ISEIS

Journal of Environmental Informatics 32(2) 82-97 (2018)

Journal of Environmental Informatics

www.iseis.org/jei

The Hillslope Length Impact on SWAT Streamflow Prediction in Large Basins

A. Malagò[1] *, O. Vigiak[1], F. Bouraoui[1], L. Pagliero[2], and M. Franchini[3]

[1]*European Commission Joint Research Centre (JRC), Directorate-Sustainable Resources, via Enrico Fermi 2749, Ispra 21027, Italy*
[2]*Department of Civil Enginnering, Hydraulic Section, Katholieke Universiteit Leuven, Kasteelpark Arenberg 40, Leuven B-3001, Belgium*
[3]*Engineering Department, University of Ferrara, via Saragat 1, Ferrara 44100, Italy*

Received 22 May 2015; revised 23 February 2016; accepted 13 April 2016; published online 17 March 2018

ABSTRACT. The objective of this study was to assess the impact of hillslope length on Soil and Water Assessment Tool (SWAT) streamflow predictions in large basins using three methods for hillslope length calculation (the SWAT method, L1; a 3D analysis method, L2; a constant value, L3) combined with two DEMs (pixel size of 25 and 100 m), for a total of six DEML configurations that were tested in the Upper Danube (132000 km^2). The delineation of subbasins and HRUs were kept unchanged in all configurations, thus isolating the DEM impact on streamflow from that of subbasins delineation. The configurations were independently calibrated in 98 gauged stations located in headwater subbasins (period 1995 ~ 2006), and validated in 150 gauged stations (period 1995 ~ 2009). The analysis of streamflow prediction was extended to its components (surface runoff, lateral flow and baseflow) using performance criteria and residual analysis, and the comparison of different components of water yield was pursued. Calibration and validation showed that all configurations simulated monthly streamflow acceptably (PBIAS < 25% for more than 70% of 150 gauged stations). DEM pixel size had negligible effect of streamflow and its components. The default hillslope length (L1) resulted in large overestimations of lateral flow. L2 resulted in the best performance as well as L3 method. Given that L2 method takes into account the topographic convergence of flow, the configuration of DEM100 and L2 is recommended for SWAT application in large basins in order to obtain reliable streamflow predictions.

Keywords: SWAT, hillslope length, DEM, streamflow, water yields, multi-variables-site calibration

1. Introduction

A main challenge for distributed hydrological modelling at large scale concerns the identification of the best input datasets in order to predict hydrological processes and water quality accurately. Digital elevation models (DEMs) represent the topography that drives surface runoff, and are arguably one of the most important data input to several hydrological models (Wechsler, 2007). With regard to the eco-hydrological model Soil and Water Assessment Tool (SWAT; Arnold et al., 1998), a considerable amount of research has addressed the impact of DEM pixel size on model outputs (Cotter et al., 2003; Chaubey et al., 2005; Chaplot et al., 2005; Di Luzio et al., 2005; Dixon and Earls, 2009; Lin et al, 2013; Chaplot, 2014; Zhang et al., 2014). Generally, research demonstrated that streamflow was unaffected by DEM pixel size, nutrients were slightly affected, but sediment yield was greatly influenced by DEM pixel size (see for instance Chaplot 2005; Chaplot 2014; and Zhang et al., 2014). In most studies, the finer DEM was assumed to be the most representative and the most adequate to represents the streamflow, nutrients and sediment yields. Henceforth, calibration of SWAT was conducted on the smaller pixel size and kept unchanged after DEM resampling to larger sizes. This procedure showed that SWAT water quality is sensitive to delineation and number of model units (Chaplot, 2005; Zhang et al., 2014).

Conversely, the impact of DEM derivatives (topographic attributes) has been explored less frequently (i.e. Wu et al., 2008; Yao et al., 2010; Bieger et al., 2015), even tough a profound understanding of DEM derivatives is essential for assessing the relationships between watershed topography and hydrologic processes (Wu et al., 2008). Among several DEM derivatives, Zhang et al. (2014) showed that streamflow was most sensitive to the slope gradient. Sharma et al. (2014) explored the impact of DEM pixel size on DEM derivatives and on streamflow and sediment outputs. They concluded that DEM pixel size had a greater impact on sediment yields than on streamflow, but recommended to further investigate the impact of derivatives on the shape of hydrograph, peak runoff, and time to peak.

Among DEM derivatives, the hillslope length plays an important role in predicting streamflow and sediment yield through the peak runoff. Currently, the default method to set hillslope length in SWAT is based on a look-up table that relates hillslope length to the subbasin slope gradient, albeit generally the hillslope length is not only dependent on slope

* Corresponding author. Tel.: +39 0332 789640.
E-mail address: anna.malago@ec.europa.eu (A. Malagò).

ISSN: 1726-2135 print/1684-8799 online
 doi: 10.3808/jei.201700384

gradient (Hickey, 2000). Often the hillslope length is arbitrarily set to reflect field conditions or calibrated to achieve the desired spatial variation of streamflow components within a watershed (Bieger et al., 2015), sometimes overcoming the recommended range (values below 122 m) by Arnold et al. (2012a). For instance, Spruill et al. (2000) calibrated the streamflow in a small watershed (5.5 km^2) changing the hillslope length from 37 m (basic SWAT calculation) to 500 m as they observed a decrease of the absolute deviation between observed and simulated streamflow by increasing the hillslope length. In reviewing 64 SWAT watershed studies, Arnold et al. (2012b) pointed out that only in two of them hillslope length was adjusted for calibrating directly the sediment yields. At large scale, Rouholahnejad et al. (2014) calibrated the streamflow of 144 gauging stations in the Black Sea Basin (2.3 × 106 km^2) adjusting the SWAT hillslope length (derived from a 90 m DEM pixel size) in a range of ±40%, while Abbaspour et al. (2015) decreased the SWAT hillslope length and increased the slope (based on a resampled 700 m DEM pixel size) in order to calibrate the simulated hydrographs. Furthermore, Bonuma et al. (2014) decreased by 25% the default 50 m hillslope length to calibrate the sediment yields of the Arroio Lino watershed (4.8 km^2).

Over the last 20 years several Geographic Information System (GIS) procedures have been developed for estimating the hillslope length using DEM three-dimensional analysis (e.g. Zhang et al., 2013). Concerning the SWAT method, Kim et al. (2009) suggested modifying the SWAT GIS pre-analysis for assessing hillslope length, especially if average slope in the watershed is > 25%. Yao et al. (2010) assessed the effect of 10- and 30-m DEM pixel size on hillslope length estimations using two GIS-methods (Flow Length and Flow accumulation) to see which of them resulted in hillslope length most similar to observed field measures. The authors found out that both the DEM pixel size and the GIS-method had an influence on hillslope length, but with coarser DEM the Flow accumulation method produced higher errors than Flow Length method. Furthermore, the erosion predictions obtained using 10 m and 30 m DEMs pixel size with Flow Length method were slightly closer to the observations than those obtaining using Flow Accumulation method. More recently, Grieve et al. (2016) explored three techniques (slope-area analysis, inversion of drainage density and hilltop flow routing) for extracting hillslope length from high DEM resolution. The authors concluded that the hilltop flow routing produced more robust results (compared to field measurements) than the other investigated methods across a wide range of landscape morphologies.

However, all these studies were more oriented to define which method better reproduced the observed hillslope length rather than to assess the impact of hillslope length methods on model streamflow outputs and on hydrological processes representation.

Conversely, the aim of this study was to assess the impact of hillslope length estimation on streamflow and its components (surface runoff, lateral flow and baseflow). The assessment comprised three methods (L1, L2 and L3) to estimate hillslope length based on two DEMs of different resolution (25 and 100 m pixel size) for a total of six model configurations. The methods consisted of a slope-hillslope empirical relation (SWAT method; L1), a GIS-DEM flow accumulation algorithm (L2; Zhang et al., 2013), and the application of a constant value (50 m; L3). They were selected since they differ for complexity of implementation (L3 is the simplest one, followed by L1 and L2) and have different level of dependence with respect to the DEM pixel size (L2 is directly dependent, L1 is indirectly dependent, while L3 is independent).

The Upper Danube Basin (132000 km^2) was chosen as a study area because of its diversity in land cover and use, topography, and availability of gauging stations. The impact of hillslope length and DEM pixel size on streamflow and its components was evaluated using numerical, graphical and qualitative methods. With respect to previous studies (i.e. Zhang et al., 2014 and Chaplot, 2014), since it was demonstrated that the number of subbasins in the watershed remained unchanged when DEM pixel size is up to 300 m (Zhang et al., 2014), the subbasins delineation was kept constant to isolate the impact of DEM derivatives rather than on watershed delineation. Furthermore, differently from other studies (Chaplot, 2005; Lin et al., 2010; Zhang et al., 2014; Chaplot, 2014) where calibration was based on the finer DEM pixel size or even not performed, we considered that independent calibration of each configuration was essential to assess the impact of hillslope length on streamflow predictions: if independent calibration would generate similar predictions of streamflow and its components, then we could conclude that hillslope length impact was negligible.

2. Materials and Methods

2.1. Study Site

The study area is the Upper Danube Basin, which covers about 132000 km^2 across Austria, Germany, Czech Republic and Slovakia, extending from the Danube source down to the Gabcikovo Reservoir near Bratislava (Figure 1). The southern part of the Upper Danube Basin includes major parts of the Alps up to the watershed in the crystalline Central Alps and then the adjoining Swabian-Bavarian-Austrian foothills belt. The northern part of the basin is significantly smaller; it is confined by the heights of the Swabian and Franconian Alb, the Upper Falconian Forest, then parts of the Bavarian-Bohemian forests down to the Austrian Mühl and Waldviertel and the Bohemian-Moravian Upland (Schiller et al., 2010). The amount of precipitation in the Upper Danube Basin shows a distinct gradient with the altitude. It rises from 650 ~ 900 mm/y in the lowland areas to more than 3,000 mm/y in the high mountain exposed to the west and north (Rank et. al., 2005). Major land uses within the watershed are forest (38%) and pastures (18%) mainly extended in Alpine regions, and cropland (34%) in flat areas.

