are estimated a priori from available data bases, such as HYPRES (Wösten et al. 1999), CORINE Land Cover (Batista e Silva et al. 2012), Soil Geographical Database of Europe (King et al. 1994), SRTM (Jarvis et al. 2008), CCM2 (Vogt & Foisneau 2007), among others.

Monthly averaged Leaf Area Index (LAI) maps are available for the entire Europe. They are estimated from satellite observations, through the Normalized Difference Vegetation Index (NDVI). Meteorological data needed as input from the Lisflood model are precipitation, average daily temperature, and daily potential evaporation rate (from free water surface, bare soil and reference crop). These maps are to be provided for each time step of simulation (i.e., daily or sub daily), although for temperature and evaporation, daily or longer term averages are required by the model.

The latest Lisflood model developments for the European setup include the simulation of 182 lakes and 34 large reservoirs, and the implementation of monthly maps of water use from the SCENES project (Kamari et al. 2008), which are assumed constant throughout the current century.

2.2 LISVAP

When evapo-transpiration maps are not directly available for use in Lisflood, they can be estimated from other weather parameters though different literature formulations. To this aim, a software named Lisvap (Burek et al. 2013) was developed at the JRC-IES. Lisvap is a pre-processor that calculates potential evapo(transpi)ration from gridded meteorological observations, that can be used as input to Lisflood. The approach is based on the Penman-Monteith equation or the Hargreaves equation. The calculation of potential evapo(transpi)ration is complicated somewhat by the fact that the different datasets that are usually available are rather heterogeneous. For instance, incoming solar radiation can be estimated from sunshine duration or cloud cover data. Some data suppliers do not offer this kind of information, but provide pre-calculated grids of components of the radiation balance instead. Wind speed is sometimes provided in the form of U and V components. Vapour pressure is sometimes used in place of dew point temperature. Lisvap has been designed to handle this heterogeneity in a straightforward way. It contains various options that allow the user to select the data to use, and combinations of different data sources (e.g. vapour pressure and dew point temperature) can be combined within one Lisvap run. Just like Lisflood, Lisvap is implemented in the PCRaster Environmental Modelling language (Wesseling et al. 1996), wrapped in a Python based interface. It runs on any operating system for which Python and PCRaster are available. Currently these include 32-bits Windows (e.g. Windows XP, Vista) and a number of Linux distributions.

2.3 MODEL PERFORMANCE AT CALIBRATED STREAMFLOW GAUGES

The validation of Lisflood on the streamflow regimes in different regions of Europe, including past observed events, is performed by driving the hydrological model with an observed meteorological dataset and then comparing the resulting discharge with observed time series at a number of stations. To this end, the meteorological input data needed to run the hydrological model must cover the domain of interests (i.e., typically the whole drainage area upstream all the selected streamflow gauges), possibly with no gaps in space and time. Also, if high-flow events are of interest, as in HELIX, the space and time resolution of the input data and of the hydrological model must be fine enough to enable the model to reproduce adequately the statistics and the magnitude of the observed runoff. For Lisflood simulations over Europe, we have opted for a model setup on a 5 km regular grid and 1-day time step, which was shown to be capable of simulating high flow conditions in river basins with upstream area larger than 4000 km^2 (Bartholmes et al. 2006).

A semi-automated calibration procedure was developed to perform global optimization for numerous European catchments where historical discharge measurements were available at selected gauges along their river network. The approach used is based on the hydroPSO R package (Zambrano-Bigiarini & Rojas 2014). The procedure implements a state-of-the art version of the Standard Particle Swarm Optimization (Kennedy & Eberhart 1995). This algorithm was selected for calibrating 9 model parameters, due to its flexibility, easy implementation (programming), low memory and computational requirements, low number of adjustable parameters, and efficiency.

The main improvements that occurred since the previous (i.e., November 2011) Lisflood calibration for Europe include: model development (e.g., inclusion of the water use component, reservoir component), improvement of the meteorological data (better spatio-temporal coverage, better data quality checks), extensive corrections of river network for the European domain, increase of the number of calibrated stations to a total of 693, longer time series of discharge observations which allowed an average of 7.5 years of calibration periods, and calibration of reservoir parameters for catchments with simulated reservoirs. For the calibration exercise over the European domain, Lisflood was forced with temperature, precipitation and potential evapotranspiration raster maps extracted from the EFAS-Meteo dataset (Ntegeka et al. 2013), also described in the HELIX deliverable D7.1.

In order to validate the performance of the calibrated model, simulations were compared against observations for a selected validation period of up to 10 years of daily values. For the majority of the calibrated stations the Nash–Sutcliffe Efficiency (NSE) was skillful (i.e., above zero), while for approximately 60% and 30% of stations NSE for validation period exceeded 0.5 and 0.75 respectively. The spatial pattern of the NSE indicated that the poorest performance occurred for areas with anthropogenic impacts on river regimes (e.g., reservoirs, important water withdrawals etc), especially in the Iberian Peninsula (see Figure 1). An example score card of the parameter calibration for the Thames River at Kingston on Thames (UK) is shown in Figure 2.

The calibrated model set up was also used to generate a continuous discharge climatology for the period 1991 – 2014 for the entire European domain.

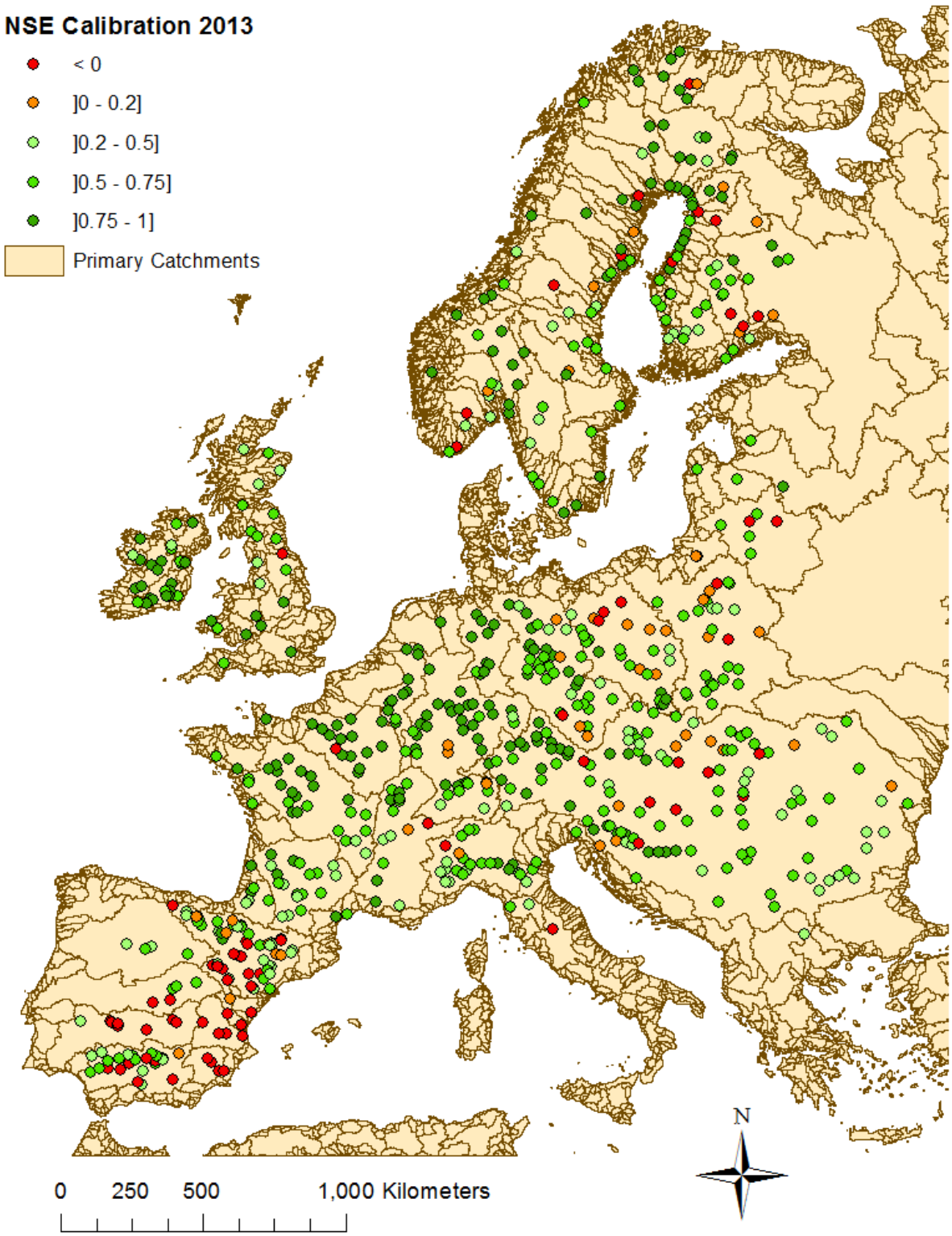


Figure 1: Spatial distribution of the NSE for the calibrated Lisflood model at 693 river stations in Europe. Boundaries of river basins are also shown.

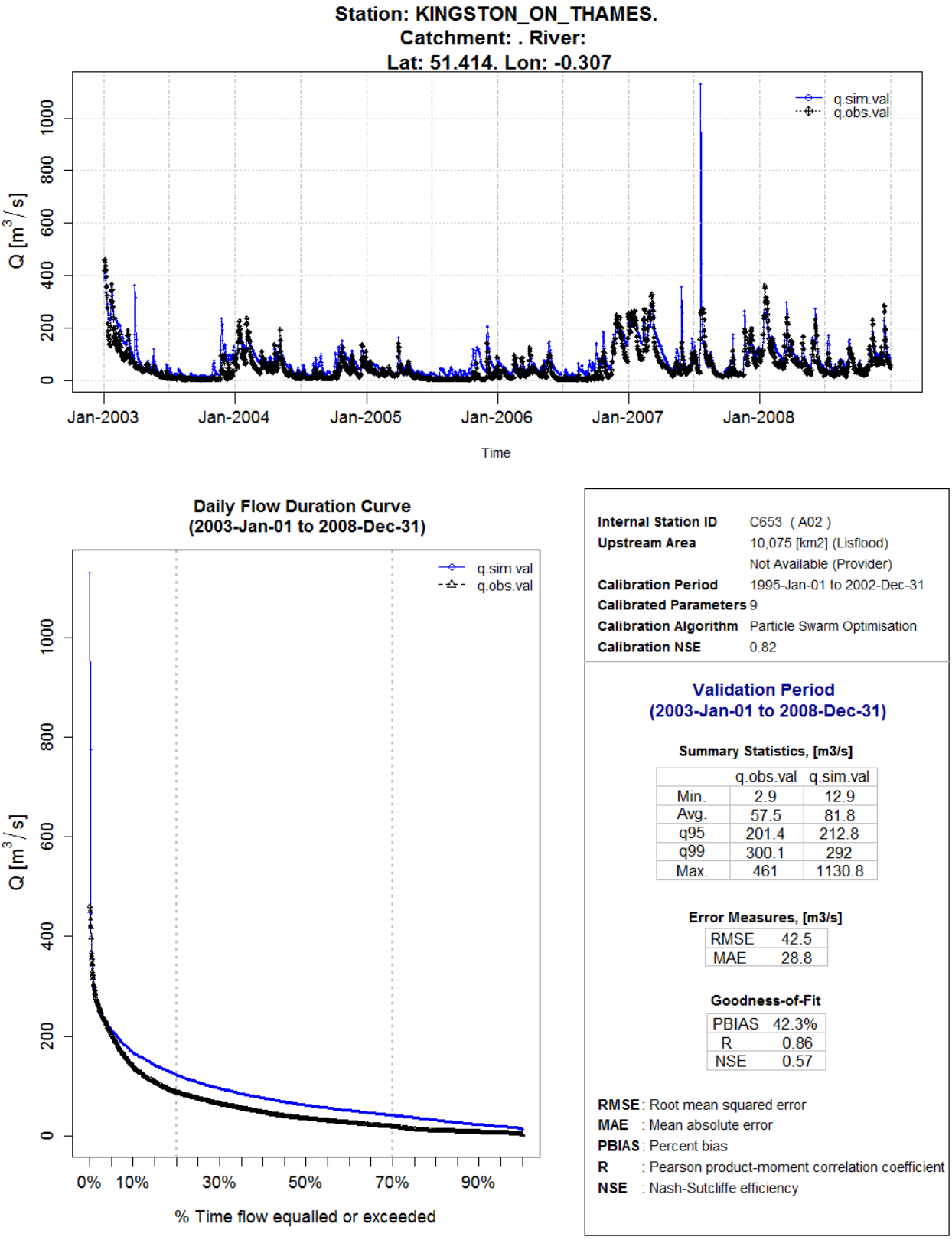


Figure 2: Score card of Lisflood calibration for the Thames River at Kingston on Thames (UK).

2.4 MODEL PERFORMANCE FOR RECENT FLOOD EVENTS

The following sub-sections illustrate the performance of the Lisflood hydrological model to reproduce some recent extreme events described in the HELIX Deliverable D7.1. To this aim, Lisflood

was forced by the EFAS-Meteo dataset over the area affected by the hydro-meteorological event and simulated discharge for the event was compared for some stations where daily observations were available.

2.4.1. Floods in Central Europe in June 2013

Lisflood simulations driven by the EFAS-Meteo dataset as input were performed on the European domain and simulated daily discharge was compared with observations in 28 river stations where streamflow data was made available (see Figure 3). Note that some recorded discharge time series include missing values, particularly in the range of extreme values, often due to power cuts following the severe weather and the damaging of the measuring devices caused by the rage of the flood flow. The comparison between simulated and observed discharges is performed for the May-June 2013 time span and is shown in Figure 4. Six relevant skill scores are calculated for the comparison shown in Figure 4 and are plotted in Figure 5 as a function of the upstream area of each river gauge. These include (from top-left in Figure 5) the Root Mean Square Error (RMSE) [m^3/s], the Normalized Root Mean Square Error (NRMSE), the Percent Bias (PBIAS), the Nash-Sutcliffe Efficiency (NSE), the Pearson Correlation coefficient (r), and the Coefficient of Determination (R2).

Figure 3: Location of the river gauges where observed daily discharge was provided for May-June 2013.

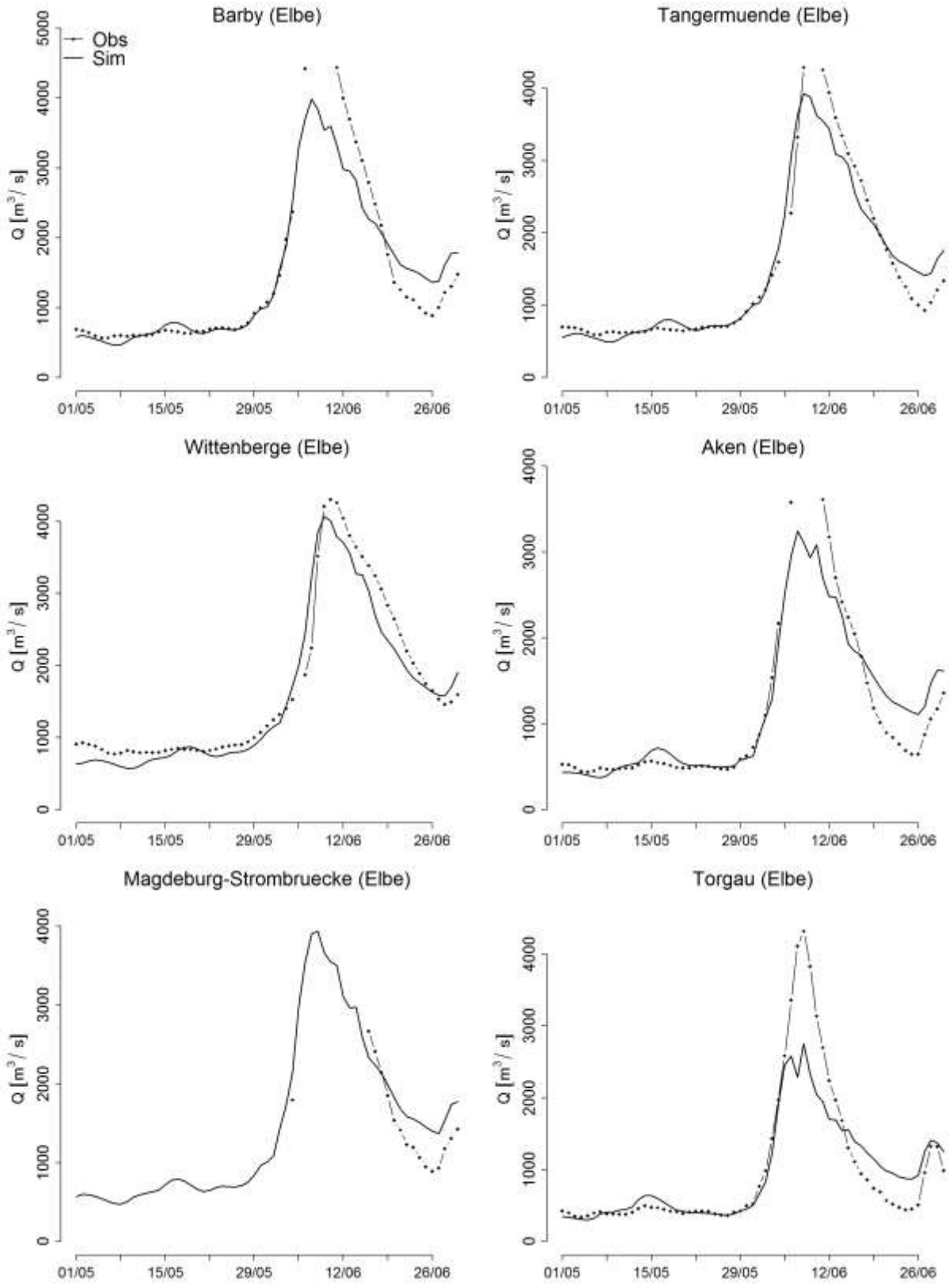


Figure 4: Observed vs. simulated discharges for river gauges shown in Figure 3.

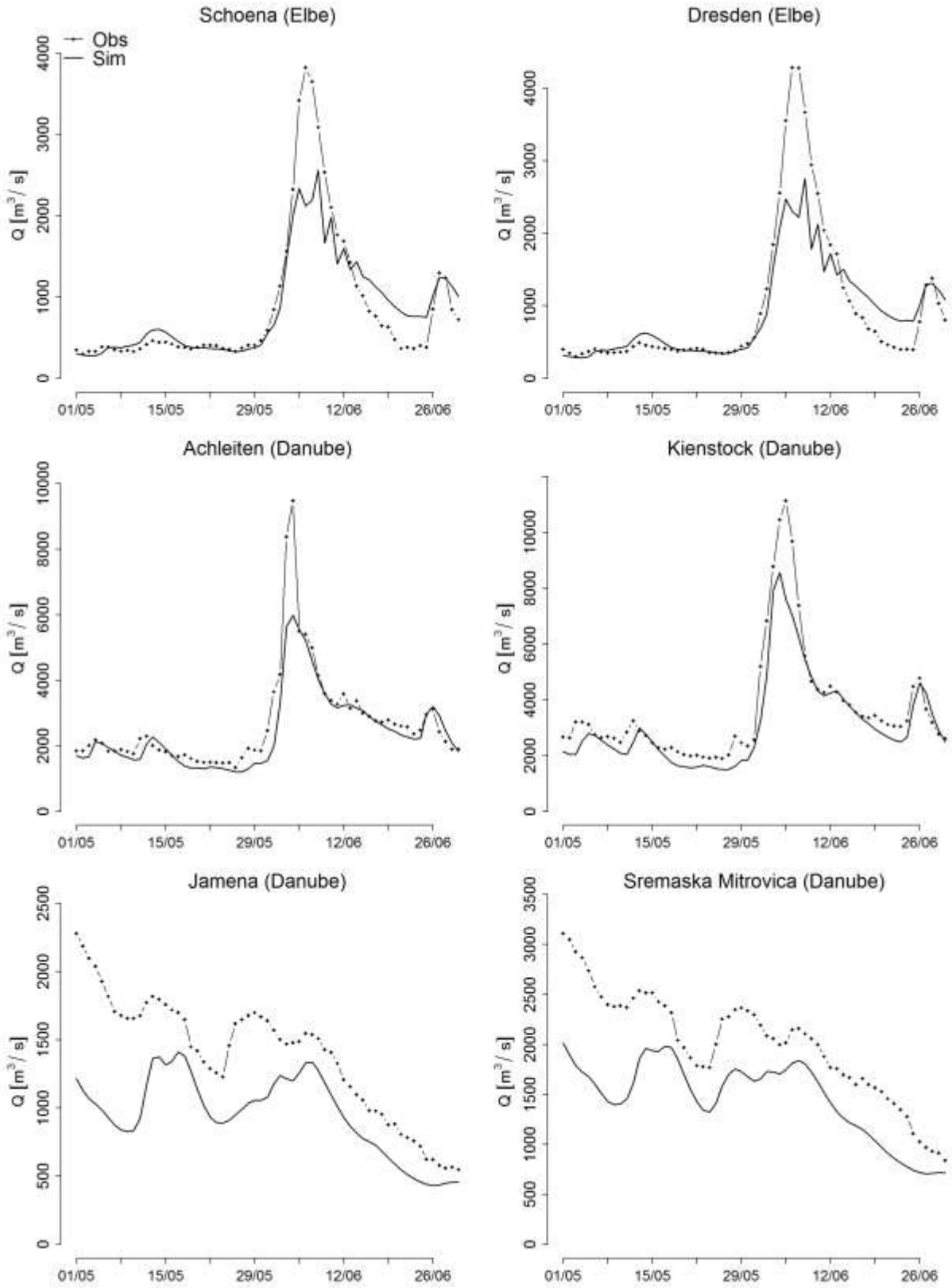


Figure 4 (continued)

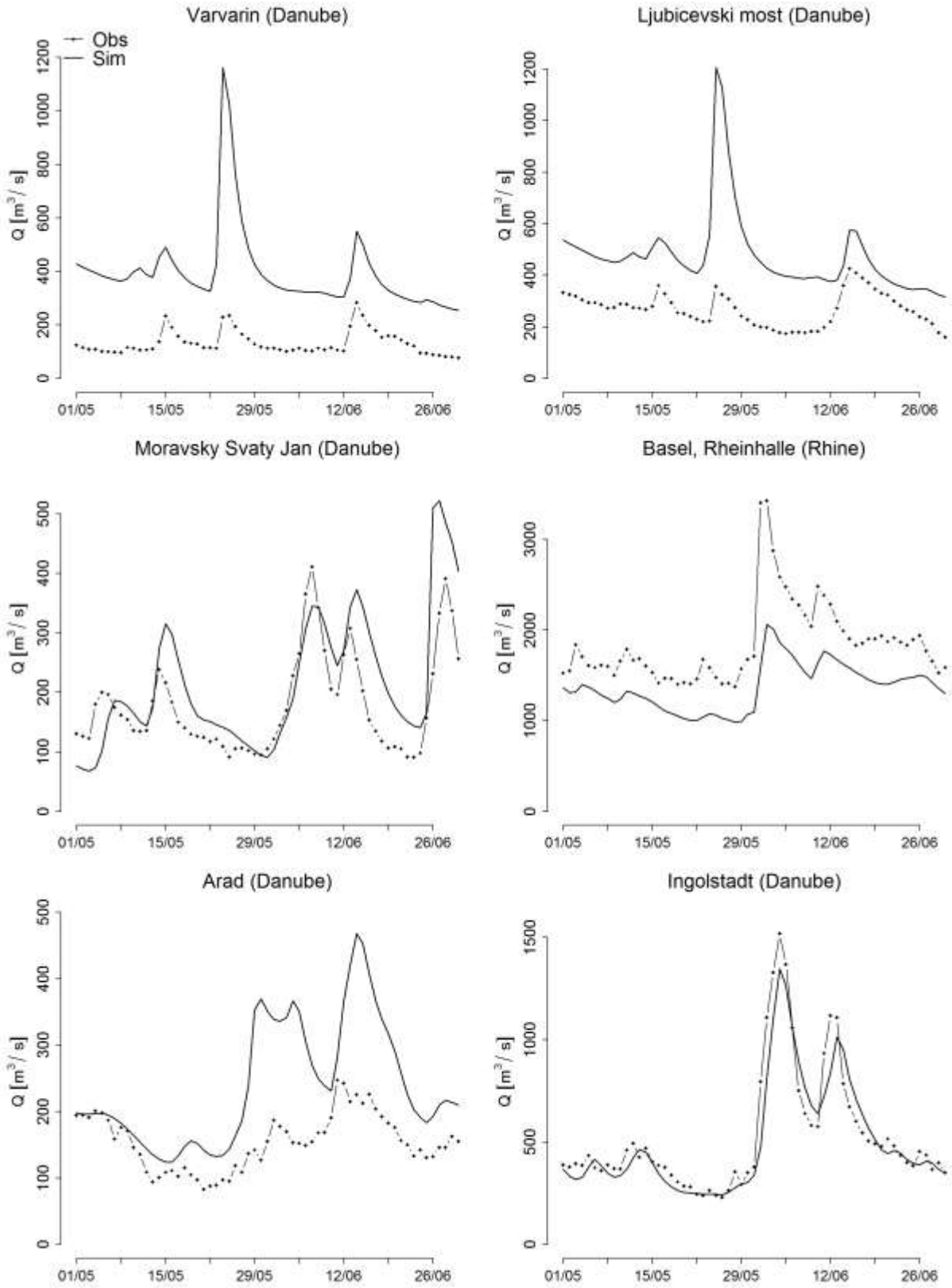


Figure 4 (continued)