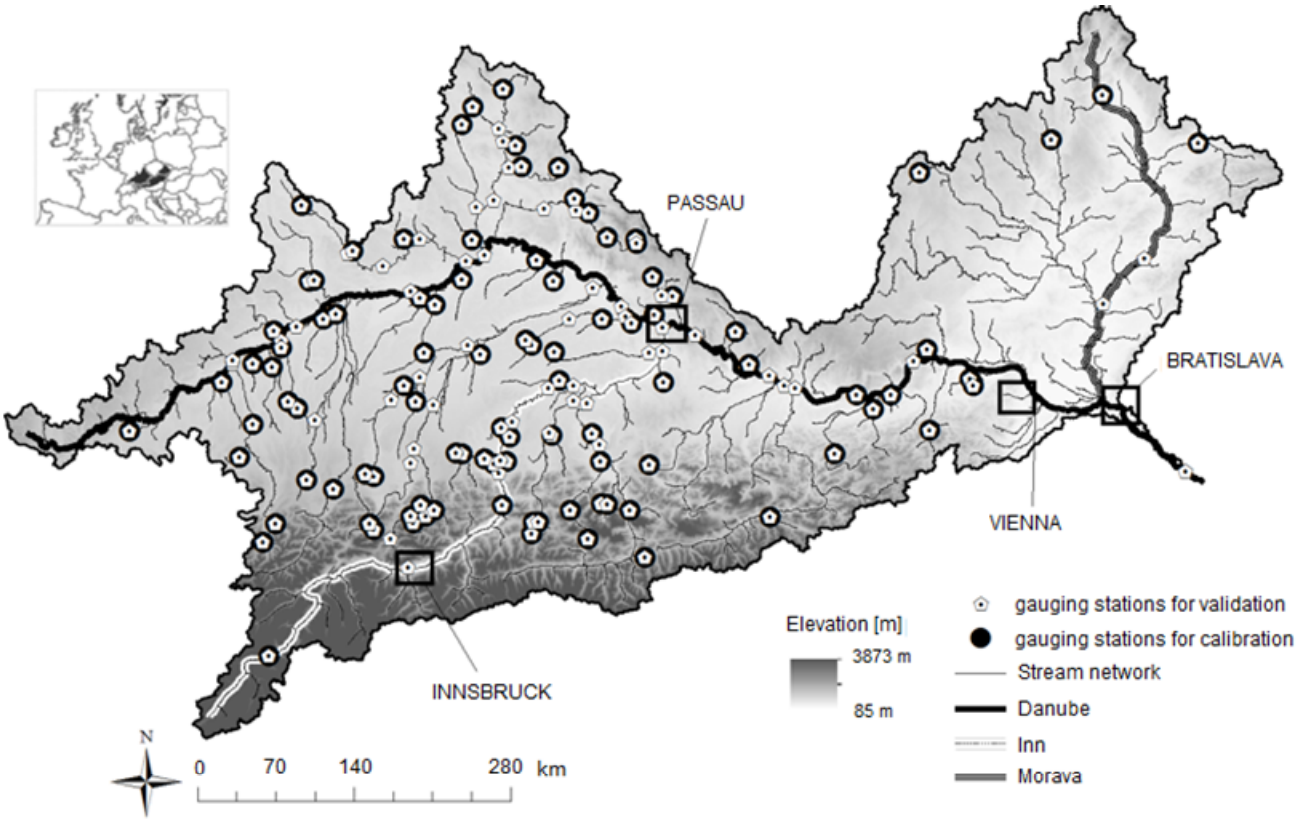

Figure 1. Overview of the Upper Danube Basin with location of gauging stations used for model calibration and validation.

The study area can be divided in three main sectors: the sector upstream of Passau (about 50000 km^2), the Inn river basin (about 25900 km^2) and the sector between Passau and the outlet near Bratislava (56100 km^2). The mean annual flow rate at Passau is around 670 m^3/s (35% of the flow rate at Vienna), with a maximum in March (880 m^3/s) and a minimum in October (520 m^3/s) (Rank et al., 2005). The mean altitude of this sector is relatively low (~ 800 m) compared to the Alpine regions (~ 1300 m). Groundwater discharge is the main mechanism forming baseflow in this part of the river. The period of high water is mainly controlled by precipitation and snow melting during late winter and early spring.

The Inn, which enters the Danube at Passau, has a hydrological regime typical of Alpine rivers and exhibits a nivo-glacial regime (Sommerwerk et al., 2009). The average discharge at its mouth is 732 m^3/s doubling the flow rate when it merges in the Danube. Maximum flow rates are observed in June/July (1200 m^3/s) and a minimum in January (400 m^3/s). In the Inn sector elevation ranges from 310 to 3,800 m; and mean annual precipitation ranges from 600 to more than 2000 mm/y (Parajka et al., 2007; Nester et al., 2011).

The sector between Passau and Gabcikovo reservoir is characterized by the influence of other Alpine rivers, with a mean annual flow rate of about 505 m^3/s (27% of the flow rate at Vienna) and by the Morava river with a mean annual discharge of 110 m^3/s (Sommerwerk et al., 2009). The Alpine rivers are characterized by a hydrological regime similar to that of the Inn, with maximum flow rates in May/June (800 m^3/s) and a minimum in January (270 m^3/s) (Rank et al., 2005), while the Morava is characterized by flow peaks in early summer (March/April) (Sommerwerk et al., 2009). Long-term mean annual runoff volume of the Danube at Bratislava is estimated around 64533 × 10^6 m^3 in the period 1876 ~ 2006 and the mean annual discharge is 2048 m^3/s (Pekárová et al., 2008).

2.2. SWAT Model

The Soil and Water Assessment Tool (SWAT; Arnold et al., 1998) is a physically-based, watershed-scale model developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields of a basin. The model is supported by online documentation (Neitsch et al., 2011; Arnold et al., 2012a), has a user-friendly GIS-embedded interface (e.g. Di Luzio et al., 2004; Olivera et al. 2006; Winchell et al., 2013), and is coupled to a semi-automatic calibration software (SWAT-CUP; Abbaspour, 2008). Some of the advantages of the model include: open access status of the source code, modelling of ungauged catchments and prediction of relative impacts of management practices, climate and land use change scenarios on water quality and quantity. In this study the version SWAT2012 (v.622) was used.

SWAT uses the DEM to subdivide the watershed into multiple subbasins with one main reach for each subbasin.

Subbasin topographic attributes, such as area, slope, elevation, are derived from DEM and assigned to all Hydrological Response Units (HRUs) within the subbasin. HRUs consist of homogeneous land use, management, topographical, and soil characteristics, and are the basic calculation unit. While a subbasin may be divided into more HRUs, in large basins a single HRU per subbasin (characterized by its dominant land use, soil type and management) is usually defined in order to reduce the computational burden of simulations (Pagliero et al., 2014; Malagò et al., 2015).

The structure of SWAT is divided in two phases: the land phase and the in-stream or routing phase. The land phase comprises the computation of HRU daily water balance, whereas the in-stream phase comprises the routing of water in each reach. The HRU daily water balance solves the change in soil water storage as a function of daily precipitation, surface runoff, evapotranspiration, infiltration in the vadose zone, and baseflow.

The HRU daily water yield is the total amount of water leaving the HRU and entering into main channel during the time step. It is estimated as:

$$WYLD = SR + LF + BF - TLOSS \tag{1}$$

where *WYLD* is the water yield (mm), *SR* is the surface runoff (mm), *LF* is the lateral flow contribution to stream flow (mm), *BF* is the baseflow contribution to stream flow from shallow aquifers (mm), and *TLOSS* is the bed transmission losses (mm) from the subbasin tributary channel. The accurate estimation of these components is of primary importance when assessing the impact of pollutant transport and for a sustainable water resources use. In the following the flow generated from HRUs is referred to as *WYLD*, whereas the flow routed in the reach (streamflow) is referred to as *Q*.

In this study, the surface runoff (*SR*, mm) was estimated using the SCS curve number method (USDA Soil Conservation Service, 1972) as modified by Williams (1985) to account for the impact of slope on the curve number. SWAT calculates *SR* according to the equation:

$$SR = \frac{(\mathrm{R}_i - I_a)^2}{(\mathrm{R}_i - I_a + S)} \tag{2}$$

where R_i is the rainfall for the day (mm), I_a is the initial abstraction which include surface storage, interception and infiltration prior to runoff (mm), and *S* is the retention parameter (mm) which is a function of the curve number (CN) for the day. The peak runoff rate (q_{peak}, m^3/s) is calculated with the modified rational formula:

$$q_{peak} = \frac{\alpha \cdot SR \cdot Area}{3 \cdot 6 \cdot t_{conc}} \tag{3}$$

where Area is the HRU area (km^2), t_{conc} is the time of concentration (h), and α is the fraction of daily runoff that occurs during the time of concentration. The time of concentration is calculated with a modified rational method (Chow et al., 1988) as:

$$t_{conc} = t_{ov} + t_{ch} \tag{4}$$

where t_{ov} is the overland flow time and t_{ch} is the channel flow time. Overland flow time is the time that water takes to travel from the furthest point in the sub-basin to a stream channel and it is computed as:

$$t_{ov} = \frac{L^{0.6} \cdot n^{0.6}}{18 \cdot slp^{0.3}} \tag{5}$$

where *L* is the average subbasin hillslope length (m), *slp* is the average subbasin slope (% or m/m), and *n* is Manning's roughness coefficient. The channel flow time is computed as:

$$t_{ch} = \frac{0.62 \cdot L_{ch} \cdot n^{0.75}}{Area^{0.125} \cdot slp_{ch}^{0.375}} \tag{6}$$

where L_{ch} is the channel length (km), *n* is Manning's roughness coefficient of the channel, and slp_{ch} is the channel slope (% or m/m).

The lateral flow (*LF*, mm) occurs whenever the water content of the soil exceeds its water content at field capacity and it is calculated as:

$$LF = 0.024\left(\frac{2 \cdot SW_{ly,excess} \cdot K_{sat} \cdot slp}{\Phi_d \cdot L}\right) \tag{7}$$

where $SW_{ly,\ excess}$ is the drainable volume of water stored in saturated zone of the HRU per unit area (mm), K_{sat} is the saturated hydraulic conductivity (mm/h), φ_d is the drainable porosity of the soil (mm/mm), and *L* is the hillslope length (m).

Multiplying the K_{sat} with *slp* gives the *LF* velocity at the HRU outlet (vlat; Neitsch et al., 2011):

$$\begin{aligned} v_{lat} &= K_{sat} \cdot \sin(\alpha_{hill}) \\ &= K_{sat} \cdot \tan(\alpha_{hill}) = K_{sat} \cdot slp \end{aligned} \tag{8}$$

where α_{hill} is the HRU gradient (degree). In SWAT it is assumed $\sin(\alpha_{hill}) \sim \tan(\alpha_{hill})$ to simplify the equation, thus *slp* is equivalent to $\tan(\alpha_{hill})$.

The estimation of the baseflow BF depends on the water recharge of shallow aquifers and the baseflow recession constant. Transmission losses TLOSS were assumed negligible. Equations (1 ~ 8) show that all streamflow components in SWAT are affected by DEM derivatives. For instance, lateral flow is strongly influenced by slope gradient and hillslope length, and the peak runoff rate is a function of hillslope length.