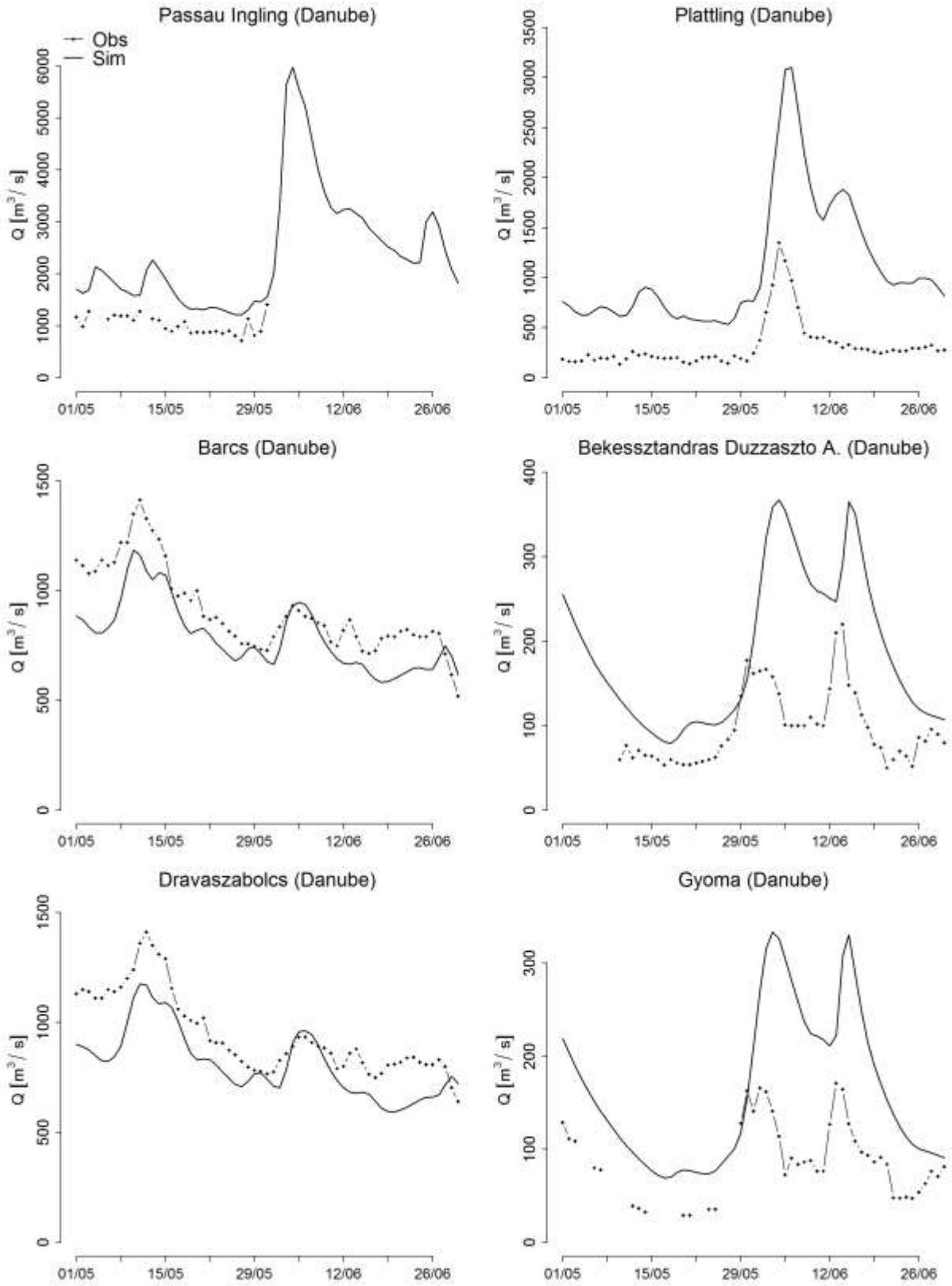


Figure 4 (continued)

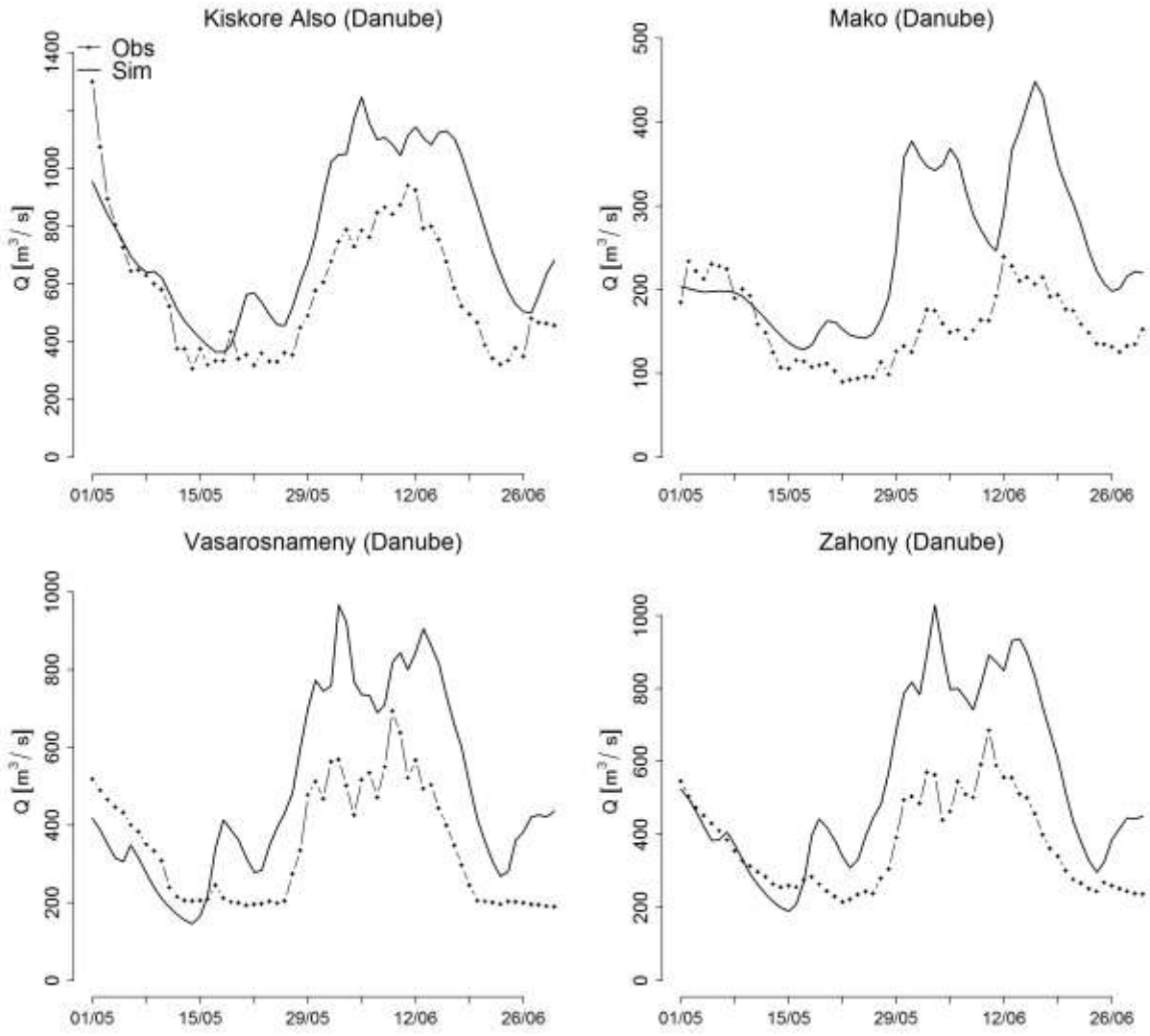


Figure 4 (continued)

Figure 5 shows a general improvement of the simulation performance for increasing upstream area, both in terms of bias and of correlation. Note that no data assimilation of discharge measurements is included in the hydrological model, as initial conditions are always estimated by updating the water balance in the river network, using the model states of the previous day and the maps of observed meteorological variables. In a number of cases the simulated discharge is in good agreement with observations for low flow conditions, while large negative bias occurs for peak discharge values. This points out the limitations given by the space and time resolution of the hydrological model and of the meteorological input data, which limits the representation of extreme discharge peaks, particularly when the event dynamics have peculiar features at finer scales than those considered in the modeling. Indeed, such an issue is less evident on river points further downstream the river network, where the simulation of discharges on average improves.

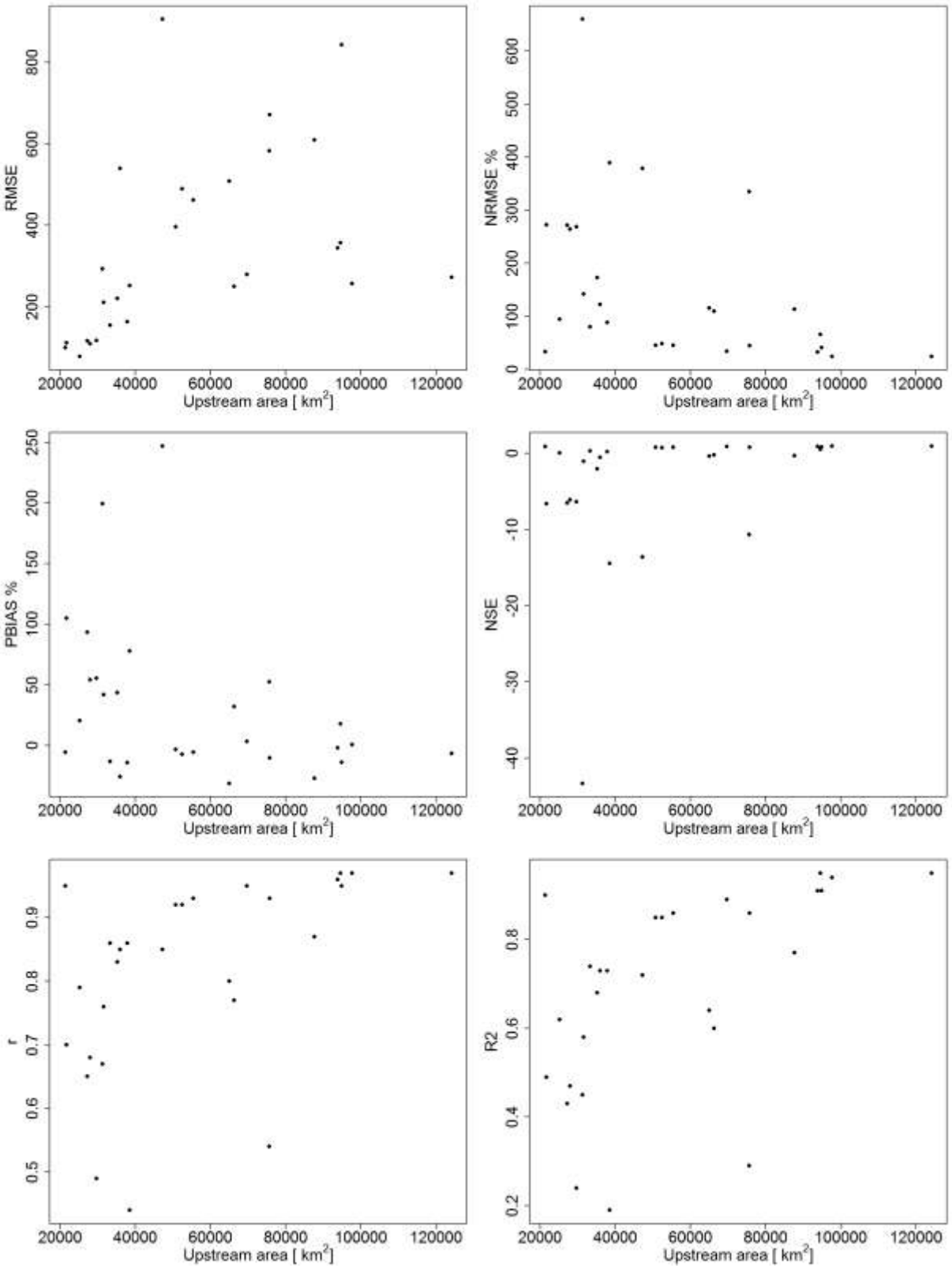


Figure 5: Skill scores of the discharge comparison shown in Figure 4.

3. JULES LAND SURFACE MODEL

3.1 JULES

JULES (Joint UK Land Environment Simulator) is a physically based community land surface model that was established in 2006. It is comprised of two parts: the Met Office Surface Exchange Scheme (MOSES; Cox et al. 1998) and the Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID; Cox 2001) component. MOSES is an energy and water balance model which is JULES' forerunner, and TRIFFID is a dynamic global vegetation model (Cox 2001; Best et al. 2011; Clark et al. 2011). In our model application for this study we do not examine vegetation dynamics thus we are focusing on the MOSES component of JULES.

The meteorological forcing data required for running JULES are: downward shortwave and longwave radiation, precipitation rate, air temperature, wind-speed, air pressure and specific humidity (Best et al. 2011).

JULES has a modular structure, which makes it a flexible modelling platform, as there is the potential of replacing modules or introducing new modules within the model. The physics modules that comprise JULES include the following themes: surface exchange of energy fluxes, snow cover, surface hydrology, soil moisture and temperature, plant physiology, soil carbon and vegetation dynamics (Best et al. 2011), with the latter being disabled for this application.

In JULES, each gridbox is represented with a number of surface types, each one represented by a tile. JULES recognises nine surface types (Best et al., 2011), of which five are vegetation surface types (broadleaf trees, needleleaf trees, C3 (temperate) grasses, C4 (tropical) grasses and shrubs) and four are non-vegetated surface types (urban, inland water, bare soil and ice). A full energy balance equation including constituents of radiation, sensible heat, latent heat, canopy heat and ground surface heat fluxes is calculated separately for each tile and the average energy balance for the gridbox is found by weighting the values from each tile (Pryor et al. 2012).

In JULES the default soil configuration consists of four soil layers of thicknesses 0.1 m, 0.25 m, 0.65 m and 2.0 m. This configuration however can be altered by the user. The fluxes of soil moisture between each soil layer are described by Darcy's law and a form of Richards' equation (Richards 1931) governs the soil hydrology. Runoff production is governed by two processes: infiltration excess surface runoff and drainage through the bottom of the soil column, a process calculated as a Darcian flux assuming zero gradient of matric potential (Best et al. 2011). There is also the option of representing soil moisture heterogeneity. In that case total surface runoff also includes saturation excess runoff. The model allows for two approaches to introduce sub-grid scale heterogeneity into the soil moisture: 1) use of TOPMODEL (Beven & Kirkby 1979), where heterogeneity is taken into account throughout the soil column, or 2) use of PDM (Moore 1985), which represents heterogeneity in the top soil layer only (Best et al. 2011). Calculation of potential evaporation follows the Penman-Monteith approach (Penman 1948). Water held at the plant canopy evaporates at the potential rate while restrictions of canopy resistance and soil moisture are applied for the simulation of evaporation from soil and plant transpiration from potential evaporation.

JULES simulates fluxes at the vertical direction only. For hydrological applications this means that the model calculates runoff production in each gridbox which needs to be routed to estimate

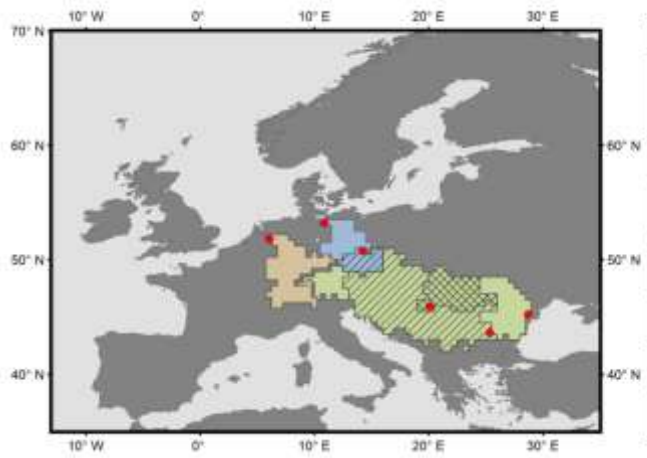
streamflow. The standard version of the JULES model until very recently (February 2015) did not account for a routing mechanism. To overcome this model limitation, we use a conceptual lumped routing approach based on triangular filtering in order to delay runoff response. This is applied after discriminating the gridboxes that contribute to runoff production of a specific basin from the gridded model output. Determination of gridboxes upstream of the gauging station location is implemented using the TRIP river routing scheme (Oki & Sud 1998).

3.2 DROUGHT EVENT IDENTIFICATION

For the task of historical validation the model was forced with a daily meteorological dataset combining the WATCH and WFDEI (Weedon et al. 2014) datasets. Model output covers the time period from 1971 to 2010 but in order to avoid the discontinuity of the transition from one dataset to the other existing in the combined data, only results from (and including) 1981 where used in this study. This analysis includes results up to 2005, in order to be consistent with the end of the historical period of the Euro-Cordex data and HELIX Deliverable D7.2.

Drought conditions have been studied for the whole European continent and for specific catchments and gauging stations in Europe. To evaluate the model at the European scale a gridded dataset of composite monthly runoff (Fekete & Vorosmarty 2011) spanning from 1986 to 1995 was used. Catchment level model evaluation was performed based on GRDC discharge data from the selected gauging stations. Selection of basins and stations was subjected to two kinds of limitations. Firstly, a criterion for selecting gauging stations was the availability of daily data for the study period, something that left no available station at the western Mediterranean region. Secondly, due to model resolution only the larger basins could be examined giving meaningful results in terms of the simulated area. The Europe study domain along with information on the catchments tested and their corresponding gauging stations are shown in Figure 6.

In our approach of modelling hydrological drought conditions, we establish a drought threshold which corresponds to the lower 10^{th} percentile of historical discharge. The threshold is computed separately for observed and modelled data. At the European scale the drought threshold is calculated for each gridbox, taking into account all the values of the 10 year runoff record. To examine the variability in the intensity of drought conditions, the lower 20^{th} and 30^{th} percentiles are also considered. At the basin level, daily varying drought thresholds are established as follows: for each Julian day k the threshold is defined as the 10^{th} percentile of a 31-day window discharge centering at day k, applied to all years of the study period (1981-2005).



Station Info						
River	Station	GRDC Station No	Latitude (dec. degrees)	Longitude (dec. degrees)	Observed cathment area [km²]	Modelled cathment area [km²]
Danube	Senta	6544100	45.93	20.08	140130	140779
Danube	Zimnicea	6742500	43.63	25.35	658400	604393
Danube	Ceatal Izmail	6742900	45.22	28.72	807000	815063
Rhine	Rees	6335020	51.75	6.40	159300	177339
Elbe	Decin	6140400	50.79	14.23	51123	53799
Elbe	Neu-Darchau	6340110	53.23	10.89	131950	133886

Figure 6: European study domain, tested basins as defined by the model's 0.5 degree resolution, gauging stations and general information on the stations.

3.3 MODEL PERFORMANCE FOR THE HISTORICAL VALIDATION PERIOD

Hydrological drought is mainly described as a period of persistent low flows compared to average state flows. For this reason model validation is orientated in evaluating how well JULES can capture the lower percentiles of runoff and discharge. Figure 7 shows the spatial variability of the 10th percentile of observed and modelled monthly runoff from 1986 to 1995 over the European domain. The comparison reveals that observed and modelled runoff are in good agreement in terms of their spatial pattern and range of values.

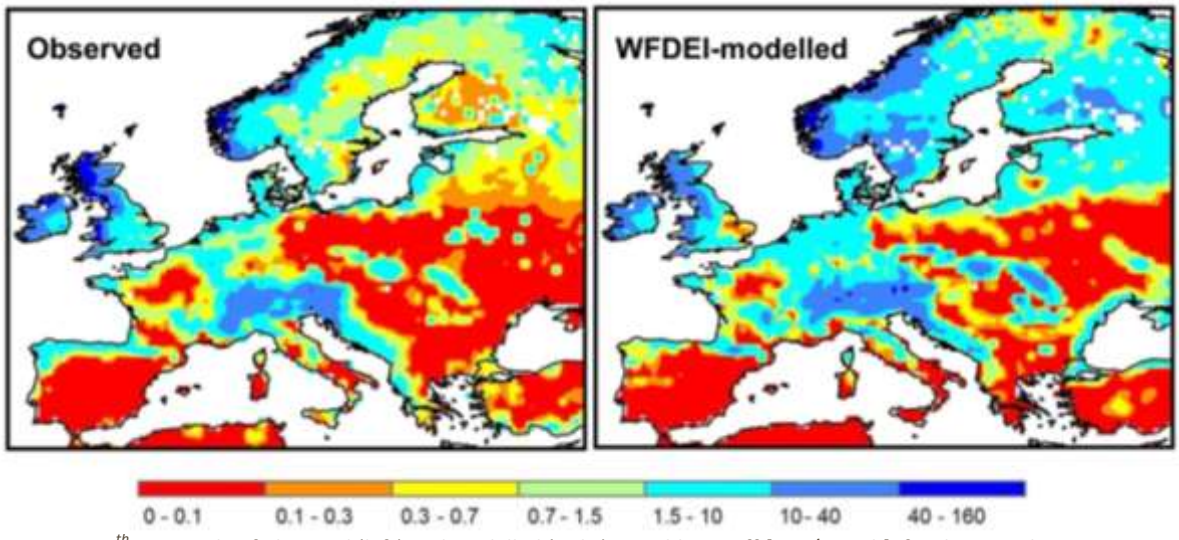


Figure 7: 10th percentile of observed (left) and modelled (right) monthly runoff [mm/month], for the period 1986 to 1995.

Moving to model evaluation at the catchment level, it is important to assess how well historical low flow conditions can be captured by the model. Figure 8 illustrates a comparison of observed and modelled low flow conditions, examined as number of days per year under the daily varying drought threshold. With the exception of one station (station Senta of river Danube), number of days under drought threshold from observed and modelled discharge values are well correlated (R^2>0.71 for all

other stations). Looking at the number of days under drought threshold for each year, although deviations are present, a relatively good match between observed and modelled values can be observed for most years. A constant sign in model bias cannot be identified, meaning that the model can either under- or overestimate the observed days under drought threshold.

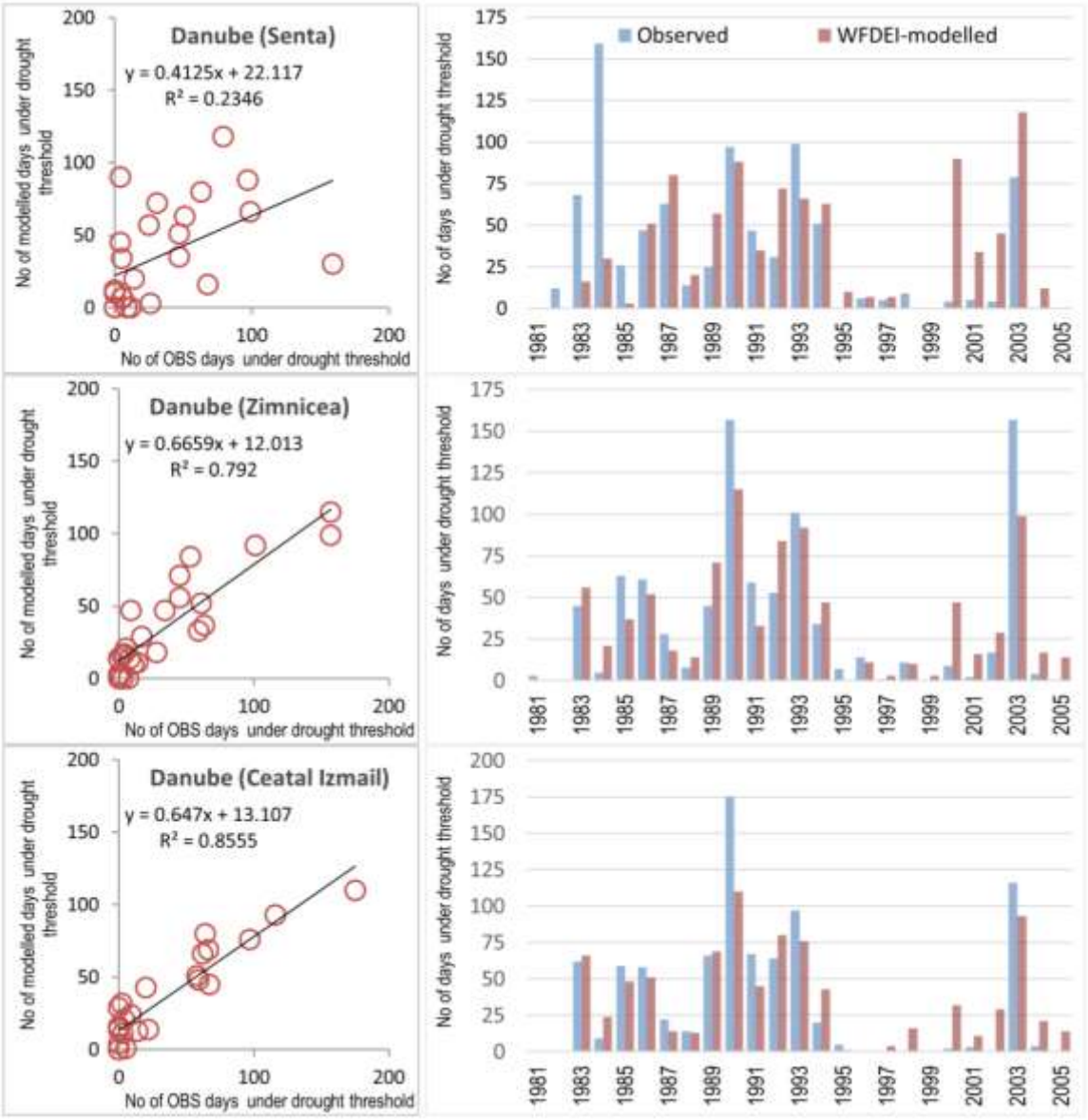


Figure 8: Comparison between observed and modelled drought climatology for each gauging station Correlation between the number of days under drought threshold as computed by observed and modelled values (left) and number of days under drought threshold per year (right).