2.3. Estimating the Hillslope Length

According to Wishmeier and Smith (1978), in this study the hillslope length was defined as the horizontal distance from the point of origin of surface runoff to the point where runoff enters a river. The best estimates for hillslope length are obtained from field measurements (Yao et al., 2010), but these are not always available or practical, especially at watershed scale (Zhang et al., 2013). In a basin the hillslope length is very difficult to calculate and it is not solely dependent on slope gradient (Bieger et al., 2015).

In this study three methods for estimating hillslope length at basin scale were investigated. The first method (L1) is the current SWAT default method that defines the hillslope length based on slope gradient using Wishmeier and Smith (1978) look-up table. The slope gradient is derived from pre-analyses of the Digital Elevation Model (DEM) at subbasin level using the ArcSWAT interface (Winchell et al., 2013). L1 takes the maximum value of 122 m when the slope gradient is less than 2%, and decreases to 9 m when the slope gradient is more than 25%. However, the original look-up table was proposed for setting hillslope length in the case of contour support practices and may not be appropriate in the absence of these. The second method (L2) estimates hillslope length from a DEM flow accumulation analysis. The freeware LS-TOOL developed by Zhang et al. (2013) was used here. LS-TOOL calculates hillslope length through a step-process that requires setting an accumulated area threshold (As, m^2) above which water flow is channelized. As is a very sensitive parameter; Zhang et al. (2013) demonstrated that when increasing As, the hillslope length increases until a maximum value that depends on DEM pixel size, after which it remains quite constant. In our study, As was defined as the area for which the maximum possible hillslope length was 122 m (i.e. the DEM pixel size times 122 m). Hence, multiplying 122 m times the DEM pixel size, As was set to 12200 m^2 for the DEM100 and 3050 m^2 for the DEM25.

The third method (L3) is that suggested by Arnold et al. (2012a) in the presence of coarse or inaccurate DEM, i.e. setting hillslope length to 50 m (L3) everywhere. It is noteworthy that the representativeness of the values of hillslope length derived for the proposed method is not object of this study due to the fact that not field data were available for the comparison.

2.4. SWAT Model Setup

The first two hillslope length methods (L1 and L2) are strongly dependent on the accuracy of the DEM, hence in this study two DEMs of different resolution were investigated: the CCM2 DEM (Vogt et al., 2007) with 100 m pixel size, and the EU-DEM (EU-DEM Metadata, 2013) with 25 m pixel size. The CCM2 DEM is a pan-European raster derived from the 3 arc-second digital elevation model from the Shuttle Radar Topography Mission (SRTM Version 4.1, about 90 m pixel size). Vogt et al. (2007) describe in detail its generation and accuracy. The EU-DEM is a continent-wide fusion dataset that covers the European Union (Bashfield and Keim, 2011; available online at the European Environmental Agency website http://www.eea.europa.eu/data-and-maps/data/eu-dem). The EU-DEM is a 25 m pixel size digital elevation model created by an automated data fusion of improved ASTER GDEM (Advanced Spaceborne Thermal Emission and Reflection Radiometer, ASTER GDEM database) with STRM data. In the following, the abbreviations DEM100 and DEM25 will be used for CCM2 DEM and EU-DEM respectively.

The combination of three hillslope length methods with two DEMs resulted in six model different configurations (called DEML configurations; Table 1). The six configurations differed only for hillslope length and DEM pixel size, while all others input were kept constant, including the subbasins and streams delineation and the definition of HRUs. By maintaining the same unit definitions in the SWAT configurations, the impact of DEM resolution on unit definition was avoided, and the impact of hillslope length could be better isolated. Subbasins and stream delineation was based on the Catchment Characterization Modelling version 2 database (CCM2, Vogt et al., 2007) for continental Europe, subdividing the basin in 753 subbasins with average area of 180 km^2.

The landuse was defined using a map of 1 × 1 km for year 2000, built from the combination of CAPRI (Britz, 2004), SAGE (Monfreda et al., 2008), HYDE 3 (Klein Goldewijk and Van Drecht, 2006) and GLC2000 (Bartholome and Belward, 2005) databases. The soil map of 1 × 1 km was obtained from the Harmonized World Soil Database (HWSD) (FAO, 2008), using top soil layer data. Except for urban areas, the combination of dominant land use and soil was used to define a single HRU per subbasin. Urban areas were considered hydrological isolated and defined as separate HRUs. In total 822 HRUs were defined in all DEML configurations.

Table 1. Characteristics of the Six SWAT Configurations Compared in this Study (DEML configurations) as Generated by the Combination of Two Digital Elevation Models (DEM) with Three Hillslope Length Methods

DEML configurations	DEM pixel size (m)	Hillslope length method	Hillslope length range (m)	Description
25L1	25	L1	9 - 122	SWAT default method based on slope calculated from DEM 25 m pixel size
25L2	25	L2	20 - 45	Application of LS-TOOL (Zhang et al., 2013) using a DEM 25 m pixel size
25L3	25	L3	50	Hillslope length fixed to 50 m
100L1	100	L1	9 - 122	SWAT default method based on slope calculated from DEM 100 m pixel size
100L2	100	L2	50 - 64	Application of LS-TOOL (Zhang et al., 2013) using a DEM 100 m pixel size
100L3	100	L3	50	Hillslope length fixed to 50 m

*L1 = SWAT default method based on look-up slope table; L2 = flow accumulation method; L3 = constant hillslope length.

Management practices for each crop included planting, fertilization, irrigation and harvesting. The timing of management actions was implemented through daily heat unit method (Arnold et al., 1998). In this study the heat units for each crop were calculated by Bouraoui and Aloe (2007) using the PHU (Potential Heat Units) program (PHU, 2007), developed at Texas Agricultural Experiment Station, while the amount of manure and mineral fertilization applied was retrieved from the CAPRI model (Britz, 2004). The auto-irrigation was selected for irrigated areas based on the MIRCA database (Portmann et al., 2008).

Reservoirs and lakes exceeding 20 km^2 (Lehner and Döll, 2004; Vogt et al., 2007) and hydropower plants of large generation capacity (> 10MW; ICPDR, 2013) installed on the main rivers were included in the model. Reservoir outflow rates were modelled using the average release for uncontrolled reservoirs method (Neitsch et al., 2011) and their volumes were set according to Lehner and Döll (2004) and Vogt et al. (2007).

The climate data including daily precipitation, temperature, solar radiation, wind speed and relative humidity were obtained from EFAS-METEO at spatial resolution of 5 × 5 km (Ntegeka et al., 2013). To account for the increase in precipitation with elevation that is typically observed in mountainous regions, four elevation bands were implemented. The 20-year climate record (1990 ~ 2009) used for the study was divided into three parts. The first five years (1990 ~ 1994) were used as a model "warm up" period to allow the model parameters to adjust to the watershed characteristics. Calibration of streamflow was performed for 1995 ~ 2006 in 98 gauged stations. Validation was performed for the period 1995 ~ 2009 for the entire dataset of gauged stations (150 gauged stations; Figure 1).

2.5. Streamflow Calibration and Validation

Daily streamflow data were collected from different sources, including the Global Runoff Data Centre (GRDC, 2010), and various national environmental agencies. The streamflow was calibrated independently for each configuration following a step-wise calibration procedure that involved sensitivity analysis, multi-variable calibration of headwater subbasins, and regionalization of the calibrated parameters (Pagliero et al., 2014; Malagò et al., 2015). Only gauged stations in headwater subbasins were used in the calibration because headwater subbasins are more likely to represent natural hydrological behaviour (Gudmundsson et al., 2012; Malagò et al., 2015) and the streamflow components are more representative than in larger basins where streamflow is often influenced by human activities (Döll et al., 2008). The daily streamflow of headwaters subbasins was divided into its main components (surface runoff SR, lateral flow LF and baseflow BF) using the SWAT filter (Lyne and Hollink, 1979).

The streamflow components of the headwater subbasins were calibrated separately using SUFI-2 (Abbaspour, 2008) method in four sequential steps that focused on different hydrological processes: snow processes, surface runoff, lateral flow, and baseflow. The steps, parameters, and performance criteria used at each step are described in detail in Malagò et al., (2015). The procedure was slightly modified by including the parameter Manning's roughness for tributary channel (CH N1) in the calibration of the surface runoff. In addition, a fifth final step of calibration was added by calibrating all hydrological parameters in a reduced range to account for any covariance of parameters belonging to different hydrological groups.

After calibration, subbasins where calibration of monthly streamflow reached "acceptable performance" of monthly streamflow simulation were selected as donors for use in the parameter regionalization. "Acceptable performance" for donors was defined if the percent bias (PBIAS, Gupta et al., 1999) was in the range $\pm$ 25% and the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) was > 0. Note that the NSE threshold was set low as to keep a sufficient number of well distributed donors, which is important to ensure robust regionalization (Harvey et al., 2012). Instead, in the overall evaluation of the calibration and validation of model performances, the satisfactory performance criteria were set to the recommended statistics of PBIAS $\pm$ 25% and NSE > 0.5 (Moriasi et al., 2007).

Once the donors were selected, a regionalization technique coupled with a classification procedure based on a similarity approach was applied. The similarity approach is based on the assumption that similar catchments behave hydrologically similarly. The regionalization consisted in transferring the calibrated parameter sets (Near Optimal Parameter set, NOP; Malagò et al. 2015) from donors to their hydrologically similar receptor subbasins.

The definition of the similarity measure is subjective and conditions the success of the regionalization (Heuvelmans et al., 2006). The regionalization was performed using the Partial Least Squares Regression method (PLSR; Wold, 1966; Geladi and Kowalski, 1986) that allows identification of similar subbasins based on the correlation between the watershed characteristics and the streamflow characteristics. In total, 19 independent variables representing the subbasin character istics ("c" matrix of PLRS regressors) and 14 dependent streamflow variables ("q" matrix of PLRS responses) were used. Malagò et al. (2015) provides a detailed description of all the variables. The PLSR analysis defined latent variables to identify "hydrological regions" using the Ward's minimum variance linkage method (Ward, 1963) together with the Euclidean distance similarity. To find the best number of "hydrological regions" two index were used: the corrected Rand index (Hubert and Arabie, 1985) and the Meilă index (Meilă, 2007). Finally, in each hydrological region a classification procedure based on a supervising clustering approach (k-NN method; Dettling and Maechler, 2012) was performed. Each subbasin in a region that was classified hydrologically similar to a given donor pertaining to the same hydrological region received the donor NOP set. The regionalization and classification analysis were performed using the statistical software R (R Development Core Team, 2008), using

the packages (collection of mathematical functions) "pls" (Mevic and Wehrens 2007), "fpc" (Henning, 2010), and "supclust" (Dettling and Maechler 2012).

The accuracy of model performance after calibration and regionalization-classification was evaluated by comparing the calibrated monthly streamflow with observed data using statistic performance criteria. Percent bias (PBIAS %) and Nash and Sutcliffe coefficient were calculated using the R package "hydroGOF" (Zambrano-Bigiarini, 2012). However, since these criteria measures only specific aspects of a model's performance, box and whiskers plots, visual appraisal of time-series, and residuals analysis (simulation-observation) were also used (Harmel et al., 2014; Bieger et al., 2012).

The comparative assessment of six DEML configurations was conducted on different aspects of the modelling: descriptive analysis of topographic characteristics, assessment of models calibration and validation, and evaluation of water yield components.

3. Results and Discussion

3.1. DEM Derivatives of Six DEML Configurations

Differences in mean elevation, slope gradient, and hillslope length of the six DEML configurations are presented in Figure 2. The two DEMs did not produce significant differences in the mean altitude of subbasins (Figure 2a), but produced small differences in slope (Figure 2b). The DEM100 resulted in lower slopes than the DEM25 confirming the results of other studies (e.g. Chaplot et al., 2005; Lin et al., 2013). The hillslope length L1 (25L1 and 100L1, Figure 2 c) increaseed step-wise until 122 m with decreasing slope. L2 markedly changed from the finest to the coarsest DEM. In the 25L2 configuration, hillslope length was around 30 m whereas in 100L2 it was around 60 m, with no clear relationship with slope gradient. Hence, the coarser DEM resulted in lower slopes and longer hillslope lengths (Lin et al., 2013). It is noteworthy that the constant hillslope length of 50 m (L3) was in between of 25L2 and 100L2 hillslope lengths. Furthermore, the distribution of hillslope length showed that L1 was also strongly correlated with elevation. Figure 2 d shows that L1 decreased with increasing elevation and at elevation higher than 1000 m L1 were shorter (9 m). Conversely, in L2 and L3 configurations hillslope length was not correlated with elevation or slope.

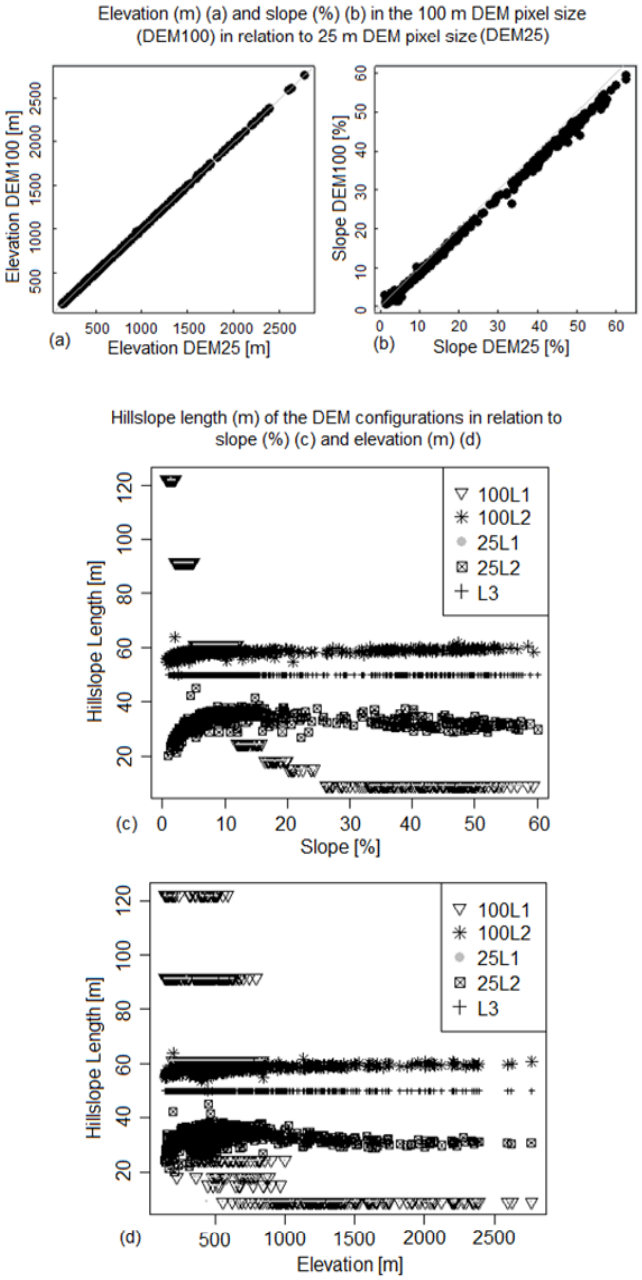

Figure 2. Relationships between the subbasin main DEM derivatives in the SWAT configurations.

3.2. Results of Model Calibration and Validation

The calibration of monthly streamflow yielded satisfactory PBIAS (PBIAS in the range of ±25%) in more than 50% of calibration headwaters in all DEML configurations (Table 2). The configurations 25L2 and 25L3 reached 61% and 62% respectively of calibrated gauged stations with satisfactory PBIAS, followed by 100L2 and 100L3 with 59%, while 25L1 and 100L1 reached the lowest percentage (56% and 57% respectively).

The median of PBIAS for the satisfactory gauged stations was in the range of -4% to -2% for all configurations, indicating a negligible underestimation of streamflow, except for 25L1 for which no bias was observed. NSE was satisfactory (NSE > 0.5) in 30% of cases; however, the median of NSE of these 30% calibrated station was close to 0.7.

The performances of monthly streamflow for the 150 validation gauged stations were better than for the calibration dataset. More than 70% of validation gauged stations reached satisfactory PBIAS in each configuration, and around 40% of the stations had NSE > 0.5 (Table 2). The highest number of subbasins with satisfactory streamflow predictions were obtained for 25L2 (73% of gauged subbasins with median PBIAS of -1.7% and 45% of gauged subbasins with median NSE close

Table 2. Summary of Calibration and Validation of Monthly Streamflow Simulations in the Upper Danube Basin for each DEML Configuration

		25L1	25L2	25L3	100L1	100L2	100L3
Calibration (1995 ~ 2006; #98 gauged stations)	Percentage of gauged stations with satisfactory PBIAS* (%)	56	61	62	57	59	59
	Median PBIAS (%) for satisfactory stations	0	-2.05	-3.6	-3.9	-2.35	-3.4
	Percentage of gauged stations with satisfactory NSE** (%)	31	32	32	31	32	31
	Median of NSE of satisfactory stations	0.68	0.67	0.67	0.66	0.66	0.69
Validation (1995 ~ 2009; #150 gauged stations)	Percentage of gauged stations with satisfactory PBIAS* (%)	70	73	72	72	72	71
	Median PBIAS (%) for satisfactory stations	-0.1	-1.7	-3.75	-2.6	-2.1	-3.3
	Percentage of gauged stations with satisfactory NSE** (%)	42	45	45	37	44	44
	Median of NSE of satisfactory stations	0.64	0.67	0.68	0.67	0.67	0.69

*Model simulation can be judged as satisfactory if PBIAS ± 25% for monthly streamflow (Moriasi et al., 2007).
**Model simulation can be judged as satisfactory if NSE >0.5 for monthly streamflow (Moriasi et al., 2007).

to 0.7) followed by 25L3 and 100L2, while 25L1 had the lowest numbers, but with the median PBIAS close to 0.

The analysis of monthly percentage residuals (simulation-observation/observation) confirmed that all configurations provide good simulations of streamflow with median residual values close to 10% in the calibration (Figure 3a) dataset and absence of residuals for 50 percentile of gauging stations in the validation dataset (Figure 3b). The interquartile range of monthly residuals for all configurations was within the interval -30% ~ 60% in the calibration dataset, indicating a slight tendency to overestimation of streamflow. The interquartile ranges were smaller, in the interval -30% ~ 40% in the validation dataset, for which the distribution of residuals was broadly symmetric around zero.

Spearman's rank correlation coefficients of the monthly streamflow residuals among the six configurations were all larger than 0.9, indicating that streamflow of DEML configurations were highly correlated. The lowest correlation were found between the configurations 100L1 and 25L3 (~ 0.93) and 100L1 and 100L3 (~ 0.93) in the calibration dataset, and between the configurations 25L1 and 100L3 (~ 0.91) and 25L1 and 25L3 (~ 0.93) in the validation one. These findings reflect the differences in the distributions of hillslope length L1 and L3. However, generally DEM pixel size and the hillslope length had a negligible impact on streamflow predictions and the spatial distribution of streamflow PBIAS showed no evident spatial pattern (e.g. Figure S1 for configuration 25L2).

Generally, low performances were associated with reaches where the hydrology was strongly influenced by human activities, in particular by diversion and regulation of streamflow, which were not accounted in SWAT, and are wide spread present in the study area. Mostly, barrages are built as cascades along river courses and more than 33 barrages are located along the Danube until Vienna, on the tributary rivers Iller, Lech, Isar, Inn, Salzach, and Enns. Additionally, there are a lot of deviations, diversions and intakes (Schiller et al., 2010).

Figure 4 shows the monthly time-series of SWAT simulations versus observations for two validation dataset stations, at the outlet of the Inn River and at the outlet of the entire Upper Danube basin, after the Gabcikovo reservoirs (locations are indicated in Figure S1). Although some discrepancies could be detected, visual appraisal of the time-series indicates good correlation between streamflow simulations and observations.

3.3. Analysis of Behavioural Parameter Sets

Figure S2 shows the distributions of NOP sets of donors for each configuration and indicates the hydrological processes upon which they were calibrated in the step-wise calibration (snow processes, step [1]; surface runoff, step [2]; lateral flow, step [3]; and baseflow, step [4]).

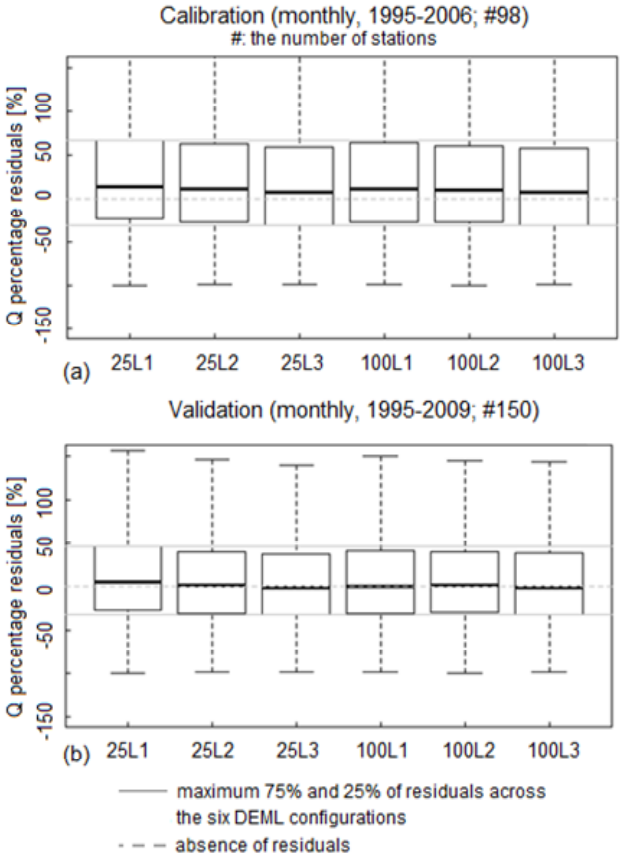

Figure 3. Box-and-whisker plots of percentage residuals between simulated and observed streamflow (Q) of the six DEML configurations for the calibration (a) and validation (b) dataset.