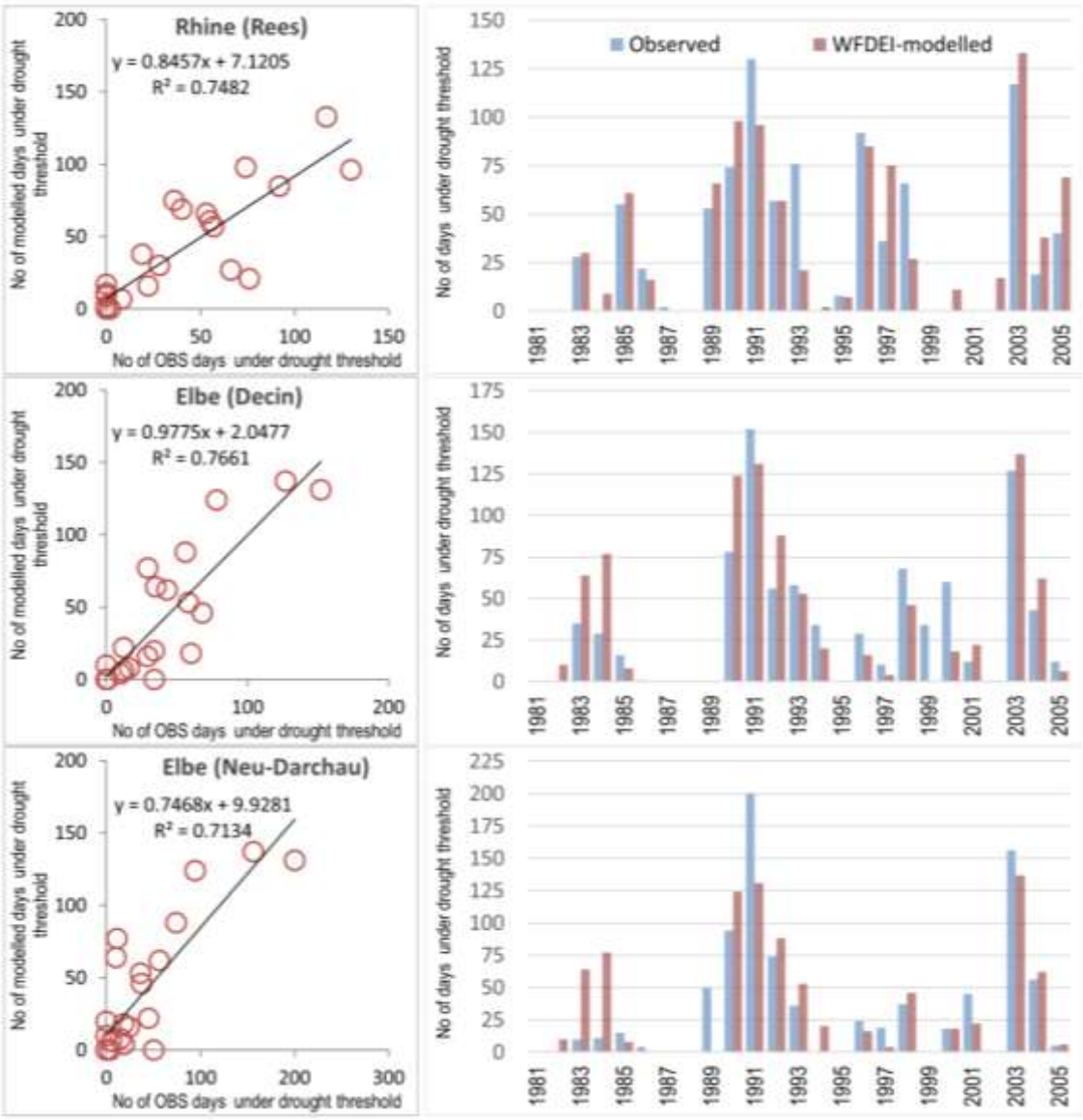


Figure 8 (continued)

3.4 MODEL PERFORMANCE FOR SELECTED PAST DROUGHT EVENTS

In this section JULES' performance in capturing two past drought events described in the HELIX Deliverable D7.1 is assessed. The first event is described as a two year drought in D7.1 (1989-1990 drought event) but here we focus on the second year (1990) as this is when hydrological drought conditions were established due to the meteorological drought of the previous year. Model performance is presented at both European and basin scale.

2.4.2. Pan-European scale

In Figure 9, the evolution in the spatial variability and intensity of the two past drought events of 1990 and 2003 is depicted. Intensity is expressed by setting three drought thresholds: the 10th, the 20th and the 30th percentile of historical monthly runoff. For 1990, observations were available thus this year's event is described by both observed and modelled data. The 2003 event is only described by the model due to the lack of a gridded observational dataset covering this period.

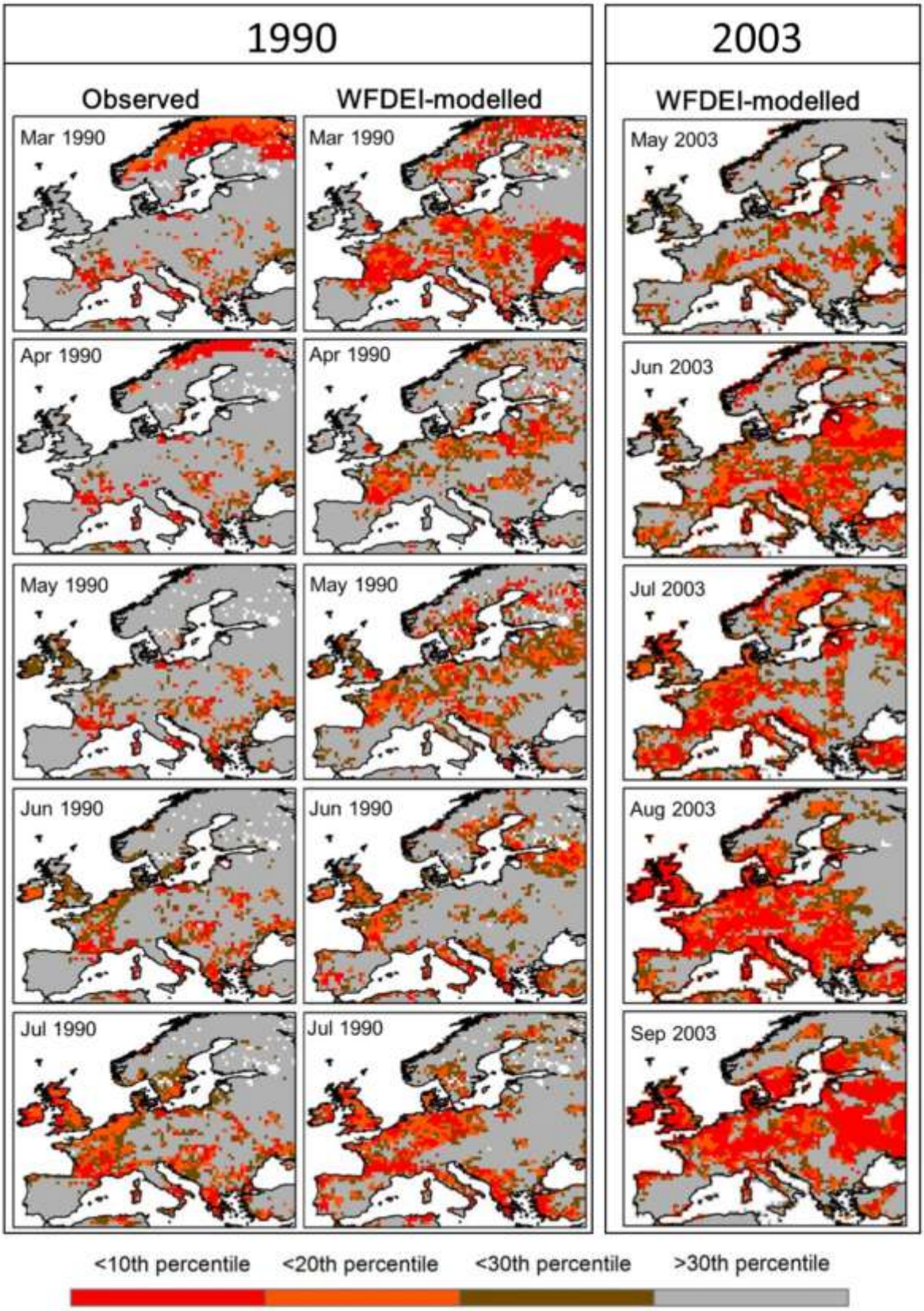


Figure 9: Evolution of drought events over the European domain, based on runoff values falling below different historically established low flow percentiles (10th, 20th and 30th percentile).

2.4.3. Basin scale

The manifestation of the two past drought events at specific basins, was assessed by applying the drought level threshold method to compare the daily discharge in the event year with the average-state historical discharge. More specifically, drought duration is assessed by the cumulative distribution of total number of days under the 10^{th} percentile drought threshold. This is derived as follows: first, daily discharge values are compared with the daily varying drought threshold. Then, binary time-series of deficit indices are produced by assigning the number 1 if the discharge of each particular day is lower than the drought threshold and the number is if it exceeds the threshold. The cumulative sum of the deficit indices time-series produces the cumulative distribution of days under drought threshold. The results of this analysis for the drought events of 1990 and 2003 are shown in Figures 10 and 11 respectively. Assessment of drought intensity is performed by accumulating the water deficit that occurred in each basin while under drought conditions. Water deficit is computed as the difference between the drought threshold discharge and the actual discharge for the days under threshold. Cumulative water deficit results for the events of 1990 and 2003 are illustrated in Figures 11 and 13 respectively.

Looking at the cumulative distributions of number of days under drought threshold for 1990 (Figure 10) it can be observed that JULES has similar distributions with the observed data for all stations, with the exception of Elbe at Decin where observations describe a plateau in drought evolution from mid-April to July 1990 which is not captured by the model. Other general observations are that JULES' modelled distribution starts to evolve earlier than the observed, and finishes with less in total days under drought threshold compared to observations for the three stations in Danube but with more days for the three other stations of Rhine and Elbe. According to the results shown in Figure 11, JULES' simulated a smaller water deficit than the observed on for Danube, with the difference between the days under threshold and water deficit between model and observations getting more pronounced as the corresponding drainage area of the basin increases. For Rhine and Elbe at Decin JULES overestimates the total water deficit while for Elbe at Neu-Darchau model and observations agree at a deficit of about 2 mm/basin area.

The temporal evolution of the 2003 drought event is described well by JULES for Danube at Senta and Ceatal Izmail, Rhine and Elbe (Figure 12), although at Danube-Senta the observed end of drought period at October 2003 is not captured by the model. For Danube at Zimnicea JULES does not capture a large drought days spike between August and October 2003, finishing with significantly less days under threshold than observed. For Rhine and Elbe at Decin JULES and observed data have a very good agreement in terms of both the shape of the distribution and the total number of days under threshold. The total water deficit during 2003 (Figure 13) is slightly overestimated by JULES for Danube at Ceatal Izmail and Elbe at Decin, significantly overestimated for Danube at Senta and Rhine and underestimated for Danube at Zimnicea and Elbe at Neu-Darchau.

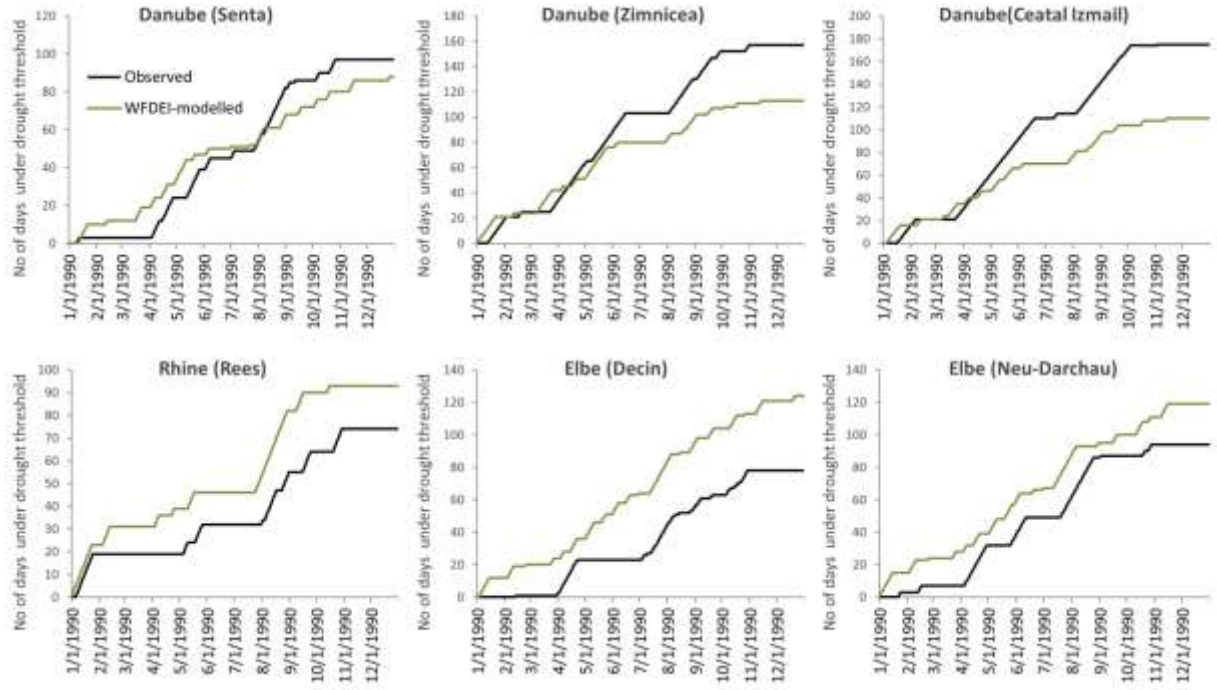


Figure 10: Cumulative distribution of total number of days under drought threshold for the year 1990.

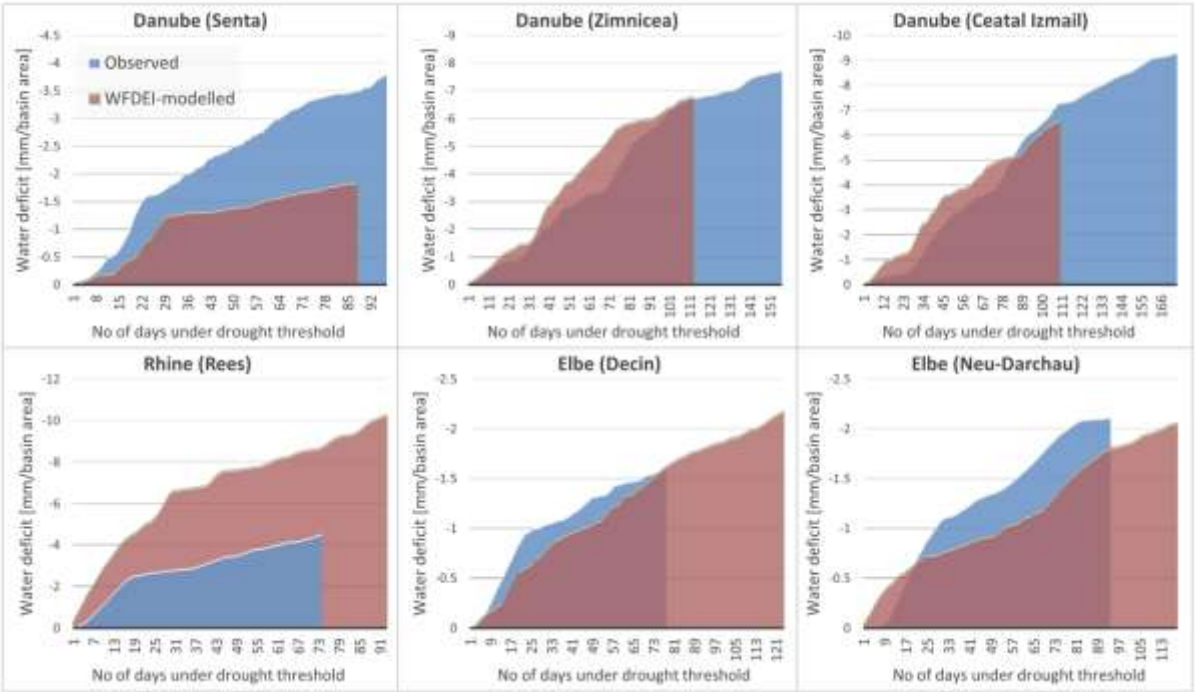


Figure 11: Cumulative sum of water deficit established during the days under the drought threshold for the year 1990.

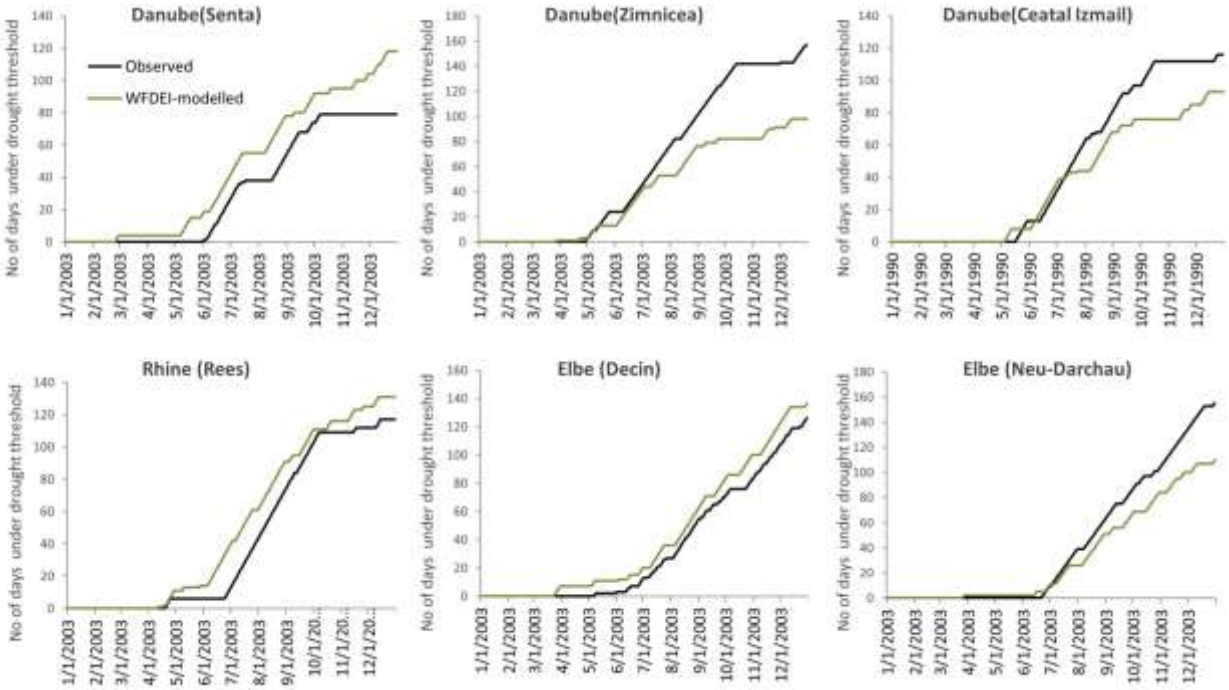


Figure 12: Cumulative distribution of total number of days under drought threshold for the year 2003.

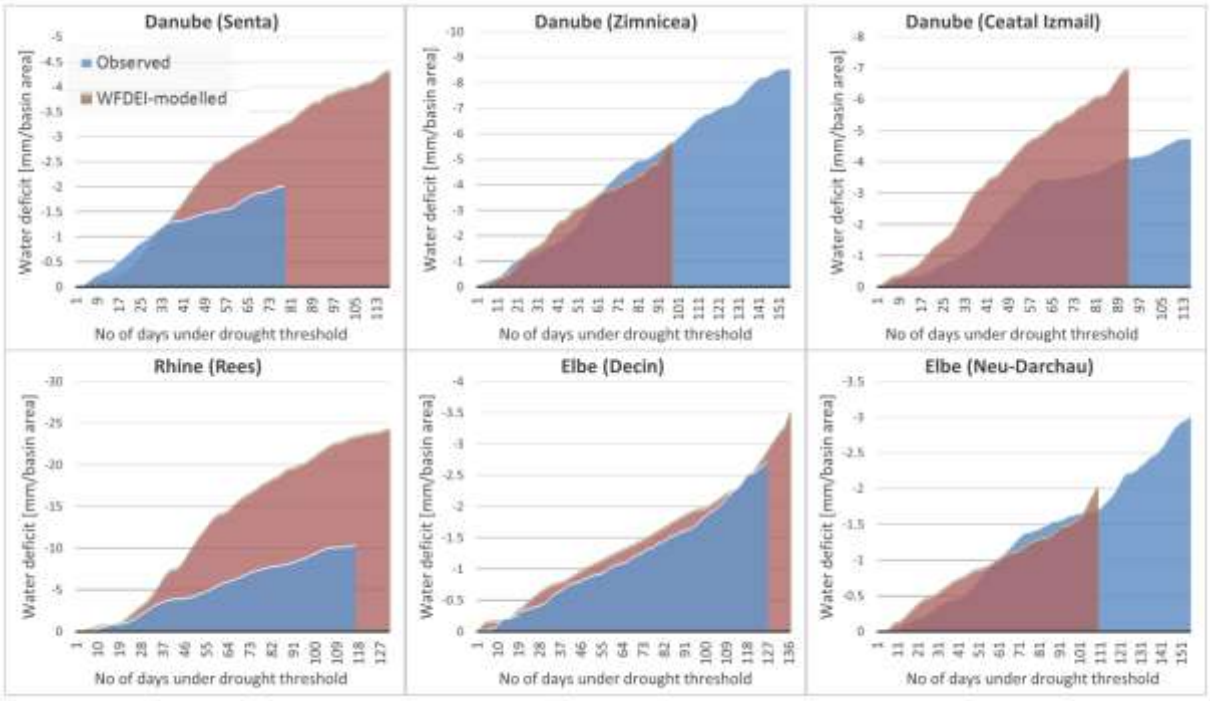


Figure 13: Cumulative sum of water deficit established during the days under the drought threshold for the year 2003.

4. LISCOAST COASTAL MODEL

4.1 NUMERICAL MODEL SETUP

Storm surges, also referred to as *meteorological residuals* or *meteorological tide*, constitute along with the waves and the tidal oscillations, the main components of extreme water levels along the coastal zone (Losada et al. 2013; Lowe et al. 2010). Storm surges are driven by wind driven water circulation towards or away from the coast and, to a smaller extent, by atmospheric pressure driven changes of the water level; i.e. the inverse barometric effect (Horsburgh & De Vries 2011). The magnitude of the storm surge depends on a number of factors comprising the size, movement, and intensity of the storm system, the nearshore local bathymetry (water depth) or the shape of the coastline.

Water level residuals η_{surge} were estimated using the open-source model Delft3D, with simulations being forced by wind velocity and atmospheric pressure fields, and without considering the tidal effects (Table 1). Modelling efforts were organized in runs carried out in batch mode, with each run simulating 5 years of water level fluctuations, after 90 days of model 'warm up' time, while water level was initially zero along the entire domain. Water level model output was obtained every three hours, and every 25 km along the coastline.