For the snow process parameters (parameters identified with [1] in Figure S2), all DEML configurations were quite similar, except for the snowfall (SMTMP, °C), the snowmelt

temperatures (SFTMP, °C) and the precipitation lapse rate (PLAPS). SFTMP and SMTMP were generally lower and had a larger range in 25L2 (with a median around -0.3 °C for SFTMP and 1.3 °C for SMTMP). The highest median value for SFTMP and SMTMP were obtained for 25L3 with 0.63 °C and 2 °C respectively. The snow pack temperature lag factor (TIMP, adimensional) was slightly different across the six configurations, with minimum median values of 0.24 (100L1) and maximum of 0.43 (25L2). The temperature lapse rate (TLAPS, °C) was used to adjust temperature for elevations band in each subbasin; in all DEML configurations the median TLAPS was around -5 °C indicating a decrease of temperature around 5 °C with a 1 km increment of elevation. The increment of precipitation per km of elevation (PLAPS, mm/km) was higher for configuration 25L2 and 100L1 (around 65 mm/km of median value) than the others (around 50 mm/km of median value). The median initial snow content (SNOEB) in each elevation band ranged between 112 mm for configurations 100L1 and 160 mm for 100L2.

The surface runoff parameters (parameters identified with [2] in Figure S2) did not greatly differ among the DEML configurations. However, CN2 adjustment (rate value) had a larger range in 100L1 with a median relative increase of 12%, whereas in the other configurations CN2 increase was around 14%. The median of CH N1 slightly increased from L1 to L3 both for DEM25 (from 0.11 to 0.13), while for DEM100 is about 0.12. The melt factor for snow on December (SMFMN) was more sensitive than the melt factor on June (SMFMX) and configuration 100L1 and 25L2 had the highest and lowest median of 4.7 and 2.9 respectively. It is noteworthy that these parameters regulate the melt factor at daily time step, so should belong to snow processes group. However, their effect is to change the position in time of runoff hydrograph, so they were included in the "surface runoff step of step-wise calibration" (see Malagò et al., 2015) because that is the phase where runoff hydrograph is calibrated at daily time step.

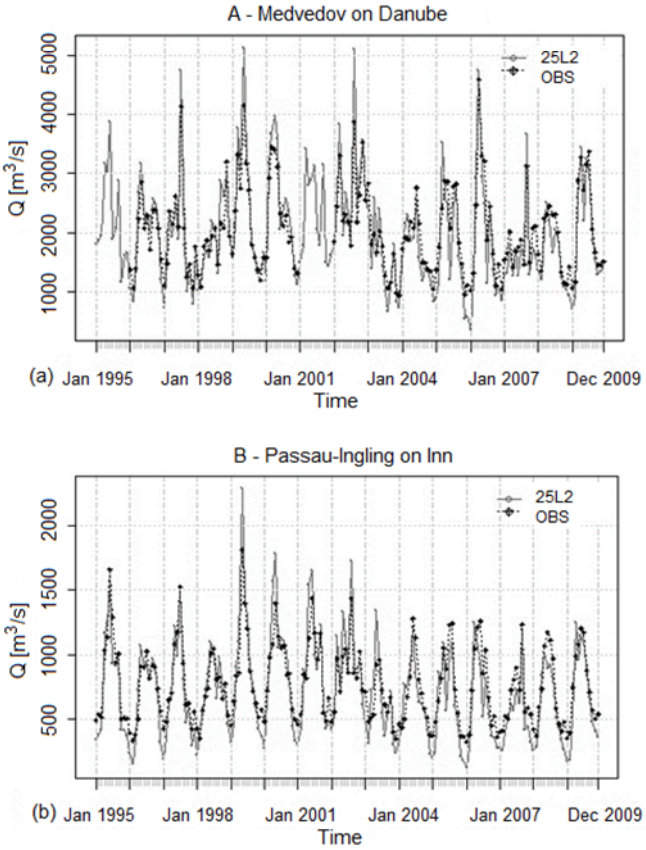

Figure 4. Comparison of SWAT outputs for 25L2 with observations (OBS) at two selected monitoring gauging stations (locations A-B shown in Figure S1).

In the calibration of lateral flow (see the parameters identified with [3] in Figure S2), the soil plant compensation factor (EPCO) and soil evaporation compensation factor (ESCO) had different ranges between the model configurations, implying differences in the water balance yield component estimations. The EPCO regulates water uptake by plants and it is a function of water required by plant transpiration and the amount of available water. If the upper layers in the soil profile do not contain enough water to meet the potential water uptake, an increase in the EPCO parameter allows water uptake to occur from the lower soil layers. EPCO values were higher in 25L2 with median values around 0.5, while for the others the median values were below 0.5, allowing less water uptake from the lower soil layers. The ESCO parameter represents the influence of capillarity on soil evaporation in each soil layer; it ranges between 0.01 and 1. When ESCO is low, the model extracts more water from lower soil layers to meet the evaporative demand. The configurations 25L2, 100L1 and 100L3 had median values around 0.4 and maximum interquartile between 0.26 and 0.5. Instead, 100L2 and 25L3 had respectively the smallest (0.2) and highest median value (0.6). The available water capacity (SOL AWC) adjustment (rate value) was lower for 25L1 and 100L1 (median value around 0.1) than other configurations (median values from 0.15 to 0.2), indicating the necessity for other configurations to increase of about 15 ~ 20% the water capacity in soil layers. The saturated hydraulic conductivity (SOL K) required the same adjustment of about 25% in all configurations, excepted for 100L1 that required 20% of increment.

Among the groundwater parameters (see the parameters identified with [4] in Figure S2), the recharge of deep aquifer (RCHRG DP) had the highest median value of around 0.18 for configurations 100L1, 100L2, 100L3 and 25L3, while 25L1 and 25L2 had the smallest median values (around 0.10) and interquartile. In addition, the median of the threshold depth of water in the shallow aquifer (GWQMN, mm H_2O) increased from L1 to L3 (both DEM pixel size) from approximately 500 to 600 mm, as well as the groundwater revap coefficient (GW REVAP) from 0.07 to 1 for DEM25. In addition, the median of the threshold depth of water in the shallow aquifer for revap was the highest for 100L1 (343 mm).

The baseflow alpha factor (ALPHA BF) values were very similar in each configuration, with median values around 0.75 indicating a rapid hydrological response of the subbasins. About the groundwater delay parameter (GW DELAY, days), 100L1 reached the lowest median value of about 27 days followed by 25L2, 100L2 and 100L3, while 25L1 and 25L3 reached the highest median values of 43 and 48 days respecttively.

It may thus be gathered that the calibrated parameters were able to reflect the behaviour of the subbasins changing the DEML configurations. In particular, quite different parameter values were obtained between the configurations during the adjustment of lateral flow (i.e. SOL AWC and ESCO) and baseflow processes (i.e. RCHRG DP, GWQMN and GW REVAP) indicating that the DEML configurations may have a markedly impact on the prediction of the related streamflow components.

3.4. Impact of Hillslope Length on Streamflow Components

While the impact of hillslope length on the total streamflow predictions appeared negligible, the analysis of streamflow components revealed some differences between DEML configurations as illustrated in Figure S3. Figure S3 shows the comparisons between monthly simulated and observed surface runoff, SR, (Figure S3a), lateral flow, LF (Figure S3b) and baseflow, BF (Figure S3c) for the 98 calibrated gauging stations in the period 1995 ~ 2006 (see Table 1), that were considered more representative of the natural hydrological behaviour.

The interquartile of monthly SR ranged from 0.12 to 2.8 m^3/s for all configurations (Figure S3 a), with median values around 0.75 m^3/s that fell on the observed interquartile (0.24 ~ 1.54 m^3/s). Spearman's rank correlation coefficients of the SR monthly residuals (simulation-observation) were all very high (ρ > 0.97). These findings indicate that all DEML configurations simulated monthly SR well and in very similar way, thus the impact of DEM pixel size and hillslope length was negligible.

The interquartile range of lateral flow was overestimated in all configurations (maximum interquartile range from 0.17 to 1.3 m^3/s) compared to the observations (between 0.05 ~ 0.3 m^3/s) (Figure S3 b). In particular, the median LF ranged from 0.44 m^3/s (100L3) to 0.62 m^3/s (100L1), while the median of observation was around 0.13 m^3/s. The Spearman's rank correlation coefficients ρ between LF monthly residuals of lateral flow of the calibration dataset (Table 3) reveal that configurations 100L2 and 25L3 were high correlated (ρ = 0.96), as well as 25L2 ~ 25L3 and 100L2 ~ 100L3 reflecting that L3 resulted in lateral flow values similar to L2 method. Conversely, configurations 100L1 ~ 100L2 and 100L1 ~ 25L3 had the lowest ρ coefficients, around 0.40, followed by 100L1 ~ 25L2 (ρ = 0.54) and 100L1 ~ 100L3 (ρ = 0.60), reflecting the differences in hillslope length distributions. Furthermore, the configurations 25L3, 25L1 and 25L2 were strongly correlated to 100L3, 100L1 and 100L2 respectively, indicating that the DEM pixel size didn't affect the lateral flow estimations.

Table 3. Spearman's Rank Correlation Coefficient ρ between the Monthly Residuals of Lateral Flow (LF, m^3/s) of DEML Configurations for the Calibrated Dataset (#98 gauged stations, 1995-2006)

	25L1	25L2	25L3	100L1	100L2	100L3
25L1	1					
25L2	0.79	1				
25L3	0.66	0.93	1			
100L1	0.86	0.54	0.42	1		
100L2	0.64	0.91	0.96	0.41	1	
100L3	0.68	0.86	0.90	0.60	0.90	1

The maximum interquartile range of baseflow was between 0.3 and 3.5 m^3/s with a median around 1.5 m^3/s, slightly lower the median of observation (around 2 m^3/s). Spearman's rank correlation coefficients ρ of BF residuals were all very high (ρ > 0.9), indicating that, as for surface runoff, DEM pixel size and hillslope length had negligible impact of baseflow estimation.