Prior to the main core of simulations, several efforts took place to optimize and validate the model, testing different model setups, varying the grid extent/resolution and applying nesting or domain decomposition. The numerical grid setup that was finally selected is a regular grid of 0.2° resolution which included Europe and the largest extent of the N. Atlantic (spanning from 40° W to 47° E and from 26° N to 73° N; Figure 14), as it was proved to be the best compromise in terms of data quality, model stability and computational times. For better analysis of the model performance and the storm surge scenarios, the European coastal zone was divided into 10 regions on the grounds of the geographical and physical setting: Black Sea, East, Central and West Mediterranean, South- and North-North Atlantic, Bay of Biscay, as well as North, Baltic and Norwegian Sea (Figure 14).

Table 1. Information about the model setup and the simulations

Model setup	**Information**
Storm surge model used	Delft3D version 5.01.00.4018
Processes simulated	Wind/pressure-driven ocean circulation
Grid	Regular, 0.2° (40° W-47° E; 26° N-73° N)
Atmospheric forcing	ERA-INTERIM (validation), CMIP5 (scenarios)
Period simulated	01/01/2008 to 01/06/2014
Model output	Water level every 3 hr and 25 km of coast

4.2 MODEL VALIDATION

Before starting the climate scenario simulations a validation case was run, spanning from 01/01/2008 to 01/06/2014, since this period was characterized by good availability of ground-truth data and increased marine storm activity including high impact events (Vousdoukas et al. 2012; Slingo et al. 2014; Bertin et al. 2014; Breilh et al. 2013). Delft3D was forced by atmospheric pressure and wind fields obtained from the ERA-Interim database (Dee et al. 2011) and validation took place

against water level time series available from the JRC Sea Level Database (http://webcritech.jrc.ec.europa.eu/SeaLevelsDb). Typical tidal harmonic analysis was applied to obtain the residual storm surge water levels η_s which were compared directly with the model output and evaluated in terms of the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i}^{n}\left(\eta_{s,measured}^{i} - \eta_{s,\mathrm{model}}^{i}\right)^2}{n}} \quad (1)$$

Where n is the number of measurements in the storm surge time series. The relative RMSE error (%RMSE) was also estimated in order to take into account spatial variations in the range of the SSL:

$$\%RMSE = \frac{\sqrt{\frac{\sum_{i}^{n}\left(\eta_{s,measured}^{i} - \eta_{s,\mathrm{model}}^{i}\right)^2}{n}}}{\max\left(\eta_{s,measured}\right)} \times 100 \quad (2)$$

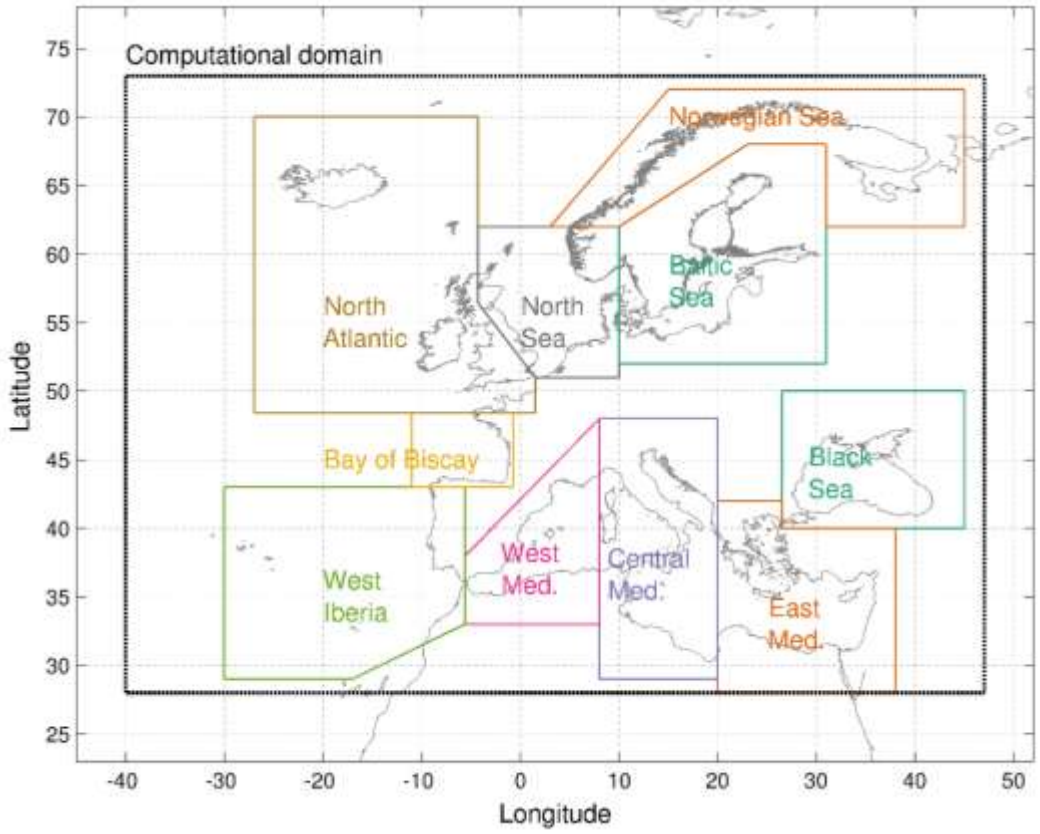


Figure 14. Map of Europe showing the Delft3D model domain (black, dashed line) showing the 10 different coastal regions defined for the analysis of the model results (color continuous lines).

4.3 MODEL PERFORMANCE

The data from the 184 tidal gauge stations were processed in order to identify data gaps or to mask periods with low-quality data, resulting in a set of 110 stations with periods of valid ground-truth

residual water level data coinciding with the simulation period. RMSE values ranged from 0.06 m to 0.29 m, while %RMSE varied between 10% and 29% (Table 2).

Table 2. Overview of model performance along the 10 defined European regions: number of available data points, as well as mean, maximum and minimum values of the RMS error in m and as a percentage of the SSL range.

Region name	No stations	RMSE mean	RMSE min	RMSE max	%RMSE mean	%RMSE min	%RMSE max
Black Sea	2	0.14	0.12	0.16	25%	22%	28%
East Mediterranean	9	0.11	0.06	0.29	24%	13%	29%
Central Mediterranean	21	0.14	0.09	0.28	18%	16%	20%
West Mediterranean	9	0.10	0.09	0.13	19%	15%	21%
West Iberia	1	0.10	0.10	0.10	17%	17%	17%
Bay of Biscay	7	0.10	0.07	0.12	16%	15%	17%
North Atlantic	36	0.12	0.08	0.23	14%	10%	20%
North Sea	14	0.17	0.11	0.22	16%	12%	19%
Baltic Sea	7	0.15	0.09	0.21	14%	11%	18%
Norwegian Sea	4	0.09	0.07	0.10	13%	11%	15%

Following the initial filtering of the tide gauge records, the region with the largest number of acceptable quality tidal gauge records was the North Atlantic (36), followed by Central Mediterranean (21), the North Sea (14) and East/West Mediterranean (9); while only one station was found in the West Iberia and 2 in the Black Sea (Table 2). Most of the Mediterranean, the Atlantic coast, and the Norwegian Sea were characterized by absolute RMSE values below 0.1 m, while RMSE>0.15 m were observed along the N. Adriatic and the North Sea (Figure 15a). The latter high RMSE values appeared to be related to the higher η_s range, as implied by the relatively low %RMSE values in the same areas (Figure 15b). The highest %RMSE values were observed along the Aegean Sea (%RMSE>0.2), while overall model performance was poorer along the Black and Mediterranean Sea (mean %RMSE ranging from 18% to 25%). On the contrary, the lowest %RMSE was observed in the Norwegian Sea (mean %RMSE=13%, see Table 2), where the model showed poor predictive skill for specific stations (Figure 15), mostly related to (i) low tidal gauge data quality, i.e. data gaps and rogue measurements; and/or (ii) the effect of the rather coarse grid resolution to the model's ability to resolve the water circulation processes along areas with complex morphology, such as the Sea of Marmara, or the Aegean Sea.

Q-Q plots indicate that the model was capable of reproducing the probability density function of the measured SSL (Figure 16), which was the crucial skill given the extreme value statistical analysis to follow. The model appears to overestimate the lower values in some locations in S. Europe (e.g. Figure 16c,f), while the opposite was observed in the North, Baltic and Norwegian Sea (e.g. Figure 16g-j); however the lower values are of minor interest for the scope of the present study. On the other hand, the extremes appear to be underestimated in several cases (e.g. Figure 16b-f), a possible artifact of the coarse atmospheric forcing and simulation grid. However, the model's capacity to simulate SSL is considered satisfactory given that the scope of the study is to project trends under climate change scenarios and not to provide accurate operational forecasts.

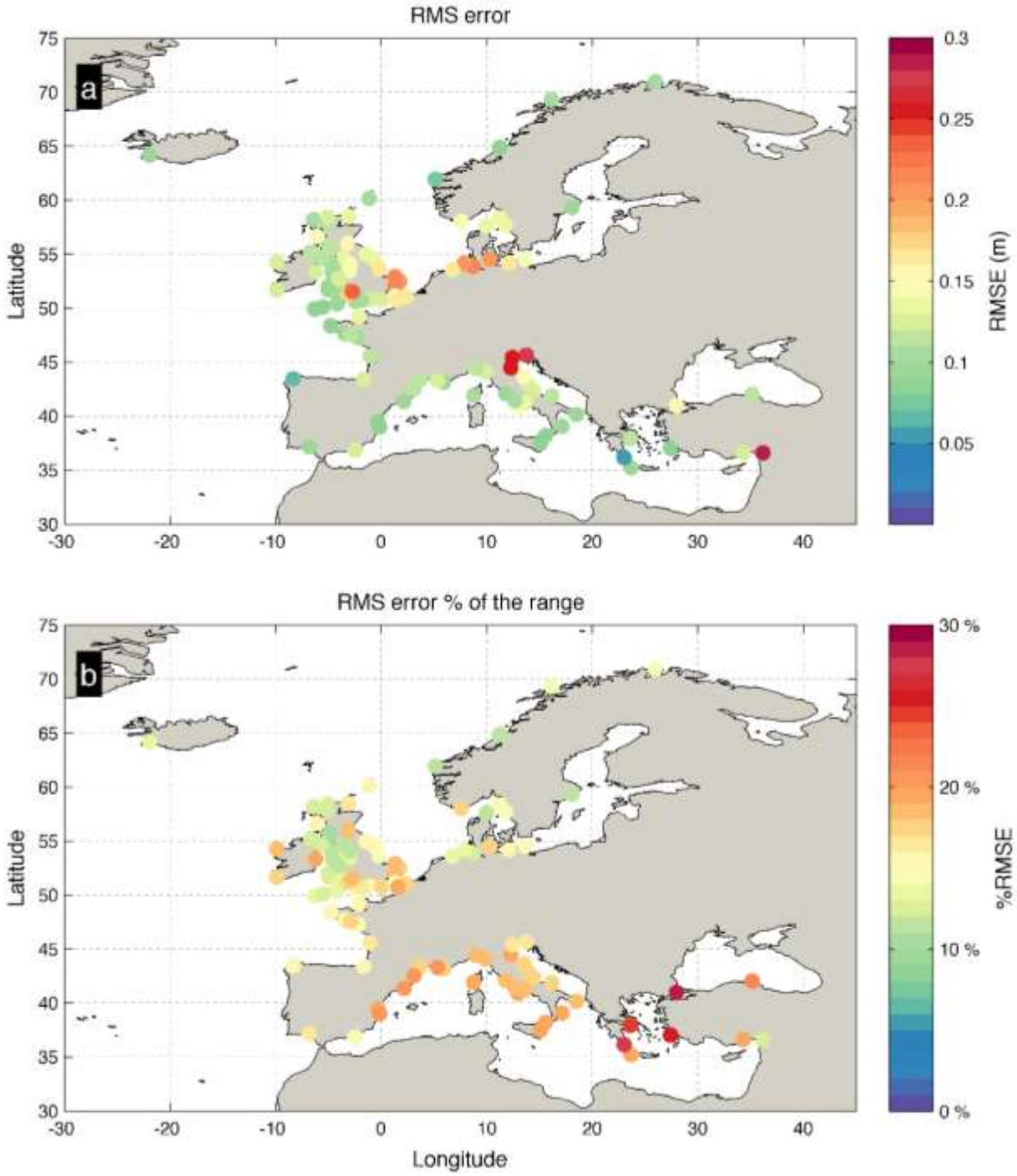


Figure 15. Model validation performance: scatter plot showing RMS error in m (a) and as a percentage of the SSL range (b) for all the available tidal gauge stations.

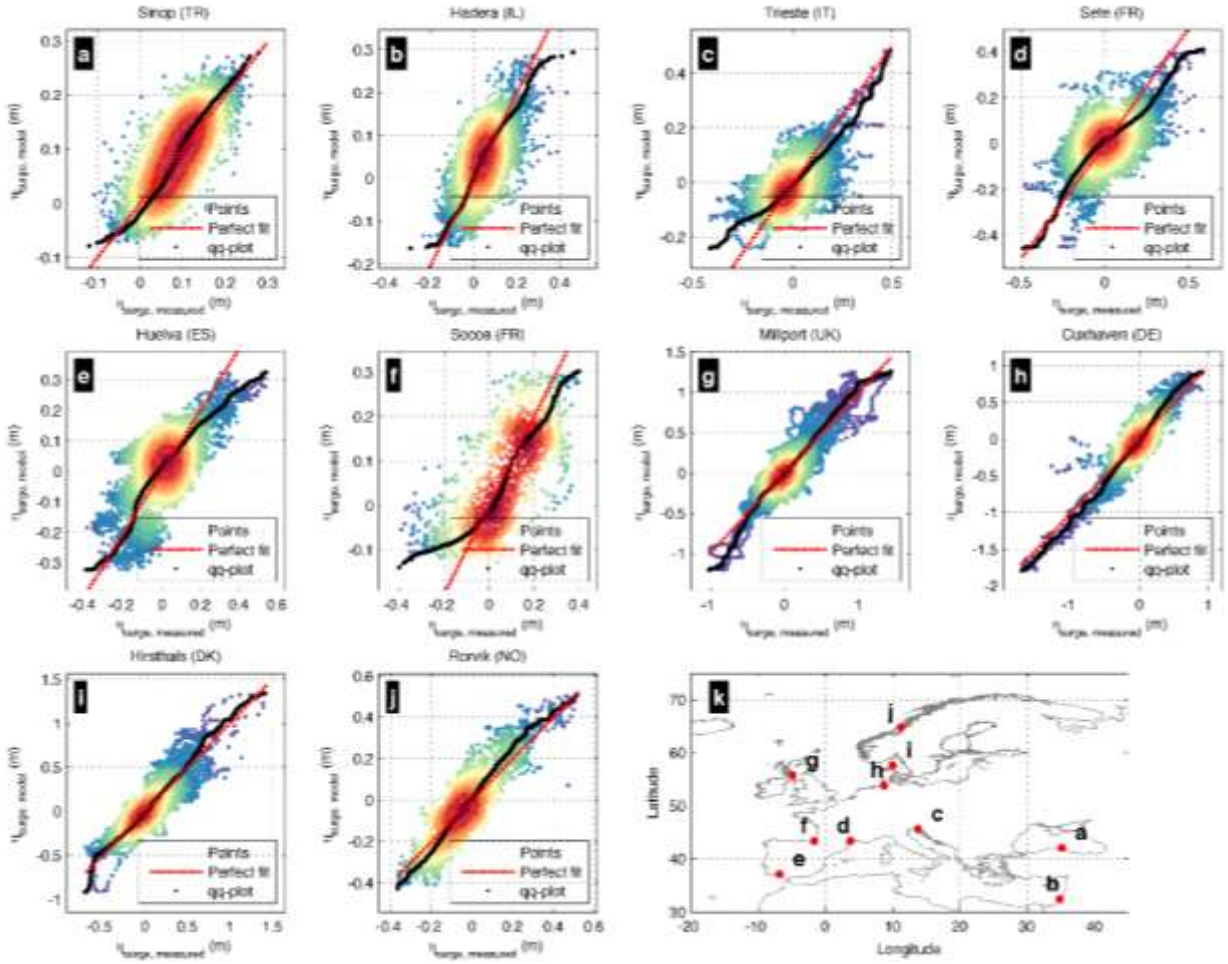


Figure 16. Model validation performance: scatter plots comparing measured and simulated SSL for all the tidal gauge locations; colors express point density (increasing from blue to red), the red dashed line expresses the perfect fit, while the black dots show the q-q plots of the two time series. The inset map (k) shows the location of the displayed tide gauge records, one from each of the 10 European regions.

5. ORCHIDEE-CROP

The ORCHIDEE land surface model has recently undergone further developments to include a crop model fully coupled with the carbon and water fluxes simulated by ORCHIDEE (see HELIX Deliverable 7.2 Section 5.2.1. for details). Evaluation of ORCHIDEE-crop for extreme events has not yet been performed, but to provide context for the future projections presented in Deliverable 7.2, a preliminary comparison of ORCHIDEE-crop simulations of present-day yields with previously-published estimates has been carried out for key European crops (Figure 17)

The focus was on ORCHIDEE-crop driven by climate simulated by 5 EURO-CORDEX projections with different combinations of Regional Climate Model and driving General Circulation Model, as used for Deliverable 7.2 but for the historical period of the simulation (See HELIX Deliverable 7.2 Table 5.1). As described in HELIX deliverable 7.2 section 5.1, some of the RCM simulations were bias-corrected and some were not. The preference would have been to use bias-correction for all crop model simulations, as the existence of critical thresholds in meteorological influences on crops can mean that systematic biases lead to unrealistic responses to changes in climate. However, bias-corrected projections were not available for all RCMs. Moreover, the use of bias-correction can mean that shortcomings in the climate simulations are not evident, potentially leading to over-confidence in the results. This will however mean that some of the ORCHIDEE-crop simulations will be affected by any systematic biases in the RCM simulations.

The ORCHIDEE-crop yield simulations at 0.11° (~12.5km) resolution were compared against a dataset at 1.125° resolution produced though a modelled combination of agricultural datasets in the year 2000, yield statistics at the country level, and estimates of Net Primary Production from remote sensing (Iizumi et al., 2014). The historical ORCHIDEE-crop simulations were from 1976-2005, and the Iizumi et al. dataset was available from 1982-2006. Bearing in mind that winter wheat sowed in 2005 did not mature before 2006, the ORCHIDEE-crop and Iizumi et al. results were therefore compared for the overlap period of 1982-2004. Figure 17 presents the mean yields for wheat, maize and soybean from OCHIDEE-crop and Iizumi et al. over this overlap period. The figure shows the results for the simulation which gave the best spatial correlation between ORCHIDEE-crop and Iizumi et al. results, to provide an upper bound on the skill of the simulations. The results shown for wheat (Figure 17b) are those simulated by ORCHIDEE-crop driven by the IPSL_RCA GCM/RCM combination, with bias-correction. Those for maize (Figure 17d) and soybean (Figure 17f) are from simulations driven by the MPI_RCA and MPI_CCLM RCM/GCM combination, with no bias-correction.

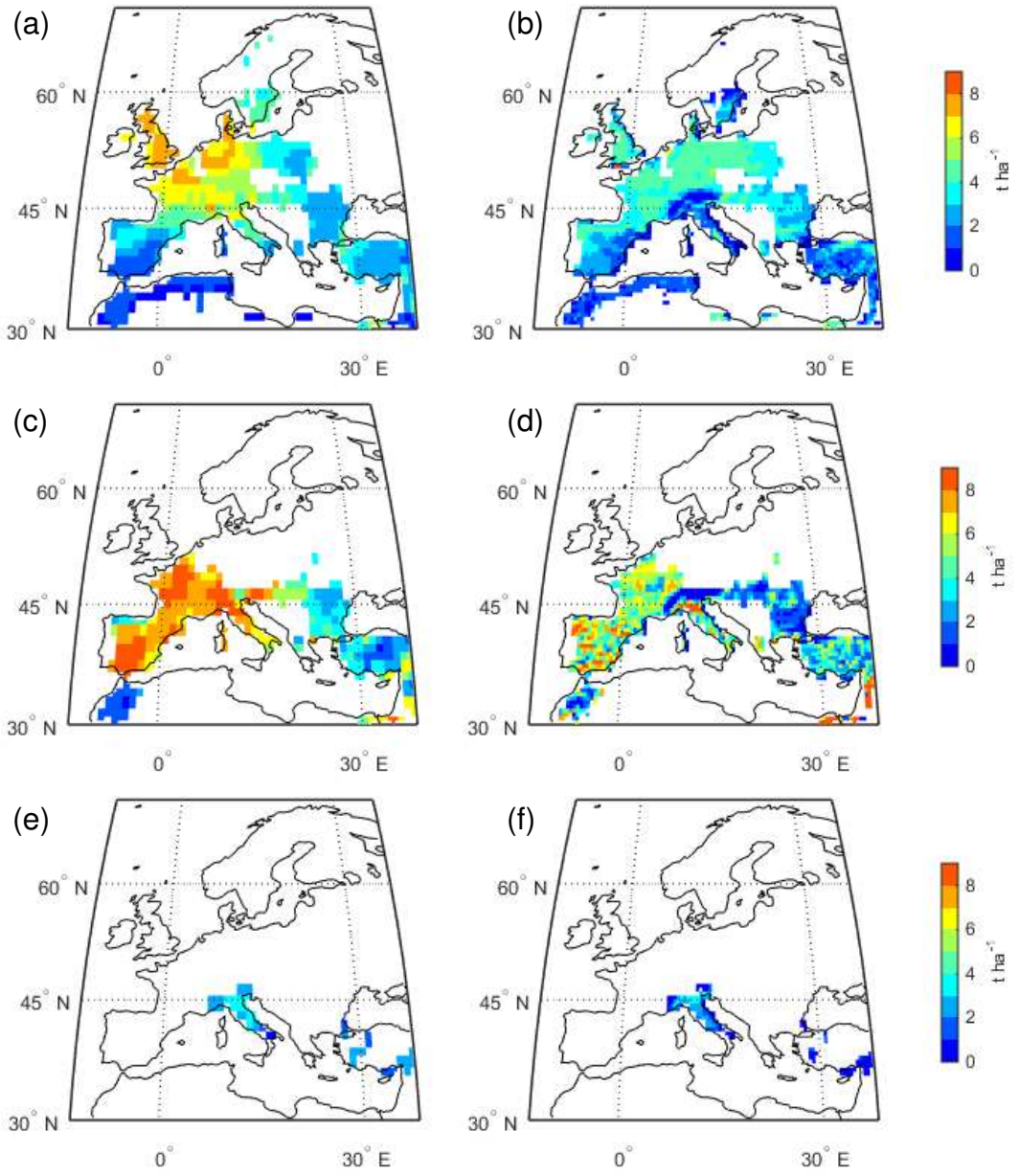


Figure 17. *Spatial distribution of the average yield during 1982-2004 from Iizumi et al. (2014; left panels) and from the ORCHIDEE-crop simulation driven by selected EURO-CORDEX climate forcings (right panels) for (a,b) wheat, (c,d) maize and (e,f) soybean, where panels (b), (d) and (f) show the ORCHIDEE-crop simulation with the largest spatial correlation with Iizumi et al results. Note that different EURO-CORDEX forcings were used for each crop – this figure presents results for highest spatial correlation with the Iizumi et al. dataset for each crop. The ORCHIDEE-crop simulation of wheat (b) were driven by climate simulated by the IPSL_RCA GCM/RCM combination with bias-correction. That of maize (d) was driven by MPI_RCA climate, and that of soybean (f) was driven by MPI_CCLM climate.*

For wheat, pattern of variation in yields across Europe simulated by ORCHIDEE-crop driven by IPSL_RCA climate was broadly similar to that in the Iizumi et al., dataset, with higher yields in north-west Europe including the UK, northern France, Belgium, the Netherlands, Luxembourg, Denmark and Germany, and the lowest yields around the Mediterranean (Figure 17 a,b). However, ORCHIDEE-crop generally simulated lower yields than Iizumi et al. (2014), with particularly large differences in north-western Europe with the result that the gradient in yields across the region was not as large in ORCHIDEE-crop as Iizumui et al. Differences were smaller in eastern and southern Europe, although ORCHIDEE-crop yields were still generally less than those of Iizumi et al. The spatial correlation R value was 0.49 . Since the IPSL_RCA climate data was bias-corrected, it is assumed that biases in the driving climate data are not the main cause of the differences between ORCHIDEE-crop and Iizumui wheat yields. The reasons for these differences require further investigation.