While the method used to subdivide the streamflow into its components introduced a certain degree of uncertainty (Huyck et al., 2005), an explanation of LF overestimation and its sensitivity to hillslope length can be found in SWAT equations (Equations 7 and 8). The assumption that $\sin(\alpha_{hill}) \sim \tan(\alpha_{hill})$ is not valid in steep slopes for which the difference between tangent and sine is not negligible (Bieger et al., 2015). As a consequence, at steep slopes lateral flow resulted overestimated. The overestimation of LF means an underestimation of the amount of soil water that is available for percolation to the groundwater (Bieger et al., 2015), thus leads to the underestimation of baseflow. In Figure S3c shows that configuration 100L1 produced the highest overestimation of lateral flow (median of value around 0.62 m^3/s) that corresponded to a larger underestimation of baseflow compared to other configurations (median around 1.4 m^3/s, Figure S3b), confirming this mechanism.

In addition, it was observed that the lateral flow monthly residuals increased with the elevation as showed in Figure S4. This figure shows the monthly residuals for three classes of elevation: from 0 to 500 m, from 500 to 1000 m and over 1000 m. In the class of the highest elevation, 25L1 and 100L1 reached the highest residuals (median values of around 0.7 m^3/s), indicating that the SWAT default configuration (both with finer and courser DEM) was not able to reproduce correctly the lateral flow in mountainous and steep subbasins.

3.5. Impact of Hillslope Length on Water Yield Predictions at Different Spatial and Temporal Scales

Figure S5 shows the impact of DEML configurations on the HRU water yield and its components (in mm) at annual and monthly time scale as mean during the simulation period (1995 ~ 2009) compared to the related interquartile of 98 calibrated stations for the same period. Mean annual and monthly HRU surface runoff (SR) was very similar across all configu-

rations (Figure S5a and b), confirming that DEM pixel size and hillslope length did not impact the generation of surface runoff, nor the time of concentration (hr) or peak runoff rate (m^3/s). The median of time concentration was around 10 hours and the median of peak runoff rate was approximately 20 m^3/s in all configurations. These results concur to Jarihani et al. (2015) findings on the impact of DEM-pixel size on hydrodynamic model, which demonstrated that the peak runoff rate did not change at DEM pixel size ranging from 30 to 250 m and that the time of concentration did not change until a pixel size of 120 m. However, the long-term annual and monthly mean of SR were close to the 75 percentile of observations, highlighting a slightly long-term overestimation of observations.

Differences in HRU lateral flow (LF) and baseflow (BF) (Figure S5c, d, e, f) confirmed the sensitivity of LF to hillslope length due to Equations 7 ~ 8 and its linked impact on BF already observed with the streamflow analysis. The configurations 25L1 and 100L1 produced the highest values of lateral flow, in the range of 60 ~ 85 mm/y dropping out to the interquartile of observations (from 9 mm in 2008 to 57 mm in 2002), while 100L2, 100L3 and 25L3 were within the observed range (Figure S5c). The peak of mean monthly LF was observed in May; configurations 25L1 and 100L1 reached the maximum value of 13 mm/month followed by 25L2 with about 7 mm/month and then the other configurations with values in the range of 4 ~ 5 mm/month (Figure S5d).

For BF, the impact of DEML configurations was the opposite than for LF, with configuration 100L2 generating the highest baseflow, from 194 mm/y (year 2004) to 344 mm/y (year 2002), while 100L1 generated the lowest, from 160 mm/y (year 2004) to 306 mm/y (year 2002; Figure S5 e). The largest differences in monthly baseflow were from April to November, and the maximum difference between 100L1 and 100L2, about 5 mm/month, was observed in May (Figure S5 f).

Mean annual and monthly water yields (WYLD, Figure S5 g, h) didn't show remarkable differences between DEML configurations, except for 25L1, which resulted in the highest monthly water yield, particularly for the period March-September (Figure S5h). Figure S5 shows also that BF and WYLD were within the observed interquartile in all configurations both for annual and monthly time step.

3.6. Impact of Hillslope Length on Water Balance

Figure 5a, c and e shows the influence of DEML configurations on the proportions of evapotranspiration (ET) and water yield (WYLD), while the proportions of water yield components (surface runoff, lateral flow and baseflow) are shown in Figure 5b, d and f.

The analysis was limited to the whole Upper Danube and two selected main tributaries with comparable drain area, the Morava and Inn river basins that however differ in elevation, slopes, climate, and land cover. The Inn river basin (25920 km^2) is characterized by a mean elevation of 1300 m, mean of

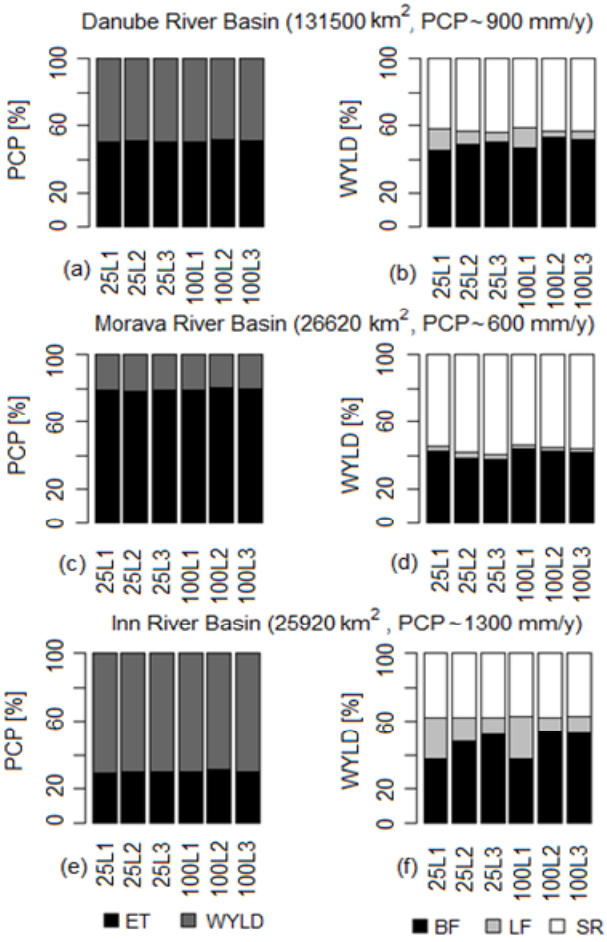

Figure 5. Bar plots of mean annual partitioning of precipitation (PCP) in evapotranspiration (ET) and water yield (WYLD) (a, c, e, all in %) and the water yield in surface runoff (SR), lateral flow (LF) and baseflow (BF) (b, d, f, all in %) for the entire Upper Danube Basin, the Morava and Inn River Basins.

precipitation of 1200 mm, median slope of 27% and the forest covers more than 50% of total area. Instead, the Morava (26628 km^2) river basin is characterized by a mean elevation of 380 m, precipitation around 600 mm, negligible slope (mean value of 5%), and the cropland is dominant covering more than 60% of total area.

The impact of DEML configurations on ET and WYLD was negligible. This is not surprising considering that the same HRUs were kept in all DEML configurations, therefore interception, infiltration and surface runoff generation were not affected by the configuration. ET in the Upper Danube the ET was 50% of total precipitation (Figure 5a), while in the Morava it was about 80% of total precipitation (Figure 5c), and in the Inn river only about 30% (Figure 5e).

Instead, some differences between the DEML configurations could be observed in the components of the water yields (Figure 5b, d, f), namely in the partitioning between lateral flow and baseflow, whereas surface runoff was very similar across configurations in all instances. For the whole Upper Danube (Figure 5b), the surface runoff was approxi-

mately 42% of total WYLD, while the lateral flow varied from 4% for 100L2 to 12% for 100L1. Baseflow volumes were higher where lateral flow volumes were lower, and varied from 46% (100L1) to 53% (100L2) of total WYLD. Differences in lateral flow and baseflow were even more noticeable in the Inn (~ 29920 km^2, Figure 5 d). In the Inn, the surface runoff was around 37% of total WYLD, and the lateral flow changed from 8% of 100L2 to 25% 100L1 and 25L1. Again, the baseflow were higher where the lateral flow was lower approximately 38% for configurations 25L1 and 100L1, while it was 54% for 100L2.

It is noteworthy that the highest lateral flow of 25L1 and 100L1 corresponds to very low values of hillslope length (median value approximately 15 m) in mountainous and steeper subbasins. Instead, at increasing values of hillslope length, lateral flow tended to decrease. Even in the Morava river basin, where lateral flow was small due to the high evapotranspiration and the small water yield, differences in LF and the impact of hillslope length could still be detected. The median hillslope length was 60 m for configuration 25L1, 100L1 and 100L2, and the lateral flow was respectively 2.7%, 3.3% and 1.7% of total water yield. In 25L2, the median hillslope length was around 30 m, and lateral flow was higher, about 3.5%. However, these differences were negligible.

4. The Recommended DEML Configurations

The analysis indicated that the streamflow was well predicted in all DEML configurations, thus both DEMs and the three hillslope length methods could be valid (Figure 3). This confirms findings reported in other studies (i.e. Chaplot, 2005; Lin et al., 2010; Zhang et al, 2014) that focused mainly on DEM pixel size. However, the analysis of streamflow and water yield components revealed differences between the configurations affecting SWAT simulation of the hydrological processes.

Surface runoff was well simulated in all DEML configurations (Figure 3a), albeit there was a tendency of overestimation for all DEML configurations (Figure S5a, b). The generated surface runoff was unaffected by changes in DEM, hillslope length, or the simulation of lateral flow. An increase of lateral flow should lead to a larger amount of surface runoff due to higher soil water content in the footslopes, but SWAT cannot account for this. The adjustment of the Curve Number (CN) in steep slopes as pointed out for instance by Huang et al. (2006) and Bieger et al. (2015) could partly account for this, and would be recommended. As a consequence, in the surface runoff estimations the DEML configuration had a secondary role compared to the uncertainty of the model structure such as the model inaccuracy due to over-simplification of the processes considered in the model (Wagener and Gupta, 2005; Liu and Gupta, 2007; Clark et al., 2008).

Instead, DEML configurations impacted lateral flow simulation, which uses hillslope length and slope (Equation 8). When the hillslope length increases, the lateral flow decreases. Thus in steep areas, where hillslope length is the lowest when using SWAT default method, configurations 25L1 and 100L1 produced the highest lateral flow, followed by 25L2. Conversely, in flat areas, the differences between the configurations were negligible. This could be observed at HRU scale: for instance, the highest lateral flow was observed in forest and pastures HRUs of the Alpine areas, while in the flatter cropland lateral flows was low. The same could be observed at basin level: in the Inn river basin, characterized by high elevation, steep slopes, high precipitation, the lateral flow ranged 8% for 100L2 to around the 25% for 25L1 and 100L1 (Figure 5).