For maize, the highest spatial correlation between maize yields in the two methods (R=0.31) for any of the 5 driving climate datasets was obtained with ORCHIDEE-crop driven by the MPI_RCA combination. Again,.the ORCHIDEE-crop simulated spatial variations which broadly agreed with those in Iizumui et al., with higher maize yields in France and Spain and lower yields in central Europe and the Balkans (Figure 17 c,d). As with wheat, ORCHIDEE-crop tended to simulate smaller yields than Iizumui et al., especially in western and central Europe. The closest agreement between the two was in Turkey. A noticeable feature of the ORCHIDEE maize simulations was a high level of spatial noise at high resolution, which was not evident in the Iizumui et al. dataset –this might however be at least partly a consequence of the lower spatial resolution in Iizumi et al. in comparison with ORCHIDEE-crop.

Growth of soy bean is concentrated in Italy and parts of Turkey. The highest correlation between ORCHIDEE-crop and Iizumi (R=0.56) were obtained with simulations driven by the MPI-CCLM GCM-RCM combination, without bias-correction. Here, ORCHIDEE-crop simulates slightly smaller soybean yields than the values given in Iizumi et al., and the general variation across these countries appeared to be similar in ORCHIDEE-crop and Iizumui et al. (Figure 17 e,f).

Overall, ORCHIDEE-crop simulates broad continental-scale patterns of wheat, maize and soybean yields which agree with the general patterns seen in an existing dataset of observational and model-based information. However there is a general tendency for ORCHIDEE to simulate lower yields than in the other dataset, and the differences are large in areas where the absolute yield in each of the datasets is larger than in other parts of the region. However, it should be noted that ORCHIDEE-crop results for maize and soybean shown here were from simulations driven by climate model outputs which had *not* been bias-corrected. Therefore it is possible that some of the differences in yields are due to biases in the driving climate data. Nevertheless, it is perhaps surprising that a higher spatial correlation was obtained using non-bias corrected climate data. This requires further investigation.

6. REFERENCES

Alfieri, L. et al., 2014. Advances in pan-European flood hazard mapping. *Hydrological Processes*, 28(13), pp.4067–4077.

Alfieri, L. et al., 2013. GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrol. Earth Syst. Sci.*, 17(3), pp.1161–1175.

Alfieri, L., Thielen, J. & Pappenberger, F., 2012. Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland. *J. Hydrol.*, 424–425, pp.143–153.

Bartholmes, J., Thielen, J. & Ramos, M.H., 2006. Quantitative analyses of EFAS forecasts using different verification (skill) scores. *The benefit of probabilistic flood forecasting on European scale, edited by: Thielen, J., EUR*, 22560, pp.58–79.

Batista e Silva, F., Lavalle, C. & Koomen, E., 2012. A procedure to obtain a refined European land use/cover map. *Journal of Land Use Science*, pp.1–29.

Bertin, X. et al., 2014. A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. *Coastal Engineering*, 94, pp.80–89.

Best, M.J. et al., 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4(3), pp.677–699.

Beven, K.J. & Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24(1), pp.43–69.

Breilh, J.F. et al., 2013. Assessment of static flood modeling techniques: application to contrasting marshes flooded during Xynthia (western France). *Natural Hazards and Earth System Science*, 13(6), pp.1595–1612.

Burek, P., van der Knijff, J. & Ntegeka, V., 2013. *LISVAP, Evaporation Pre-Processor for the LISFLOOD Water Balance and Flood Simulation Model - Revised User Manual*, Joint Research Centre –Institute for Environment and Sustainability. EUR 26167 EN. 36 pp.

Clark, D.B. et al., 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3), pp.701–722.

Cox, P.., Huntingford, C. & Harding, R.., 1998. A canopy conductance and photosynthesis model for use in a GCM land surface scheme. *Journal of Hydrology*, 212-213, pp.79–94.

Cox, P.M., 2001. Description of the " TRIFFID " Dynamic Global Vegetation Model.

Dee, D.P. et al., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), pp.553–597. Available at: http://doi.wiley.com/10.1002/qj.828 [Accessed July 11, 2014].

Falloon, P. et al., 2011. Validation of River Flows in HadGEM1 and HadCM3 with the TRIP River Flow Model. *Journal of Hydrometeorology*, 12(6), pp.1157–1180. Available at: http://journals.ametsoc.org/doi/abs/10.1175/2011JHM1388.1 [Accessed April 6, 2015].

Fekete, B. & Vorosmarty, C.J., 2011. ISLSCP II UNH/GRDC Composite Monthly Runoff. *ISLSCP Initiative II Collection, edited by: Hall, FG, Collatz, G., Meeson, B., Los, S., Brown de Colstoun, E., and Landis, D., Data set, available at: http://daac. ornl. go v/, from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridg*, 10.

Feyen, L. et al., 2007. Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology*, 332(3-4), pp.276–289.

Forzieri, G. et al., 2014. Ensemble projections of future streamflow droughts in Europe. *Hydrology and Earth System Sciences*, 18(1), pp.85–108. Available at: http://www.hydrol-earth-syst-sci.net/18/85/2014/hess-18-85-2014.html [Accessed April 6, 2015].

Gosling, S.N. & Arnell, N.W., 2011. Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. *Hydrological Processes*, 25(7), pp.1129–1145. Available at: http://doi.wiley.com/10.1002/hyp.7727 [Accessed March 2, 2015].

Grinsted, A., Moore, J.C. & Jevrejeva, S., 2010. Reconstructing sea level from paleo and projected temperatures 200 to 2100 AD. *Climate Dynamics*, 34(4), pp.461–472.

Haddeland, I. et al., 2011. Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. *Journal of Hydrometeorology*, 12(5), pp.869–884. Available at: http://journals.ametsoc.org/doi/abs/10.1175/2011JHM1324.1 [Accessed April 7, 2015].

Horsburgh, K. & De Vries, H., 2011. Guide to Storm Surge Forecasting.

Iglesias, A. et al., 2012. A regional comparison of the effects of climate change on agricultural crops in Europe. *Climatic Change*, 112(1), pp.29–46.

Iizumi, T., Yokozawa, M., Sakurai, G., Travasso, M.I., Romanenkov, V., Oettli, P., Newby, T., Ishigooka, Y. & Furuya, J. (2014) Historical changes in global yields: major cereal and legume crops from 1982 to 2006. *Global Ecology and Biogeography*, **23**, 346-357.Jarvis, A. et al., 2008. Hole-filled seamless SRTM data V4. *International Centre for Tropical Agriculture (CIAT)*.

Kamari, J. et al., 2008. Envisioning the future of water in Europe - the SCENES project. *E-WAter*, pp.1–28.

Kennedy, J. & Eberhart, R., 1995. Particle swarm optimization. In *Proceedings of IEEE international conference on neural networks*. Perth, Australia, pp. 1942–1948.

King, D., Daroussin, J. & Tavernier, R., 1994. Development of a soil geographic database from the soil map of the European Communities. *Catena*, 21(1), pp.37–56.

Van der Knijff, J.M., Younis, J. & de Roo, A.P.J., 2010. LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *Int. J. Geogr. Inf. Sci.*, 24(2), pp.189–212.

Losada, I.J. et al., 2013. Long-term changes in sea-level components in Latin America and the Caribbean. *Global and Planetary Change*, 104, pp.34–50.

Lowe, J.A. et al., 2010. Past and future changes in extreme sea levels and waves. *Understanding Sea-Level Rise and Variability*, pp.326–375.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30(2), pp.273–297.

Ntegeka, V. et al., 2013. EFAS-Meteo: A European daily high-resolution gridded meteorological data set for 1990-2011.

Oki, T. & Sud, Y.C., 1998. Design of Total Runoff Integrating Pathways (TRIP)— A Global River Channel Network. , 2(1), pp.7–22.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. pp. 120–145.

Pryor, M., Clark, D., Harris, P., Hendry, M., 2012. Joint UK Land Environment Simulator (JULES) Version 3.2 User Manual.

Refsgaard, J.C. et al., 2013. A framework for testing the ability of models to project climate change and its impacts. *Climatic Change*, 122(1-2), pp.271–282. Available at: http://link.springer.com/10.1007/s10584-013-0990-2 [Accessed January 28, 2015].

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, 1(5), pp.318–333.

Rojas, R. et al., 2012. Assessment of future flood hazard in Europe using a large ensemble of bias-corrected regional climate simulations. *Journal of Geophysical Research D: Atmospheres*, 117(17).

Rojas, R., Feyen, L. & Watkiss, P., 2013. Climate change and river floods in the European Union: Socio-economic consequences and the costs and benefits of adaptation. *Global Environmental Change*, 23(6), pp.1737–1751.

De Roo, A. et al., 2001. Assessing the effects of land use changes on floods in the meuse and oder catchment. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(7-8), pp.593–599.

Slingo, J. et al., 2014. The recent storms and floods in the UK.

Thiemig, V. et al., 2013. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *Journal of Hydrology*, 499, pp.324–338.

Vogt, J. & Foisneau, S., 2007. CCM River and Catchment Database Version 2.0.

Vousdoukas, M.I., Almeida, L.P.M. & Ferreira, Ó., 2012. Beach erosion and recovery during consecutive storms at a steep-sloping, meso-tidal beach. *Earth Surface Processes and Landforms*, 37(6), pp.583–593.

Wagener, T. et al., 2010. The future of hydrology: An evolving science for a changing world. *Water Resources Research*, 46(5), p.n/a–n/a. Available at: http://doi.wiley.com/10.1029/2009WR008906 [Accessed August 20, 2014].

Weedon, G. et al., 2014. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50((9)), pp.7505–7514.

Wesseling, C.G. et al., 1996. Integrating dynamic environmental models in GIS: The development of a dynamic modelling language. *Transactions in GIS*, 1(1), pp.40–48.

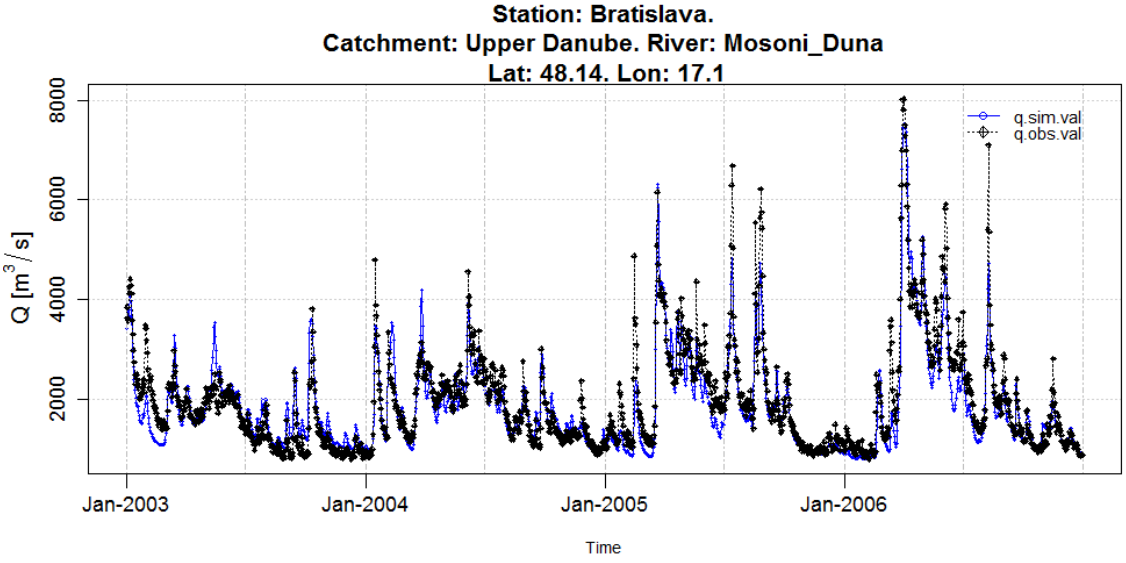
Wösten, J.H.M. et al., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma*, 90(3), pp.169–185.

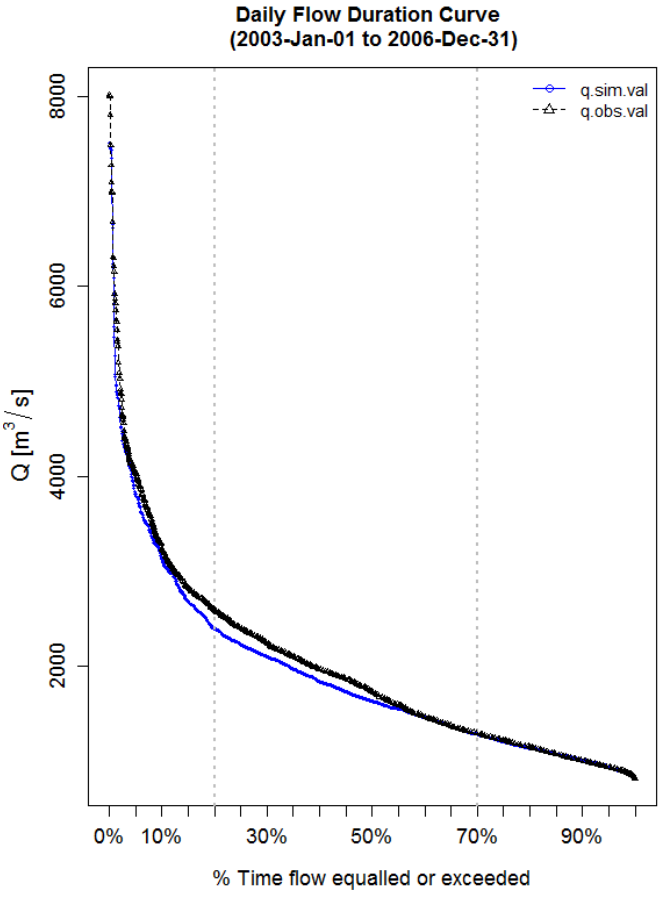
Younis, J., Anquetin, S. & Thielen, J., 2008. The benefit of high-resolution operational weather forecasts for flash flood warning. *Hydrology and Earth System Sciences*, 12(4), pp.1039–1051.

Zambrano-Bigiarini, M. & Rojas, R., 2014. *hydroPSO: Particle Swarm Optimisation, with focus on Environmental Models*,

APPENDIX A - LISFLOOD CALIBRATION PERFORMANCE

In the following, score cards of Lisflood calibration are presented. One river station is shown for each of the 18 European river basins in the simulation window with upstream area at the mouth larger than 50,000 km^2.





Internal Station ID	C046 (A04)
Upstream Area	132,550 [km2] (Lisflood)
	131,331 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.83

Validation Period
(2003-Jan-01 to 2006-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	815	817
Avg.	1968.2	1885.2
q95	4028	3806.1
q99	5889.8	5149
Max.	8020	7514.6

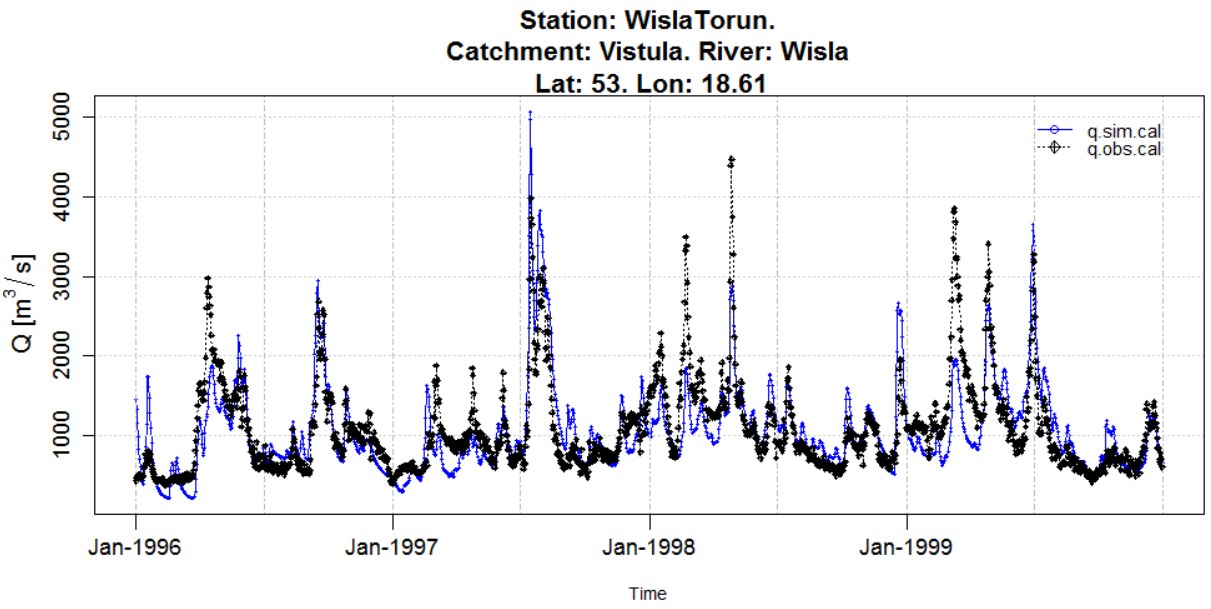
Error Measures, [m3/s]

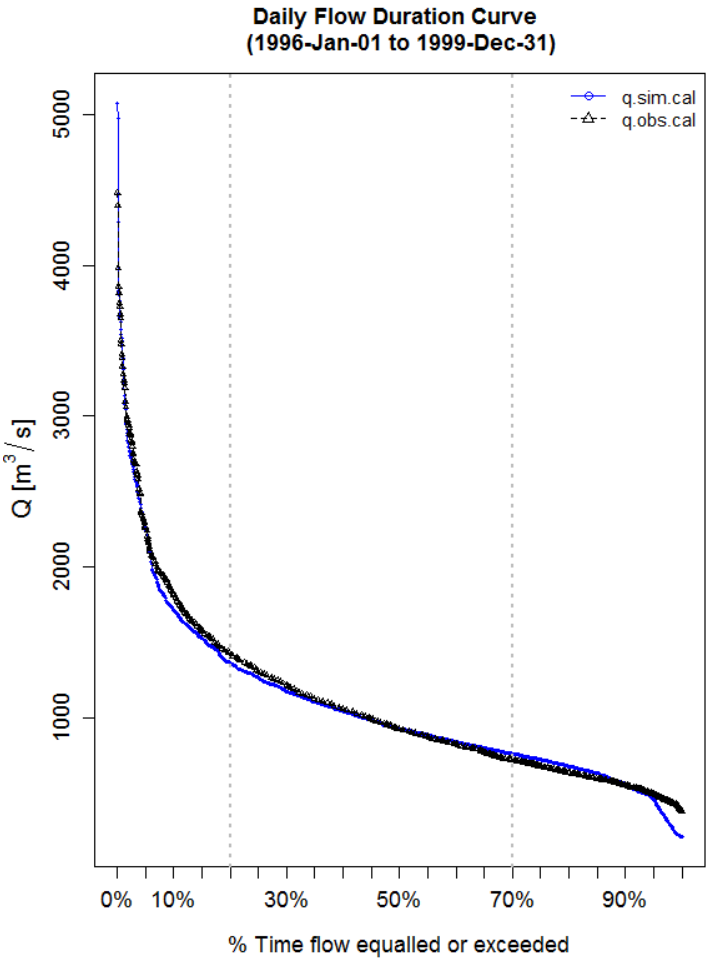
RMSE	419.7
MAE	289

Goodness-of-Fit

PBIAS	-4.2%
R	0.92
NSE	0.84

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C215 (A05)
Upstream Area	179,500 [km2] (Lisflood)
	Not Available (Provider)
Calibration Period	1996-Jan-01 to 1999-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.57

Calibration Period
(1996-Jan-01 to 1999-Dec-31)

Summary Statistics, [m3/s]

	q.obs.cal	q.sim.cal
Min.	381	210.9
Avg.	1087.3	1071.7
q95	2241	2241
q99	3280	3342.6
Max.	4480	5074.9

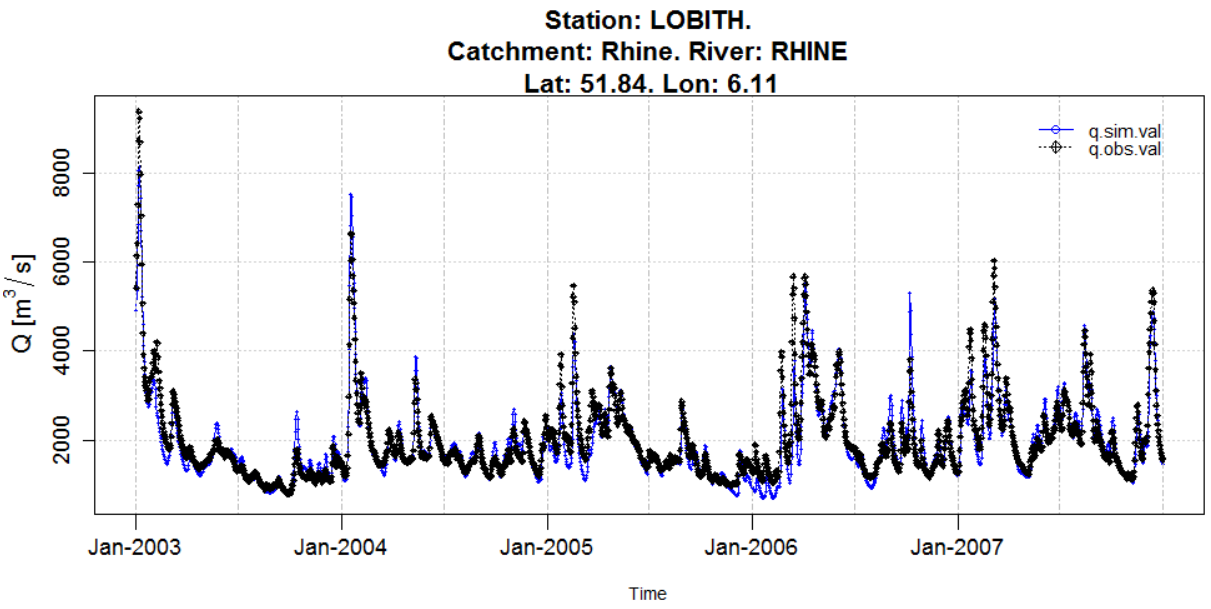
Error Measures, [m3/s]

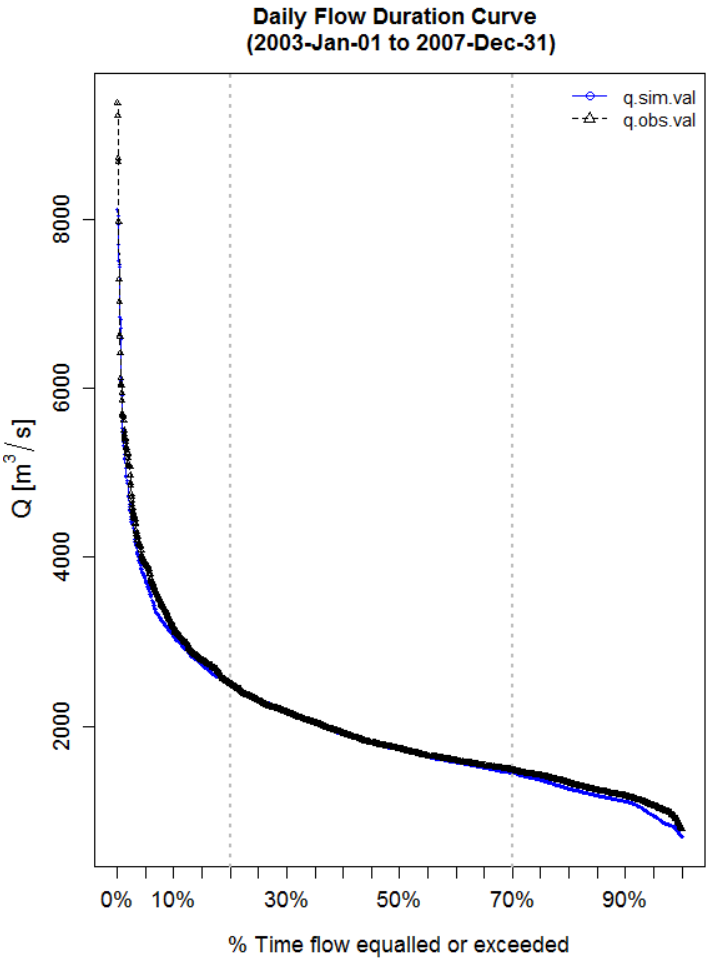
RMSE	387.7
MAE	275

Goodness-of-Fit

PBIAS	-1.4%
R	0.78
NSE	0.57

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C368 (A03)
Upstream Area	160,475 [km2] (Lisflood)
	160,800 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.93

Validation Period
(2003-Jan-01 to 2007-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	788	690.2
Avg.	2016.4	1960.8
q95	3922.2	3705.2
q99	5673.2	5441.3
Max.	9372	8117.5

Error Measures, [m3/s]

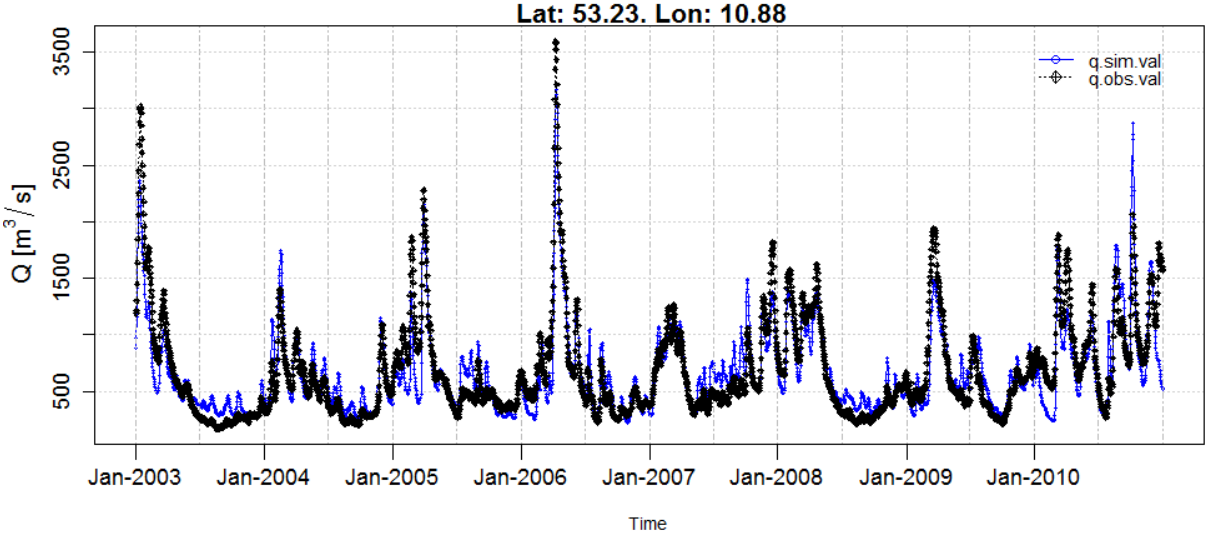
RMSE	350.9
MAE	223.1

Goodness-of-Fit

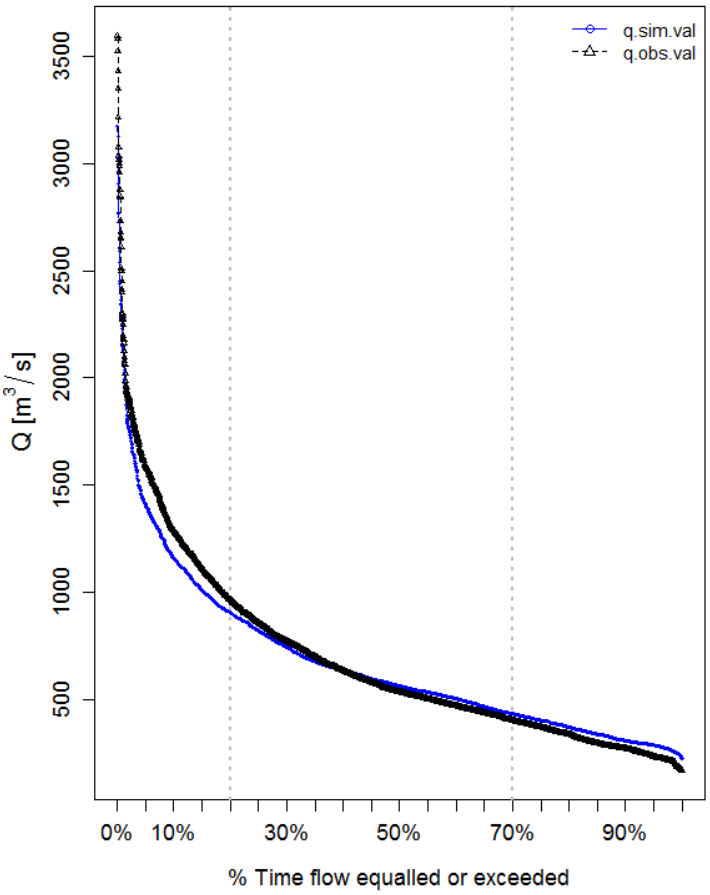
PBIAS	-2.8%
R	0.94
NSE	0.87

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: NeuDarchau.
Catchment: Elbe. River: Elbe
Lat: 53.23. Lon: 10.88



Daily Flow Duration Curve
(2003-Jan-01 to 2010-Dec-31)



Internal Station ID	C172 (A05)
Upstream Area	130,800 [km2] (Lisflood)
	131,950 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.82

Validation Period
(2003-Jan-01 to 2010-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	173	226.2
Avg.	683.2	669.8
q95	1587.3	1401.8
q99	2245.1	2127.5
Max.	3593.3	3173.5

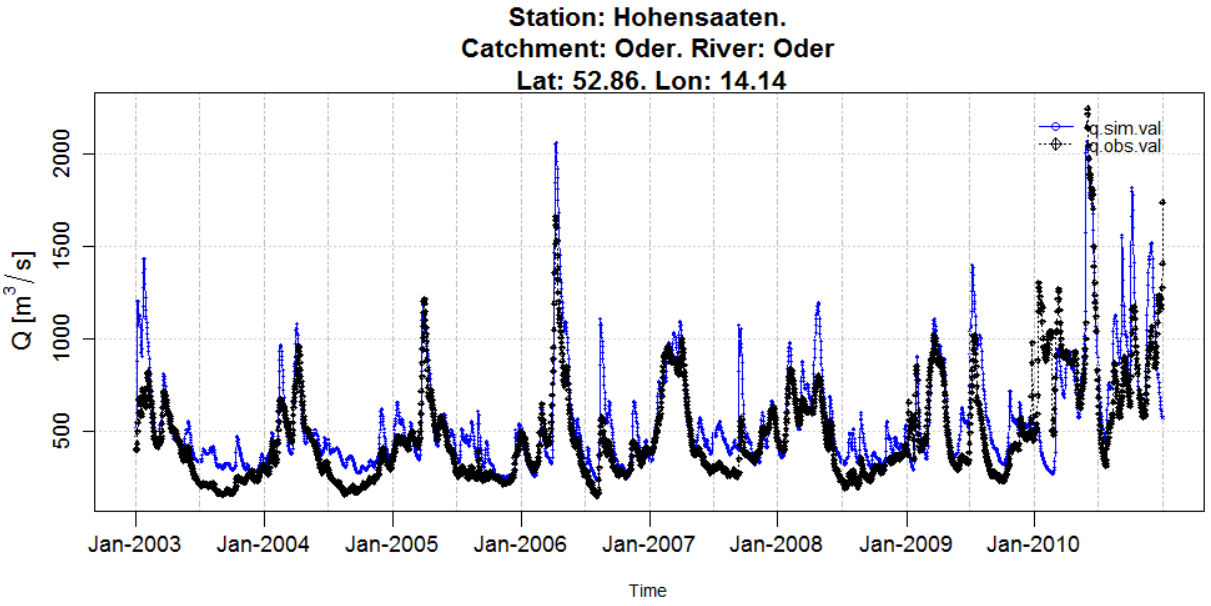
Error Measures, [m3/s]

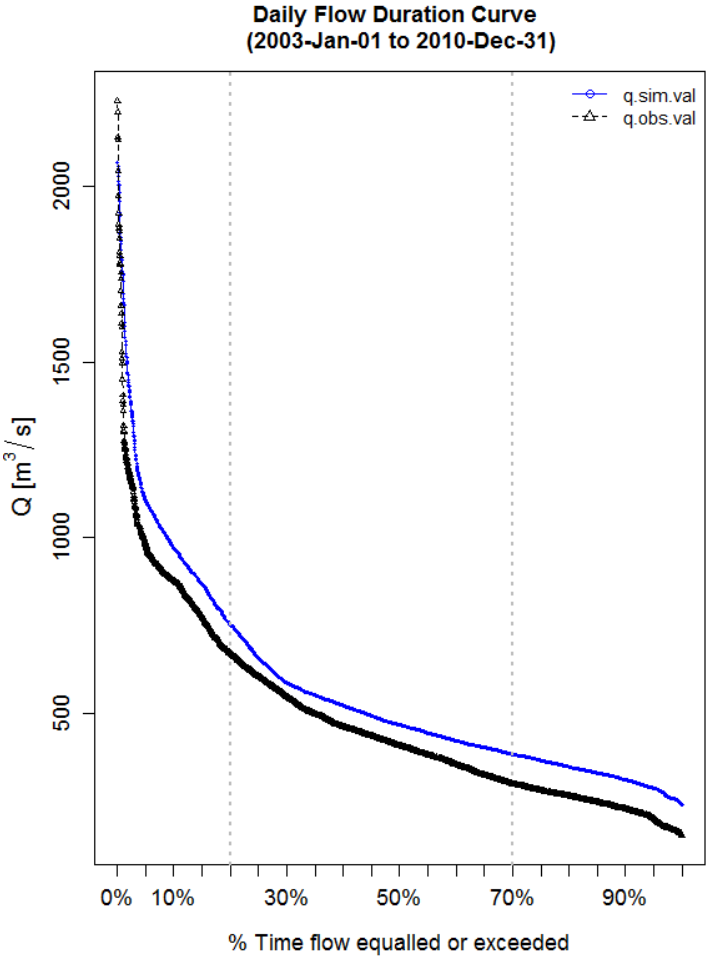
RMSE	206.6
MAE	145.1

Goodness-of-Fit

PBIAS	-2.2%
R	0.89
NSE	0.79

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C178 (A05)
Upstream Area	112,925 [km2] (Lisflood)
	109,564 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.43

Validation Period
(2003-Jan-01 to 2010-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	156	241.7
Avg.	484.6	564.1
q95	979.3	1105.5
q99	1381	1743.6
Max.	2241.6	2065.3

Error Measures, [m3/s]

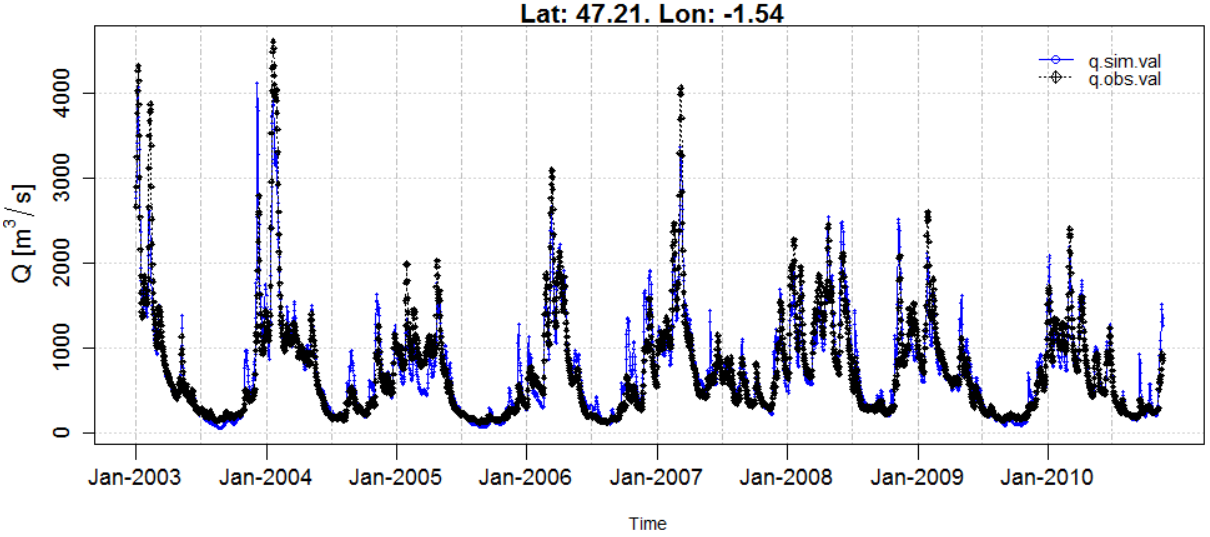
RMSE	185.4
MAE	131.1

Goodness-of-Fit

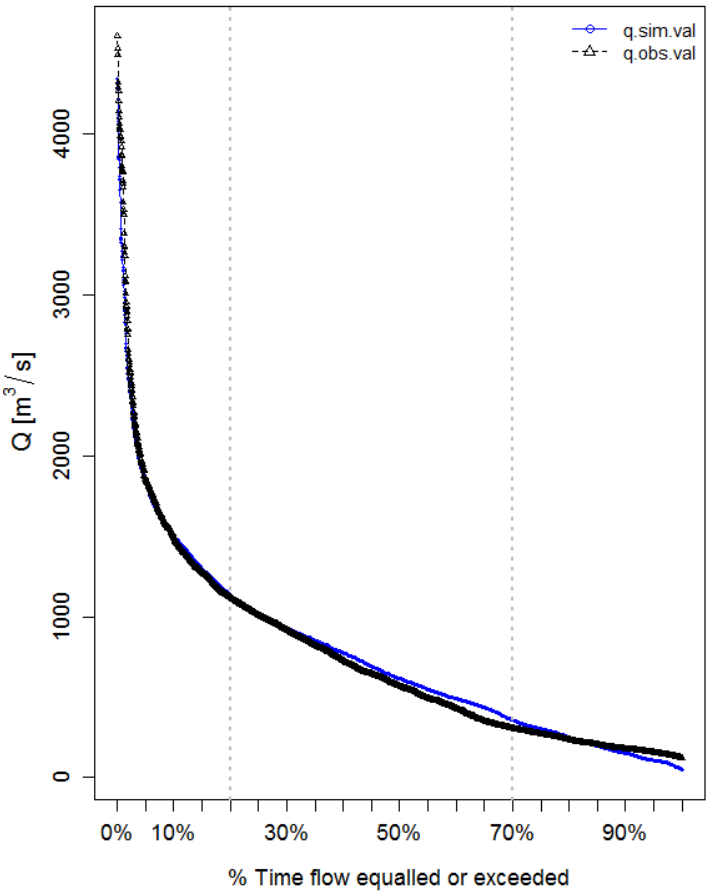
PBIAS	15.7%
R	0.82
NSE	0.54

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: La_Loire_totale_a_Nantes_Saint_Felix.
Catchment: Loire. River: Loire
Lat: 47.21. Lon: -1.54



Daily Flow Duration Curve
(2003-Jan-01 to 2010-Nov-18)



Internal Station ID	C084 (A03)
Upstream Area	111,150 [km2] (Lisflood)
	111,570 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.92

Validation Period
(2003-Jan-01 to 2010-Nov-18)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	121.9	48.2
Avg.	744.8	753.4
q95	1847.5	1827.3
q99	3677.5	3156.8
Max.	4607.9	4341

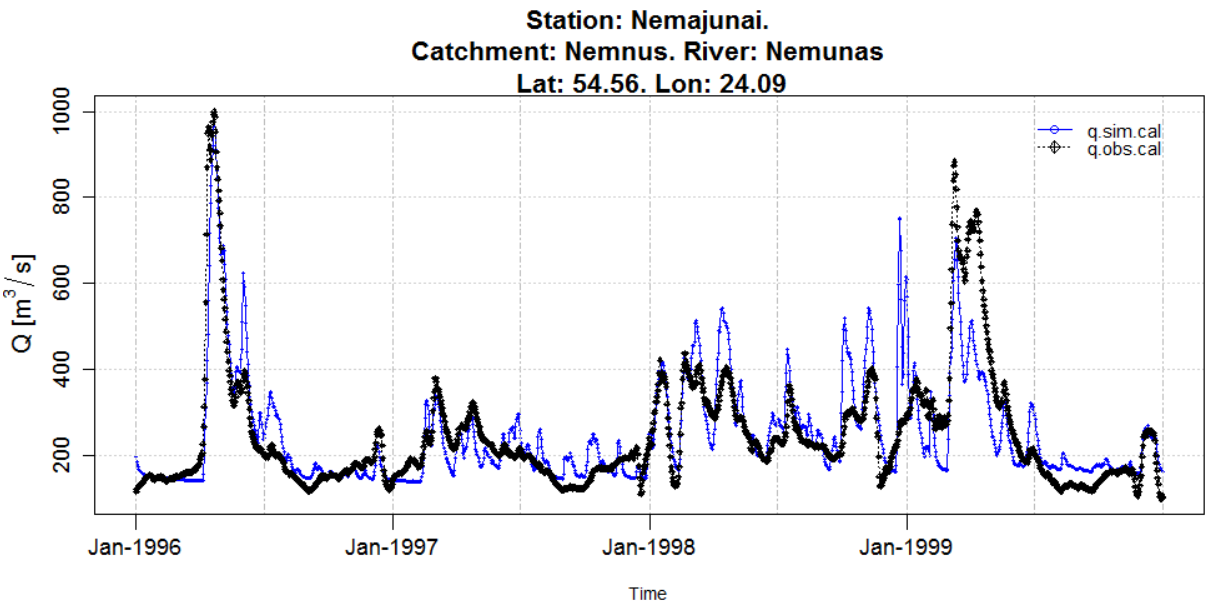
Error Measures, [m3/s]

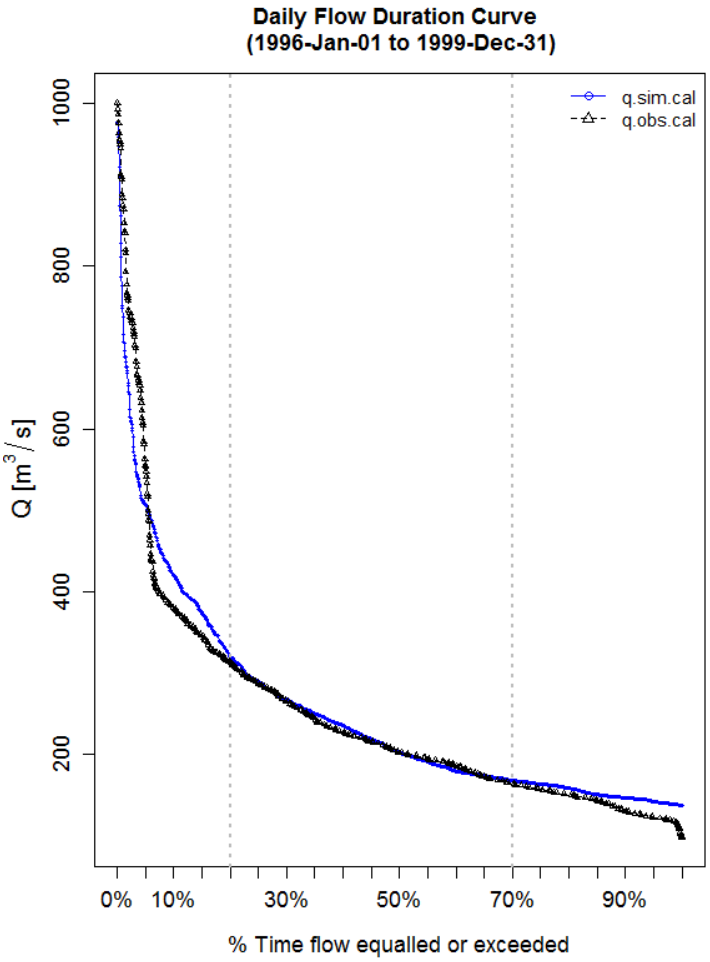
RMSE	227.4
MAE	143.1

Goodness-of-Fit

PBIAS	1.2%
R	0.94
NSE	0.88

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C174 (A05)
Upstream Area	43,575 [km2] (Lisflood)
	Not Available (Provider)
Calibration Period	1996-Jan-01 to 1999-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.69

Calibration Period
(1996-Jan-01 to 1999-Dec-31)

Summary Statistics, [m3/s]

	q.obs.cal	q.sim.cal
Min.	98.9	137
Avg.	246.5	250.3
q95	554	506.8
q99	871.6	724.4
Max.	1000	976.5

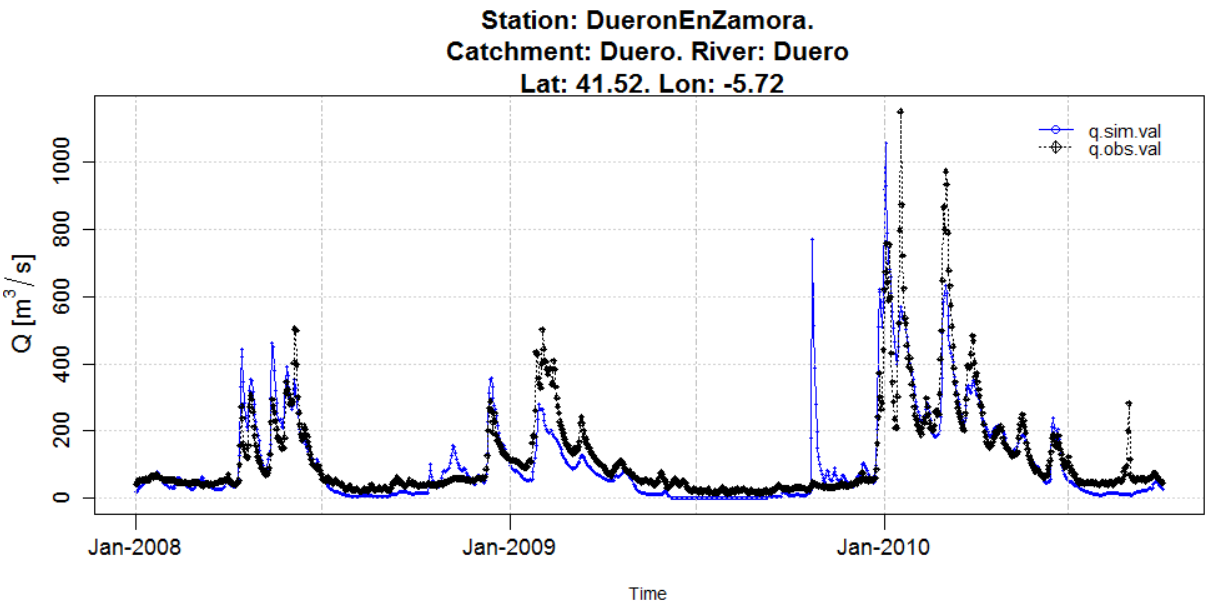
Error Measures, [m3/s]

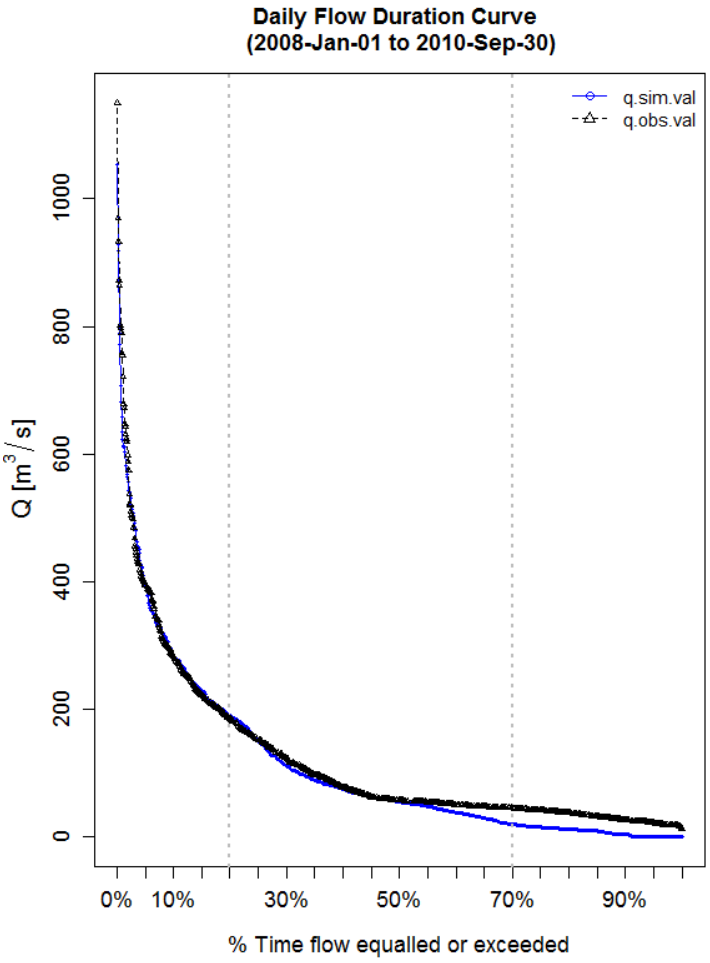
RMSE	80.3
MAE	51.8

Goodness-of-Fit

PBIAS	1.5%
R	0.83
NSE	0.69

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C113 (A01)
Upstream Area	45,375 [km2] (Lisflood)
	46,137 [km2] (Provider)
Calibration Period	2003-Jan-01 to 2007-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.46

Validation Period
(2008-Jan-01 to 2010-Sep-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	12.4	0
Avg.	119.3	107.7
q95	394	389
q99	720.7	613.5
Max.	1150	1054.4

Error Measures, [m3/s]