The high lateral flow estimations in mountainous areas in the 100L1 and 25L1 appear to be excessive (Figure S4). Indeed, all DEML configurations overestimated lateral flow in the calibrated headwaters (Figure S3), due to approximation of the sine of the slope with its tangent in Equation (8), which affected lateral flow estimation in steep slopes. Lateral flow in configurations 100L2, 100L3 and 25L3 appear more reasonable. Instead 25L2 slightly overestimated lateral flow albeit less than 25L1 and 100L1.

The overestimation of later lateral flow lead to an underestimation of baseflow (Figures S3c and 5). The configuretions 100L2, 25L3 and 100L3 appeared to predict reasonable baseflow, partly overcoming the lateral flow issue (Figures S3 and S5).

In conclusion, configuration 100L2 best suited for a spatially distributed hydrological model. Configuration 25L3 and 100L3 were comparable to 100L2, and could be considerate as simple alternatives, that however should be evaluated in each specific case study.

5. Conclusions

Hillslope length plays an important role in controlling the hydrological response of a basin, since it exerts a primary control on the fluxes transported towards the river network through the soils and hillslopes. In this study the impact of hillslope length on SWAT streamflow predictions was investigated using three hillslope length methods (L1, L2, L3) and two DEMs of 25- and 100-m pixel size given a complete overview of possible impacts of hillslope length estimations on SWAT simulated hydrological processes. These methods lead to different hillslope length distributions, but only L2 method (the three-dimensional DEM flow accumulation method) was sensitive to DEM resolution, resulting in longer hillslope length with the 100 m DEM pixel size (~ 60 m) than with the 25 m DEM (~ 30 m). Conversely, the SWAT default method (L1) was not affected by DEM pixel size.

While all configurations reached satisfactory simulation of monthly streamflow, the analysis of streamflow components (surface runoff, lateral flow, baseflow) highlighted differences in configuration outputs. The DEM resolution did not impact SWAT streamflow simulations; both DEMs were sufficient for streamflow modelling at this scale, confirming literature findings that a finer resolution DEM may not necessarily improve model simulations. Rather, the optimal DEM pixel size depends on the environmental characteristics, the

desired level of prediction, and the model output of interest (Chaplot, 2014; Zhang et al., 2014).

The hillslope length method had no impact on total streamflow and surface runoff, but had an important impact on the partition between lateral flow and baseflow. This bears important consequences in the simulation of pollutant movements, like nitrates, in the landscape. The current default method L1 was shown to produce larger errors in the simulation of streamflow components in steep areas. In particular, the lateral flow always yields higher residuals where slopes are higher and the associated hillslope lengths smaller. This reflects the reported inverse power relationship between hillslope length and lateral flow estimation (Equation 7). However, it seems that this inverse power relationship does not hold for larger catchments, where estimated values of hillslopes length lower than 10 meters might not be realistic (Figure 2 d).

The hillslope length method based on DEM analysis of flow accumulation (L2) resulted in the most adequate estimations of lateral flow and baseflow in steep regions. The L2 method is conceptually more consistent for application in spatially distributed models than the current method, since it provides a more reliable description of the landscape morphologies throughout a 3D analysis. Furthermore, the L2 method is sensitive to DEM resolution, since the use of different DEM pixel size produces different hillslope lengths, while the default ArcSWAT method returns the same values of hillslope lengths. For future SWAT development, we believe that the LS-TOOL (Zhang et al., 2013) should be integrated in ArcSWAT (interface GIS of SWAT) in order to replace the current hillslope length calculation. However, further efforts should be done for improving the representativeness of surface runoff according to the lateral flow estimations.

The combination of DEM100 m and L2 was the optimal configuration to predict streamflow in the Upper Danube and is recommended in general in large basins given that the use of DEM100 reduces the computational burden for SWAT application at large scale (i.e. time of calculation of topographic characteristics in the phase of "Automatic Watershed Delineation").

Choosing a constant hillslope length of 50 m (L3) however was a good second alternative to be considered when DEM resolution is of 25 m or coarser, especially if the DEM accuracy is low. Furthermore, given the importance of hillslope length on sediment outputs reported in literature (Chaplot, 2014; Zhang et al., 2014) the impact of hillslope length on sediment predictions was explored in Vigiak et al. (2015), in a study that confirmed the suitability of L2 method for prediction of sediment yields.

These results can be considered as representative for a wide range of landscapes since the Upper Danube is characterized by heterogeneous topography, large climate variations, several land covers/uses and soil types. They can also be considered valid for any catchment hydrological model that shares a structure for runoff and lateral flow partitioning similar to the SWAT (Equation 2 and 7). In such models, a proper representation of DEM derivatives (such as hillslope length) commensurate to DEM pixel size is thus fundamental for a correct representation of hydrological processes.

Acknowledgement. We thank the Editor and the three anonymous Reviewers for their thoughtful comments.

Supporting Material. This paper contains supporting materials which are available in its online version.

References

Abbaspour, K.C. (2008). *SWAT-CUP2: SWAT Calibration and Uncertainty Programs-A User Manual,* Department of Systems Analysis, Integrated Assessment and Modelling (SIAM), Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.

Abbaspour, K.C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., and Kløve, B. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *J. Hydrol.,* 524, 733-52. http://dx.doi.org/10.1016/j.jhydrol.2015. 03.027

Arnold, J.G., Srinivasan, R., Muttiah, R.S., and Williams, J.R. (1998). Large area hydrologic modeling and assessment: Part I. Model development. *J. Am. Water Resour. Assoc.,* 34(1), 73-89. http://dx.doi.org/10.1111/j.1752-1688.1998.tb05961.x

Arnold, J.G., Kiniry, J.R., Srinivasan, R., Williams, J.R., Haney, E.B., and Neitsch, S.L. (2012a). *Soil and Water Assessment Tool Input/Output documentation Version 2012,* Texas Water Resources Institute Technical report 436, Texas A&M University System College Station, Texas, U.S. http://swat.tamu.edu/documentation/2012-io/.

Arnold, J.G., Moriasi, D.N., Gassman, P.W. et al. (2012b). SWAT: model use, calibration, and validation. *Trans. ASABE,* 55(4), 1491-1508. http://dx.doi.org/10.13031/2013.42256

ASTER GDEM database. http://www.jspacesystems.or.jp/ersdac/GDEM/E/index.html.

Bartholome, E., and Belward, A.S. (2005). GLC2000: A new approach to global land cover mapping from earth observation data. *Int. J. Remote Sens.,* 26(9), 1959-1977. http://dx.doi.org/10.1080/01431160412331291297

Bashfield, A. and Keim, A. (2011). Continent-wide DEM Creation for the European Union. *34th International Symposium on Remote Sensing of Environment, The GEOSS Era: Towards Operational Environmental Monitoring,* Sydney, Australia 10-15 April 2011. http://www.isprs.org/proceedings/2011/isrse-34/211104015Final00143.pdf.

Bieger, K., Hörmann, G., and Fohrer, N. (2012). Using residual analysis, auto- and cross-correlations to identify processes for the calibration of the SWAT model in a data scarce region. *Adv. Geosci.,* 31, 23-30. http://dx.doi.org/10.5194/adgeo-31-23-2012

Bieger, K., Hörmann, G., and Fohrer, N. (2015). Detailed spatial analysis of SWAT-simulated surface runoff and sediment yield in a mountainous watershed in China. *Hydrol. Sci. J.,* 60(5), 2015. http://dx.doi.org/10.1080/02626667.2014.965172

Britz, W. (2004). *CAPRI Modelling System Documentation,* Final report of the FP5 shared cost project CAP-STRAT "Common Agricultural Policy Strategy for Regions, Agriculture and Trade", QLTR 2000-00394; Universität Bonn, Germany.

Bouraoui, F., and Aloe, A. (2007). *European Agrochemicals Geospatial Loss Estimator: Model development and Applications,*

EUR-Scientific and Technical Research series, ISSN 1018-5593, Office for Official Publications of the European Communities, Luxembourg.

Bonuma, N.B., Rossi, C.G., Arnold, J.G. et al. (2014). Simulating landscape sediment transport capacity by using a modified SWAT model. *J. Environ. Qual.,* 43(1), 55-66 http://dx.doi.org/10.2134/jeq2012.0217

Chaplot, V. (2005). Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO_3-N loads predictions. *J. Hydrol.,* 312(1-4), 207-222. http://dx.doi.org/10.1016/j.jhydrol.2005.02.017

Chaplot, V. (2014). Impact of spatial input data resolution on hydrological and erosion modeling: recommendations from a global assessment. *Phys. Chem. Earth,* 67-69, 23-35. http://dx.doi.org/10.1016/j.pce.2013.09.020

Chaubey, I., Cotter, A.S., Costello, T.A., and Soerens, T.S. (2005). Effect of DEM data resolution on SWAT output uncertainty. *Hydrol. Process.,* 19(3), 621-628. http://dx.doi.org/10.1002/hyp.5607

Chow, V.T., Maidment, D., and Mays, L.W. (1988). *Applied Hydrology,* McGraw Hill.

Clark, M.P., Slater, A.G., Rupp, D.E. et al. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resour. Res.,* 44(12), W00B02. http://dx.doi.org/10.1029/2007WR006735

Cotter, A.S., Chaubey, I., Costello, T.A., Soerens, T.S., and Nelson, M.A. (2003). Water quality model output uncertainty as affected by spatial resolution of input data. *J. Am. Water Resour. Assoc.,* 39(4), 977-986. http://dx.doi.org/10.1111/j.1752-1688.2003.tb04420.x

Dettling, M., and Maechler, M. (2012). *Supclust, Supervised clustering of predictor variables such as genes,* R package version 1.0-7.

Di Luzio, M., Srinivasan, R., and Arnold, J.G. (2004). A GIS-coupled hydrological model system for the watershed assess- ment of agricultural nonpoint and point sources of pollution. *Trans. GIS,* 8(1), 113-136. http://dx.doi.org/10.1111/j.1467-9671.2004.00170.x

Di Luzio, M., Arnold, J.G., and Srinivasan, R. (2005). Effect of GIS data quality on small watershed streamflow and sediment simulations. *Hydrol. Process.,* 19(3), 629-650. http://dx.doi.org/10.1002/hyp.5612

Dixon, B., and Earls, J. (2009). Resample or not? Effects of resolution of DEMs in watershed modeling. *Hydrol. Process.,* 23(12), 1714-1724. http://dx.doi.org/10.1002/hyp.7306

Döll, P., Berkhoff, K., Bormann, H., Fohrer, N., Gerten, D., Hagemann, S., and Krol, M. (2008). Advances and visions in large-scale hydrological modelling: Findings from the 11th Workshop on Large-scale Hydrological Modelling. *Adv. Geosci.,* 18, 51-61. http://dx.doi.org/10.5194/adgeo-18-51-2008

EU-DEM Metadata (2013). Available online: http://www.eea.europa.eu/data-and-maps/data/eu-dem#tab-metadata/.