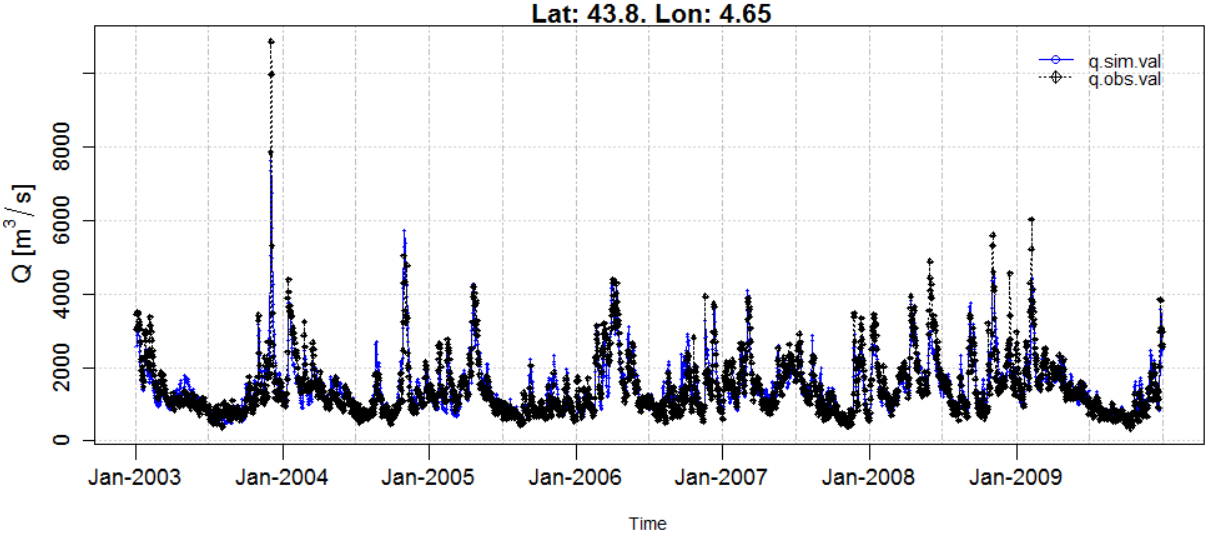
RMSE	72.4
MAE	42

Goodness-of-Fit

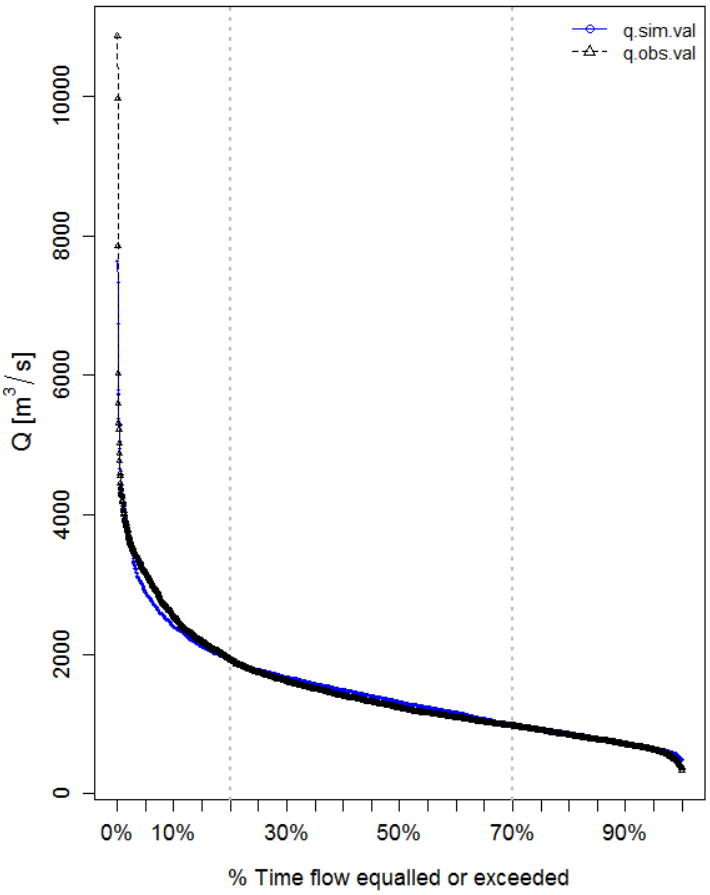
PBIAS	-9.8%
R	0.87
NSE	0.73

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: Le_Rhone_a_Beaucaire.
Catchment: Rhone. River: Rhone
Lat: 43.8. Lon: 4.65



Daily Flow Duration Curve
(2003-Jan-01 to 2009-Dec-31)



Internal Station ID	C098 (A03)
Upstream Area	95,550 [km2] (Lisflood)
	95,590 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.84

Validation Period
(2003-Jan-01 to 2009-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	341.6	483
Avg.	1458.9	1474.4
q95	3164.1	2885.4
q99	4115.1	4228.1
Max.	10861	7625.2

Error Measures, [m3/s]

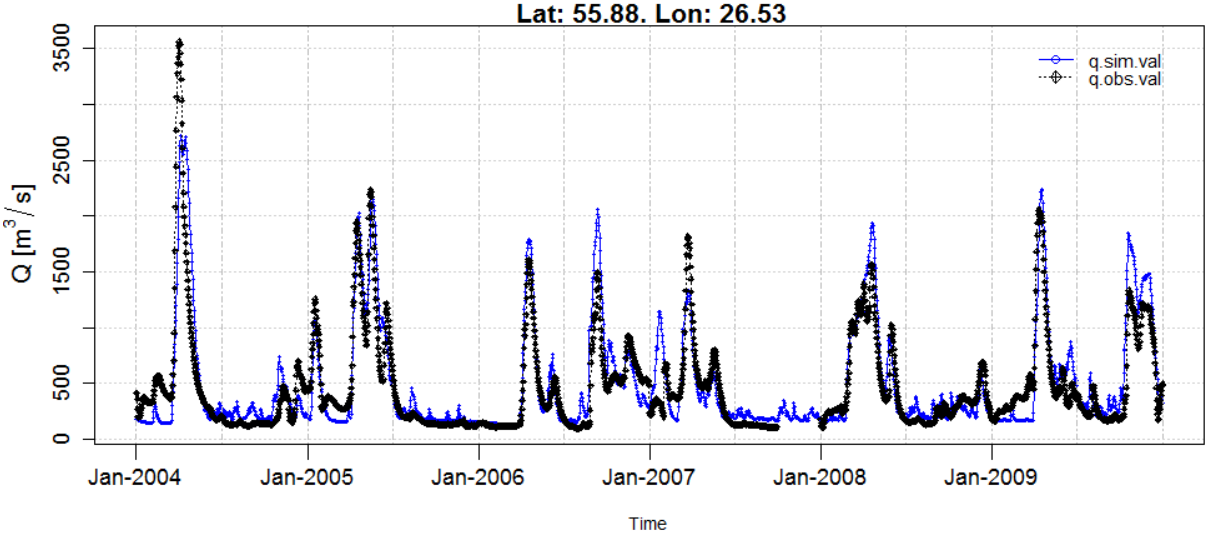
RMSE	387.4
MAE	273.7

Goodness-of-Fit

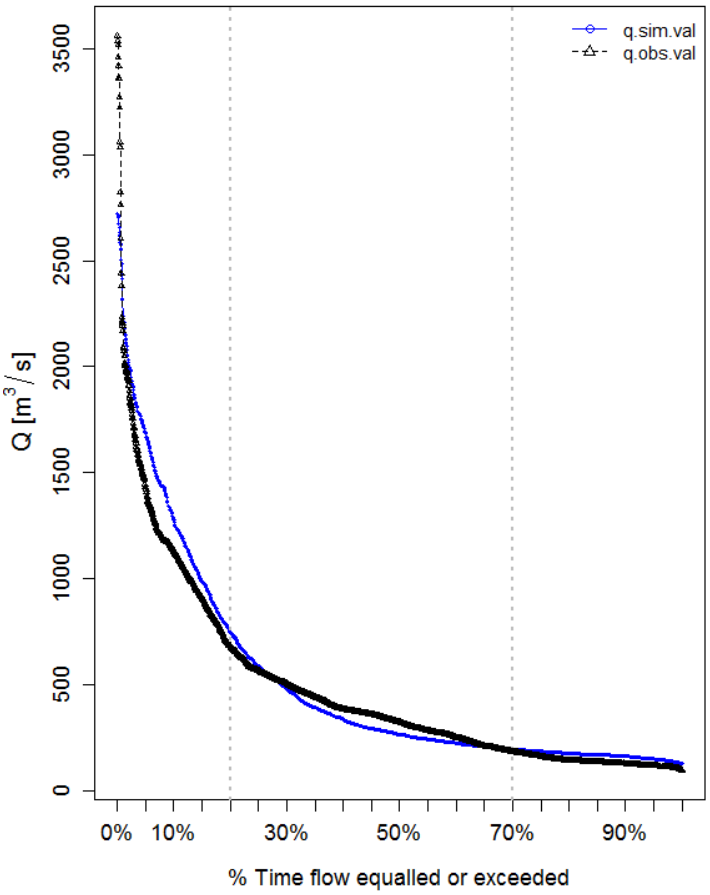
PBIAS	1.1%
R	0.89
NSE	0.78

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: DAUGAVPILS.
Catchment: Zap_Dvina_Daugava. River: DAUGAVA
Lat: 55.88. Lon: 26.53



Daily Flow Duration Curve
(2004-Jan-01 to 2009-Dec-31)



Internal Station ID	C438 (A05)
Upstream Area	64,700 [km2] (Lisflood)
	64,500 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.82

Validation Period
(2004-Jan-01 to 2009-Dec-31)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	96.6	127.1
Avg.	478.1	497.2
q95	1424	1684.7
q99	2094.8	2216.3
Max.	3561	2719.9

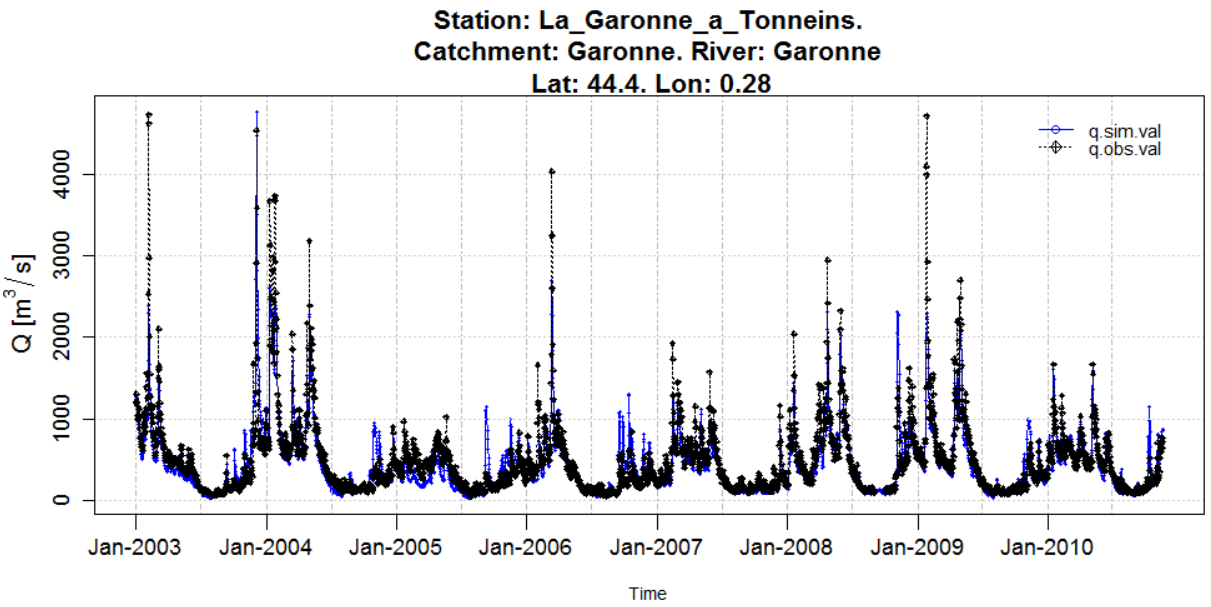
Error Measures, [m3/s]

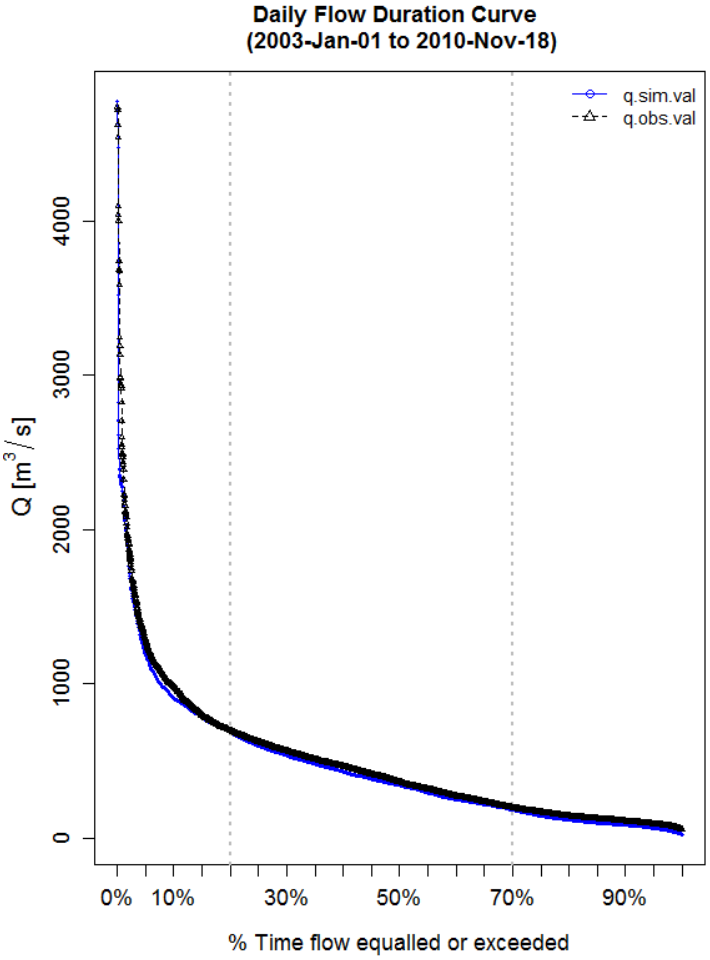
RMSE	239.9
MAE	161.2

Goodness-of-Fit

PBIAS	6.6%
R	0.88
NSE	0.74

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C069 (A03)
Upstream Area	50,425 [km2] (Lisflood)
	51,500 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.85

Validation Period
(2003-Jan-01 to 2010-Nov-18)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	53.8	20.3
Avg.	485.8	447
q95	1281	1185.9
q99	2453.7	2223.7
Max.	4735.8	4777.3

Error Measures, [m3/s]

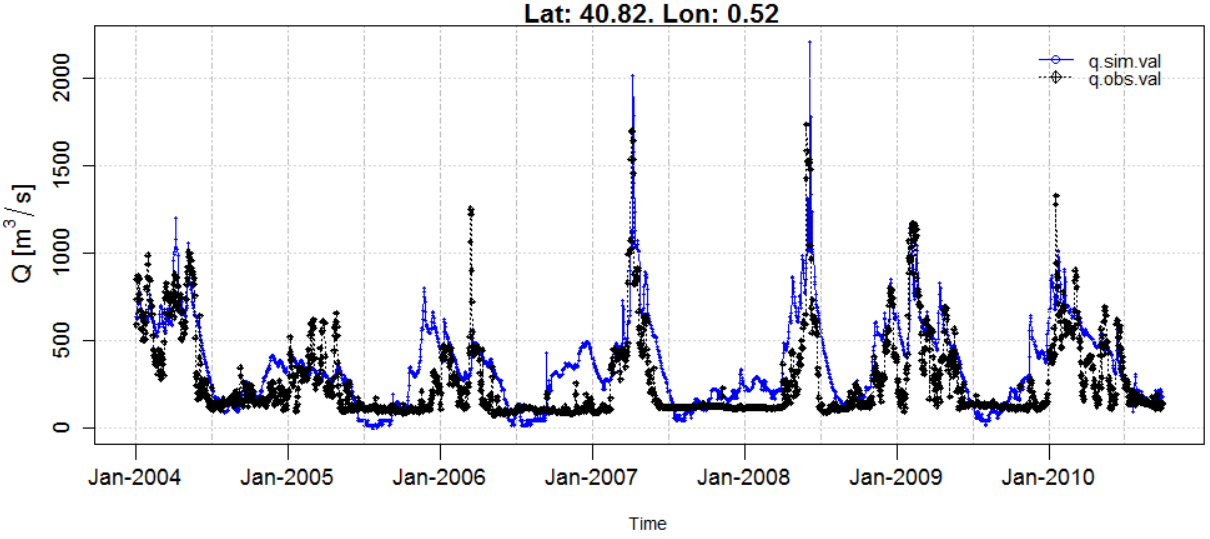
RMSE	213
MAE	122.9

Goodness-of-Fit

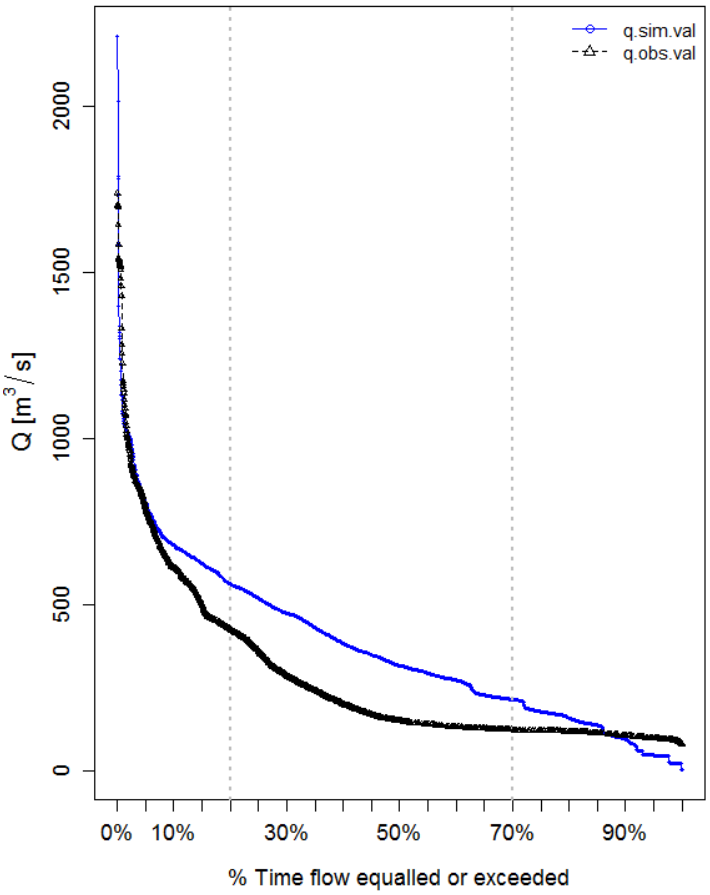
PBIAS	-7.2%
R	0.90
NSE	0.8

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: EbroEnTortosa.
Catchment: Ebro. River: Ebro
Lat: 40.82. Lon: 0.52



Daily Flow Duration Curve
(2004-Jan-01 to 2010-Sep-30)



Internal Station ID	C118 (A01)
Upstream Area	84,650 [km2] (Lisflood)
	84,230 [km2] (Provider)
Calibration Period	1996-Jan-01 to 2003-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.56

Validation Period
(2004-Jan-01 to 2010-Sep-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	77.3	1.2
Avg.	272.3	366.7
q95	784	805.3
q99	1168.3	1055.1
Max.	1737.6	2211.7

Error Measures, [m3/s]

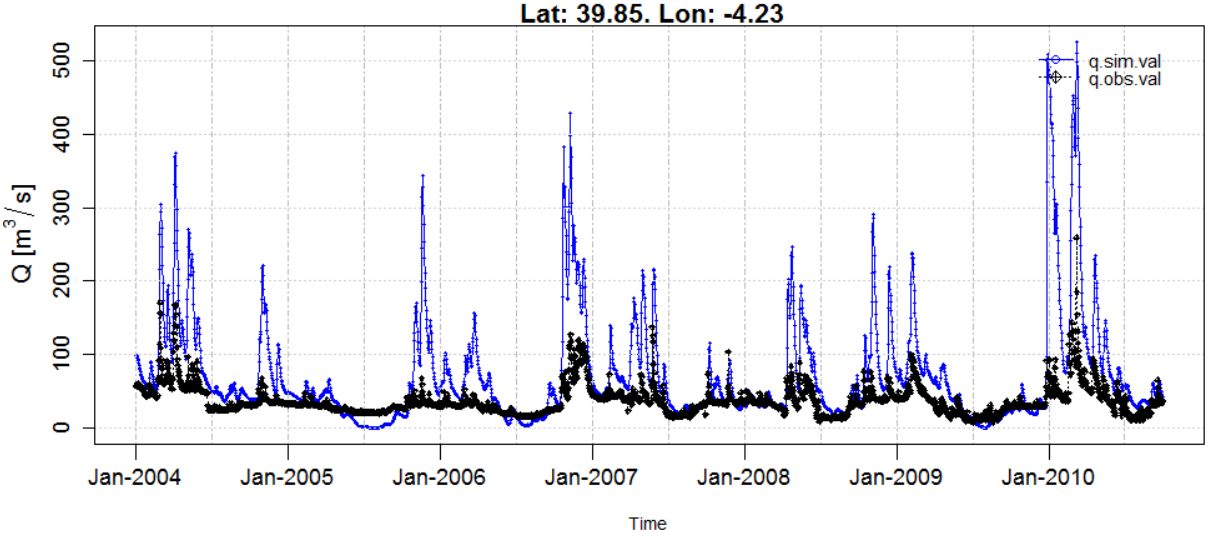
RMSE	193.2
MAE	147.4

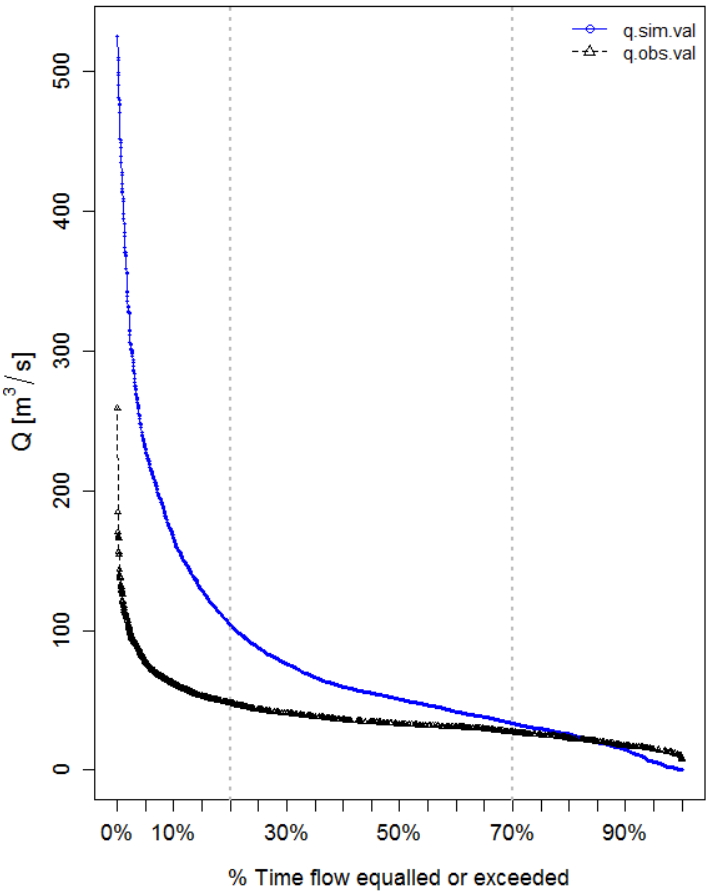
Goodness-of-Fit

PBIAS	34.6%
R	0.77
NSE	0.38

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: TajoEnLaPortusa.
Catchment: Tejo. River: Tajo
Lat: 39.85. Lon: -4.23





Internal Station ID	C147 (A01)
Upstream Area	26,950 [km2] (Lisflood)
	27,159 [km2] (Provider)
Calibration Period	1996-Jan-01 to 2003-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	< 0

Validation Period
(2004-Jan-01 to 2010-Sep-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	7.4	0.2
Avg.	37.9	74.1
q95	77	228.9
q99	118.8	401.7
Max.	258.8	525.3

Error Measures, [m3/s]

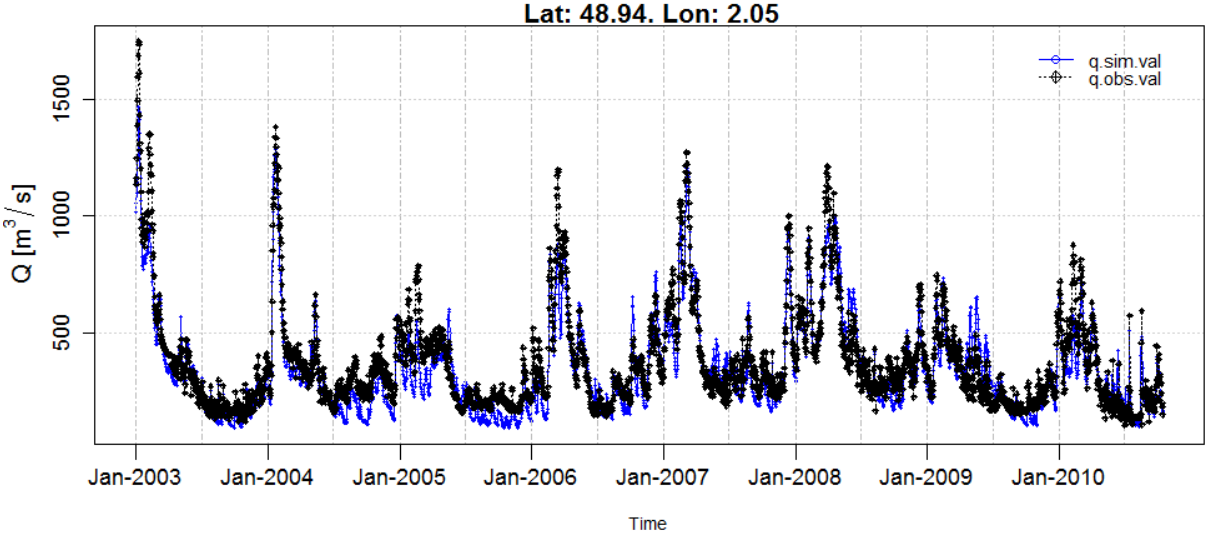
RMSE	71.4
MAE	40.5

Goodness-of-Fit

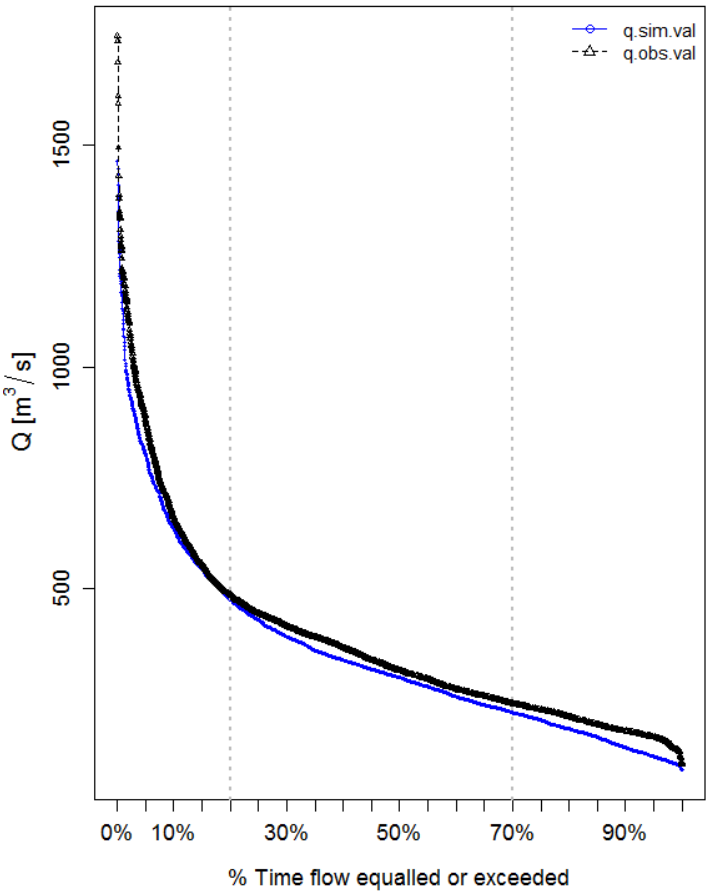
PBIAS	95.7%
R	0.76
NSE	< 0

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: La_Seine_A_Poissy.
Catchment: Seine. River: Seine
Lat: 48.94. Lon: 2.05



Daily Flow Duration Curve
(2003-Jan-01 to 2010-Oct-14)



Internal Station ID	C109 (A03)
Upstream Area	61,775 [km2] (Lisflood)
	61,820 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.91

Validation Period
(2003-Jan-01 to 2010-Oct-14)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	105	92
Avg.	380	350.5
q95	875.4	805.1
q99	1211.9	1101.8
Max.	1746.3	1465.3

Error Measures, [m3/s]

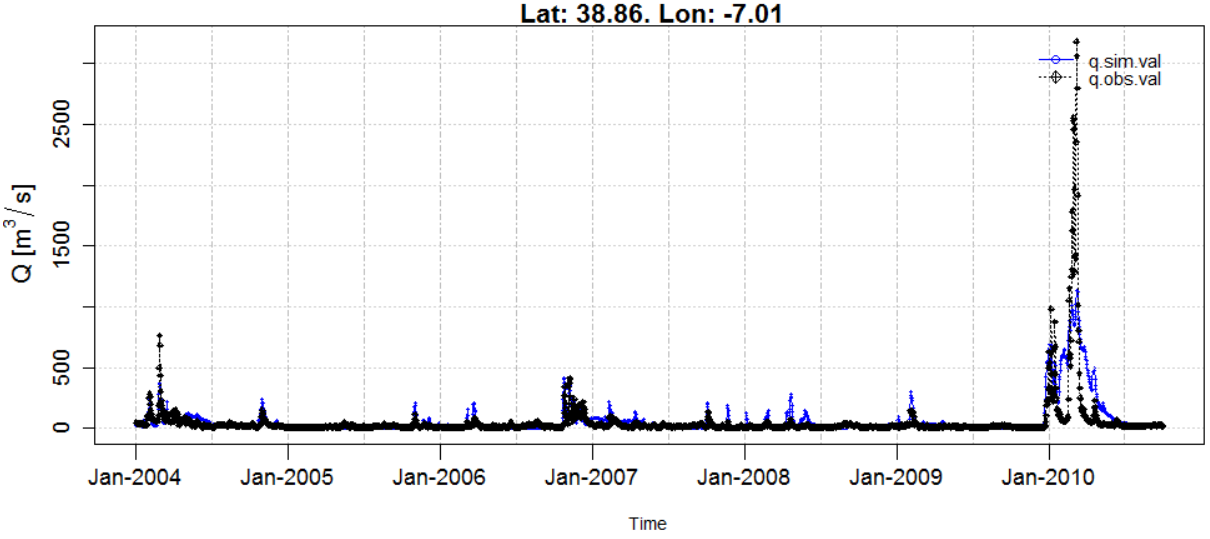
RMSE	91.7
MAE	70.5

Goodness-of-Fit

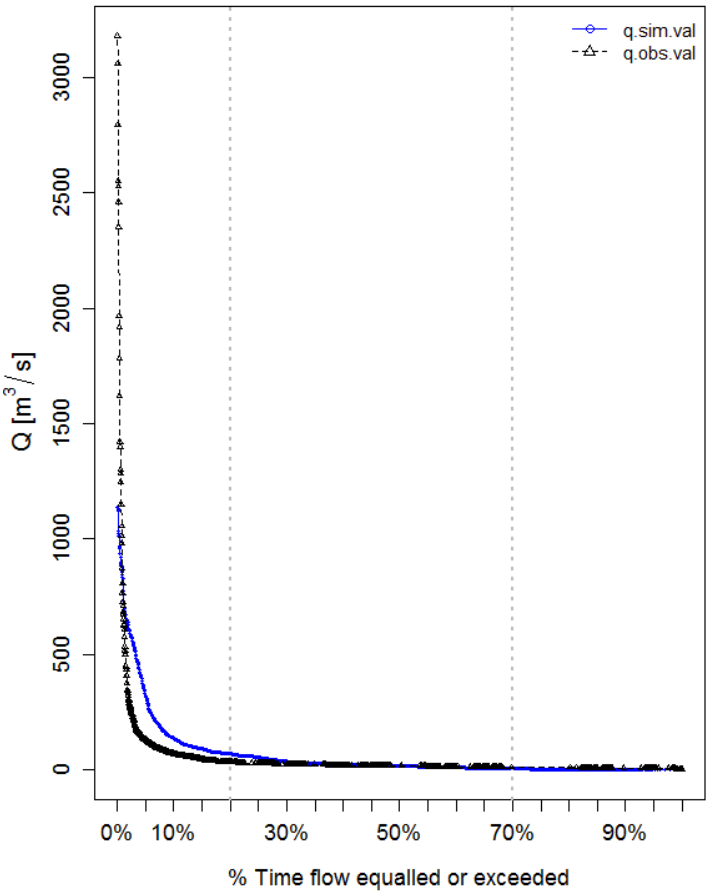
PBIAS	-7.8%
R	0.92
NSE	0.84

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency

Station: GuadianaEnAzudDeBadajoz.
Catchment: Guadiana. River: Guadiana
Lat: 38.86. Lon: -7.01



Daily Flow Duration Curve
(2004-Jan-01 to 2010-Sep-30)



Internal Station ID	C121 (A01)
Upstream Area	49,350 [km2] (Lisflood)
	48,530 [km2] (Provider)
Calibration Period	1996-Jan-01 to 2003-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.32

Validation Period
(2004-Jan-01 to 2010-Sep-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	3.2	0
Avg.	49.2	62.1
q95	130	318.6
q99	694.5	803.7
Max.	3180.1	1135.5

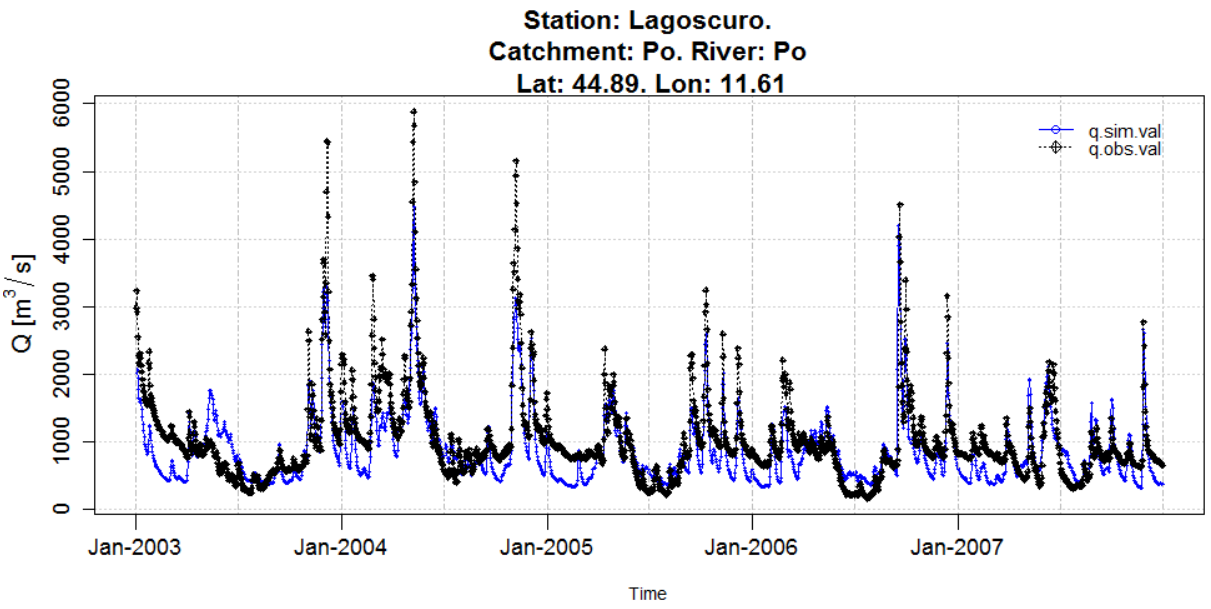
Error Measures, [m3/s]

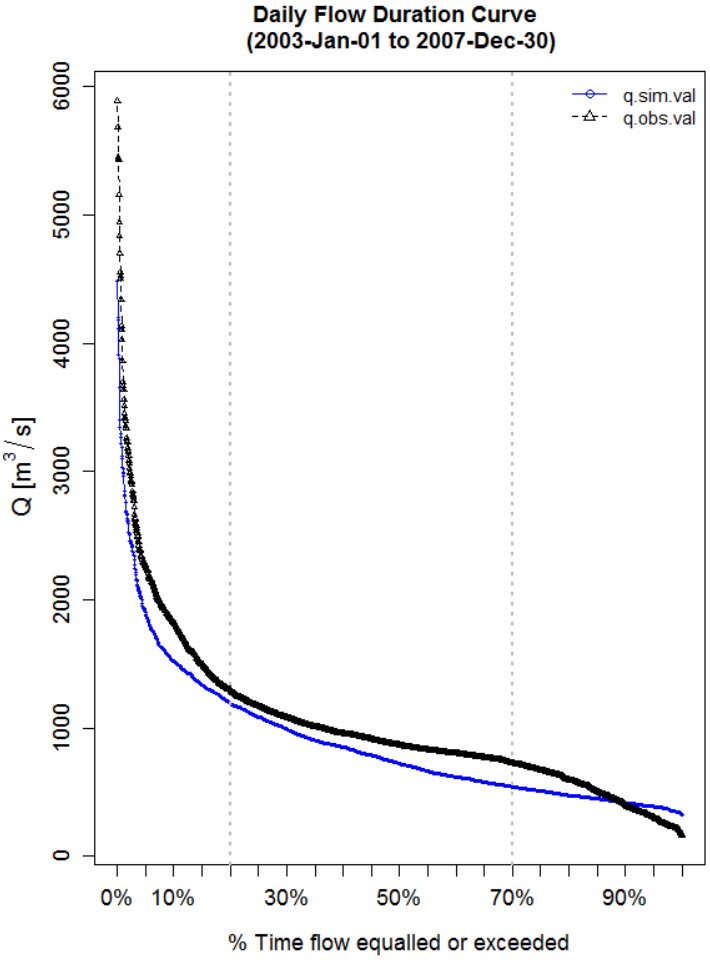
RMSE	130.4
MAE	43.3

Goodness-of-Fit

PBIAS	26.2%
R	0.72
NSE	0.52

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C140 (A03)
Upstream Area	71,875 [km2] (Lisflood)
	Not Available (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.72

Validation Period
(2003-Jan-01 to 2007-Dec-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	168.3	325.7
Avg.	1028.6	883.6
q95	2249.2	1878.1
q99	3658.4	2989.2
Max.	5885.3	4483.3

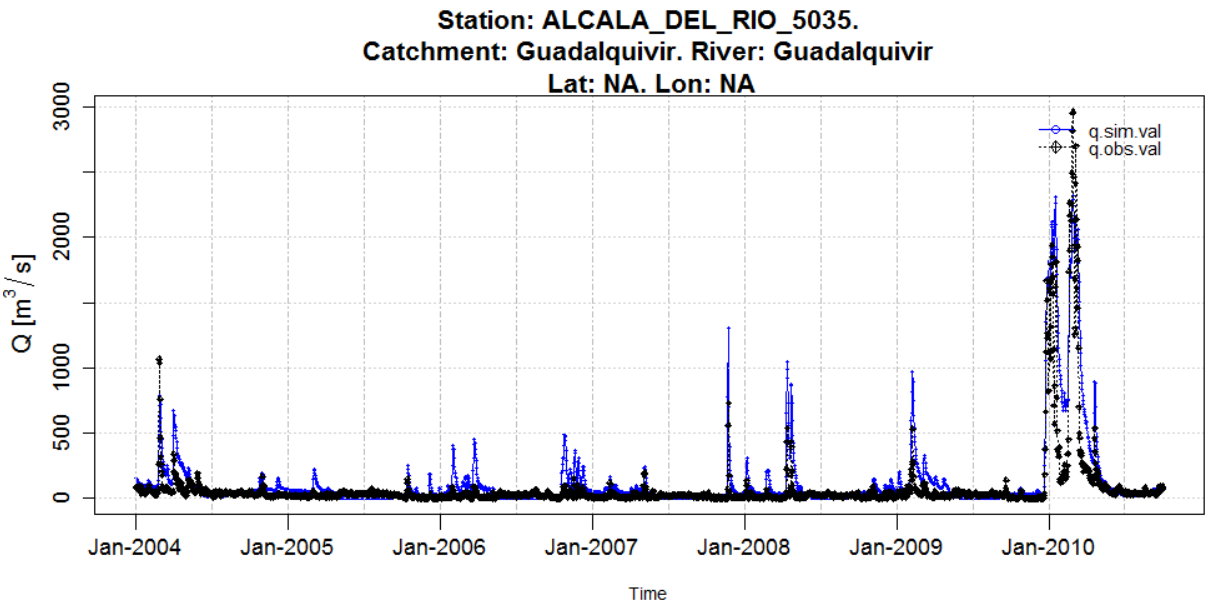
Error Measures, [m3/s]

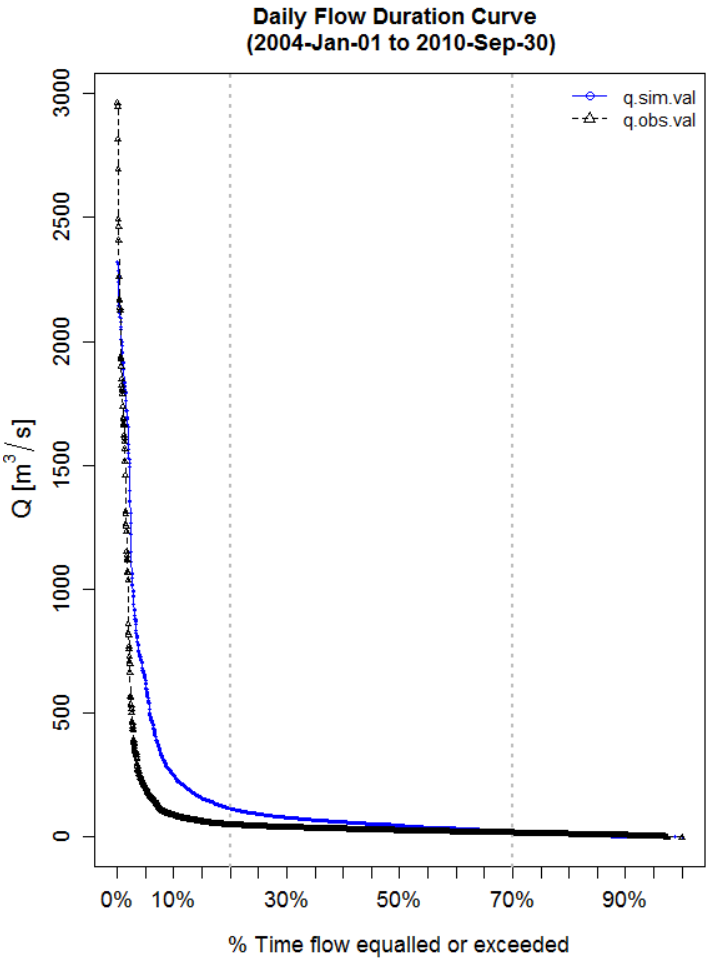
RMSE	382.8
MAE	292.1

Goodness-of-Fit

PBIAS	-14.1%
R	0.85
NSE	0.68

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	E512 (A01)
Upstream Area	48,675 [km2] (Lisflood) 197 [km2] (Provider)
Calibration Period	1996-Jan-01 to 2003-Dec-31
Calibrated Parameters	13
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.72

Validation Period
(2004-Jan-01 to 2010-Sep-30)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	0	0
Avg.	75.1	130.2
q95	199.3	612.9
q99	1688	1910.3
Max.	2962.3	2318.5

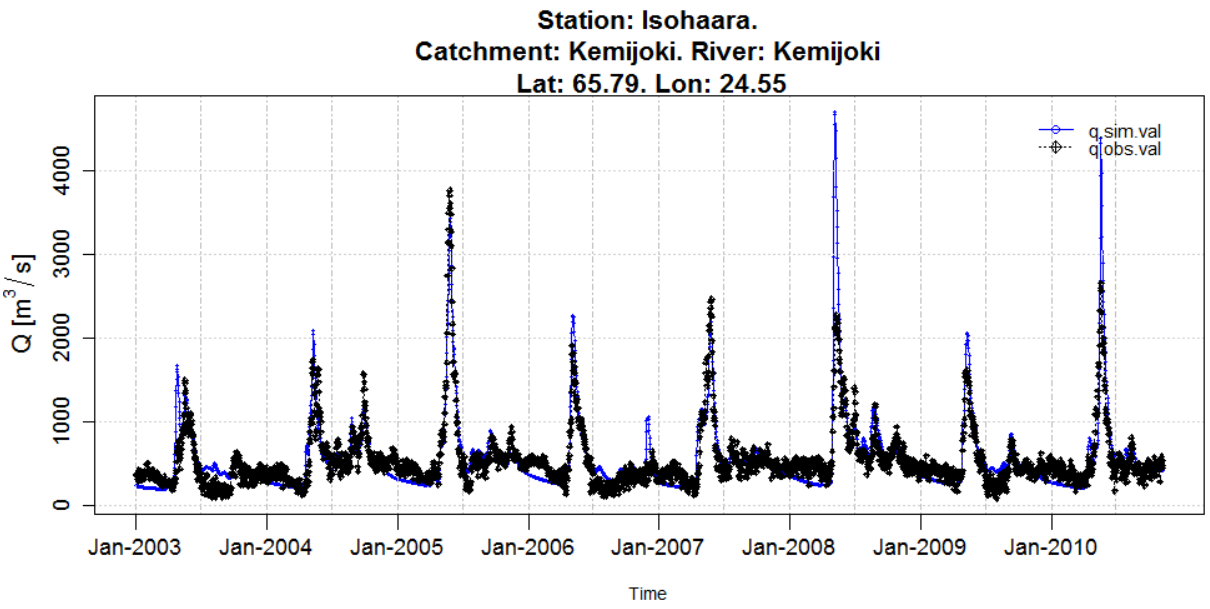
Error Measures, [m3/s]

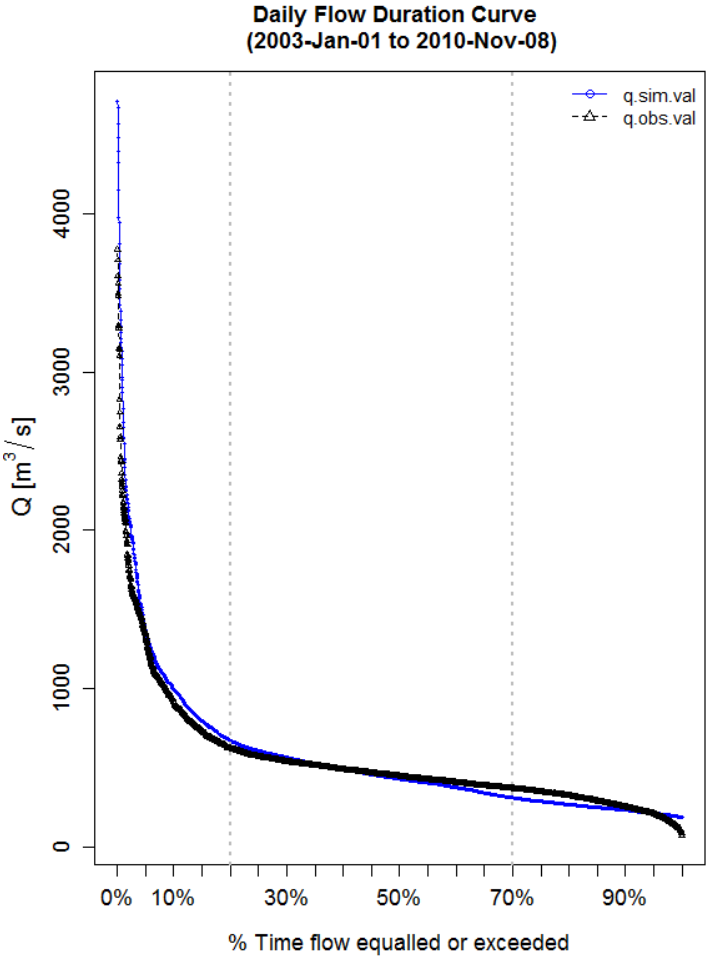
RMSE	151.4
MAE	75.4

Goodness-of-Fit

PBIAS	73.4%
R	0.89
NSE	0.66

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency





Internal Station ID	C241 (A07)
Upstream Area	52,700 [km2] (Lisflood)
	50,686 [km2] (Provider)
Calibration Period	1995-Jan-01 to 2002-Dec-31
Calibrated Parameters	9
Calibration Algorithm	Particle Swarm Optimisation
Calibration NSE	0.83

Validation Period
(2003-Jan-01 to 2010-Nov-08)

Summary Statistics, [m3/s]

	q.obs.val	q.sim.val
Min.	77.8	189.2
Avg.	544.2	558.4
q95	1342.8	1357.2
q99	2172.5	2712.2
Max.	3770	4710.4

Error Measures, [m3/s]

RMSE	234.1
MAE	146

Goodness-of-Fit

PBIAS	2.6%
R	0.87
NSE	0.64

RMSE: Root mean squared error
MAE : Mean absolute error
PBIAS: Percent bias
R : Pearson product-moment correlation coefficient
NSE : Nash-Sutcliffe efficiency