FAO/IIASA/ISRIC/ISS-CAS/JRC (2008). *Harmonized world soil database (version 1.0),* FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Geladi, P., and Kowalski, B. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta,* 185, 1-17. http://dx.doi.org/10.1016/0003-2670(86)80028-9

GRDC (2010). *The Global Runoff Data Centre,* 56068 Koblenz, Germany.

Grieve, S.W.D., Mudd, S.M., and Hurst, M.D. (2016) How long is a hillslope? *Earth Surf. Process. Landforms*. http://dx.doi.org/10.1002/esp.3884

Gudmundsson, L., Tallaksen, L.M., Stahl, K. et al. (2012). Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. *J. Hydrometeorol.,* 13(2), 604-620. http://dx.doi.org/10.1175/JHM-D-11-083.1

Gupta, H.V., Sorooshian, S., and Yapo, P.O. (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrologic Eng.,* 4(2), 135-143. http://dx.doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)

Harmel, R.D., Smith, P.K., Migliaccio, K.W. et al. (2014). Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. *Environ. Model. Software,* 57, 40-51. http://dx.doi.org/10.1016/j.envsoft.2014.02.013

Harvey, C.L., Dixon, H., and Hannaford, J. (2012). An appraisal of the performance of data infilling methods for application to daily mean river flow records in the UK. *Hydrol. Res.,* 43(5), 618-636. http://dx.doi.org/10.2166/nh.2012.110

Henning, C. (2010). *fpc: Flexible procedures for clustering,* R package version 2.0-3.

Heuvelmans, G., Muys, B., and Feyen, J. (2006). Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. *J. Hydrol.,* 319(1-4), 245-265. http://dx.doi.org/10.1016/j.jhydrol.2005.07. 030

Hickey, R. (2000). Slope Angle and Slope Length Solutions for GIS. *Cartogr.,* 29 (1), 1-8. http://dx.doi.org/10.1080/00690805.2000.9714334

Huang, M., Gallichand, J., Wang, Z., and Goulet, M. (2006). A modification to the soil conservation service curve number method for steep slopes in the Loess Plateau of China. *Hydrol. Process.,* 20(3), 579-589. http://dx.doi.org/10.1002/hyp.5925

Hubert, L. and Arabie, P. (1985). Comparing partitions. *J. Classification,* 2(1), 193-218. http://dx.doi.org/10.1007/BF019080 75

Huyck, A.A.O., Pauwels, V.R.N., and Verhoest, N.E.C. (2005), A base flow separation algorithm based on the linearized Boussinesq equation for complex hillslopes. *Water Resour. Res.,* 41(8), W08415. http://dx.doi.org/10.1029/2004WR003789

Jarihani, A., Callow, J., McVicar, T., Van Niel, T., and Larsen, J. (2015). Satellite-derived Digital Elevation Model (DEM) selection, preparation and correction for hydrodynamic modelling in large low-gradient and data-sparse catchments. *J. Hydrol.,* 524, 489-506. http://dx.doi.org/10.1016/j.jhydrol.201 5.02.049

Kim, J., Park, Y., Yoo, D., Kim, N., Engel, B.A., Kim, S., Kim, K., and Lim, K.J. (2009). Development of a SWAT patch for better estimation of sediment yield in steep sloping watersheds. *J. Am. Water Resour. Assoc.,* 45(4), 963-972. http://dx.doi.org/10.1111/j.1752-1688.2009.00339.x

Klein Goldewijk, K., and van Drecht, G. (2006). HYDE 3: Current and historical population and land cover. *In: Integrated modeling of global environmental change. An overview of IMAGE,* 2, 93-111. Netherlands Environmental Assessment Agency (MNP), Bilthoven, The Netherlands.

Lehner, R., and Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.,* 296(1-4), 1-22. http://dx.doi.org/10.1016/j.jhydrol.2004.03.028

Lin, S.P., Jing, C.W., Coles, N.A., Chaplot, V., Moore, N.J., and Wu, J.P. (2013). Evaluating DEM source and resolution uncertainties in the soil and water assessment tool. *Stoch. Environ. Res. Risk Assess.,* 27(1), 209-221. http://dx.doi.org/10.1007/s00477-012-0577-x

Liu, Y., and Gupta, H.V. (2007). Uncertainty in hydrological modeling: towards an integrated data assimilation framework. *Water Resour. Res.,* 43(7), W07401. http://dx.doi.org/10D1029/2006WR005756.

Lyne, V., and Hollink, M. (1979). Stochastic time-variable rainfall-runoff modelling. *Ins. Eng. Aust. Natl. Conf.,* 1979, 89-93.

Malagò, A., Pagliero, L., Bouraoui, F., and Franchini, M. (2015). Comparing calibrated parameter sets for the Scandinavian Peninsula and for the Iberian Peninsula. *Hydrol. Sci. J.,* 60(5), 949-967.

http://dx.doi.org/10.1080/02626667.2014.978332

Meilă, M. (2007). Comparing clusterings-an information based distance. *J. Multivariate Anal.,* 98(5), 873-895. http://dx.doi.org/10.1016/j.jmva.2006.11.013

Mevic, B., and Wehrens, R. (2007). The pls package: principal component analysis and partial least squares regression in R. *J. Stat. Software,* 18(2), 1-23. http://dx.doi.org/10.18637/jss.v018.i02

Monfreda, C., Ramankutty, N., and Foley, J. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem. Cycles*, 22(1), GB1022. http://dx.doi.org/10.1029/2007GB002947

Moriasi, D.N., Arnold, J.G., van Liew, M.W., Bingner, R.L., Harmel, R.D., and Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE,* 50(3), 885-900. http://dx.doi.org/10.13031/2013.23153

Nash, J.E., and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I-A discussion of principles, *J. Hydrol.,* 10(3), 282-290. http://dx.doi.org/10.1016/0022-1694(70)90255-6

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., and Williams, J.R. (2011). *Soil and Water assessment tool-theoretical documentation,* Texas Water Resources Institute Technical report 406. Texas A&M University System College Station, Texas, U.S. http://swat.tamu.edu/media/99192/swat2009-theory.pdf.

Nester, T., Kirnbauer, R., Gutknecht, D., and Blöschl, G. (2011). Climate and catchment controls on the performance of regional flood simulations. *J. Hydrol.,* 402(3-4), 340-356. http://dx.doi.org/10.1016/j.jhydrol.2011.03.028

Ntegeka, V., Salamon, P., Gomes, G., Sint, H., Lorini, V., and Thielen, J. (2013). EFAS-Meteo: A European daily high-resolution gridded meteorological data set for 1990-2011.

Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H., Koka, S., and Agrawal, A. (2006). ArcGIS SWAT: A geodata model and GIS interface for SWAT. *J. Am. Water Resour. Assoc.,* 42(2), 295-309. http://dx.doi.org/10.1111/j.1752-1688.2006.tb03839.x

Pagliero, L., Bouraoui, F., Willems, P., and Diels, J. (2014) Large-scale hydrological simulations using the soil water assessment tool, protocol development, and application in the Danube Basin. *J. Environ. Qual.,* 43(1), 145-154. http://dx. doi.org/10.2134/jeq2011.0359

Parajka, J., Merz, R., and Blöschl, G. (2007). Uncertainty and multiple objective calibration in regional water balance modeling Case study in 320 Austrian catchments. *Hydrol. Process.,* 21(4), 435-446. http://dx.doi.org/10.1002/hyp.6253

Pekárová, P., Onderka, M., Pekár, J., Miklánek, P., Halmová, D., Škoda, P., and Bačová Mitková, V. (2008). *Hydrologic Scenarios for the Danube River at Bratislava,* Ostrava, KEY Publishing, http://pavla.pekarova.sk/monografie.

PHU (2007). Potential Heat Unit (PHU) Program. http:// swat.tamu.edu/software/potential-heat-unit-program/.

Portmann, F., Siebert, S., Bauer, C., and Döll, P. (2008). *Global data set of monthly growing areas of 26 irrigated crops: Version 1.0,* Frankfurt Hydrology Pap. 06, Inst. of Phys. Geogr., Univ. of Frankfurt, Frankfurt, Germany.

R Development Core Team (2008). *R: A language and environ- ment for statistical computing,* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www. R-project.org.

Rank, D., Papesch, W., and Tesch, R. (2005). Runoff characteristics of the upper Danube basin: conclusions from long-term environmental isotope records. *Geophys. Res. Abstr.,* 7, 03315.

Rouholahnejad, E., Abbaspour, K.C., Srinivasan, R., Bacu, V., and Lehmann, A. (2014). Water resources of the Black Sea Basin at high spatial and temporal resolution, *Water Resour. Res.,* 50(7), 5866-5885. http://dx.doi.org/10.1002/2013WR014 132

Schiller, H., Miklos, D., and Sass, J. (2010). The Danube River and its Basin physical characteristics, water regime and water balance. *In: Hydrological Processes of the Danube River Basin, Perspectives from the Danubian Countries*. Springer. 25-77. http://dx.doi.org/10.1007/978-90-481-3423-6_2

Sharma, A., and Tiwari, K.N. (2014). A comparative appraisal of hydrological behavior of SRTM DEM at catchment level. *J. Hydrol.,* 519(27), 1394-1404. http://dx.doi.org/10.1016/j.jhydrol.2014.08.062

Sommerwerk, N., Hein, T., Schneider-Jakoby, M. et al. (2009). The Danube River Basin. *River of Europe,* 59-112. http://dx.doi.org/10.1016/b978-0-12-369449-2.00003-5

Spruill, C.A., Workman, S.R., and Taraba, J.L. (2000). Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. *Trans. ASAE,* 43(6), 1431-1439. http://dx.doi.org/10.13031/2013.3041

SRTM Version 4.1. STRM 90 m Digital Elevation Database v4.1. http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1.

SWAT-CUP version 5.6.1.2. Eawag, Zurig, Switzerland. http://www.eawag.ch/forschung/siam/software/swat/index_E.

USDA Soil Conservation Service (1972). *SCS National Engineering Handbook,* Section 4, Hydrology. Washington, DC: USDA.

Vigiak, O., Malagò, A., Bouraoui, F., Vanmaercke, M., and Poesen, J. (2015). Adapting SWAT hillslope erosion model to predict sediment concentrations and yields in large Basins. *Sci. Total Environ.*, 538, 855-875. https://doi.org/10.1016/j.scitotenv.2015.08.095

Vogt, J. et al. (2007). *A pan-European River and catchment Database,* JRC Reference Reports, EUR 229220 EN.

Wagener, T., and Gupta, H.V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environ. Res. Risk Assess.,* 19(6), 378-387. http://dx.doi.org/10.1007/s00477-005-0006-5

Ward, J.H. (1963). Hierarchical grouping to optimize an objecttive function. *J. Am.Stat. Assoc.,* 58(301), 236-244. http://dx.doi.org/10.1080/01621459.1963.10500845

Wechsler, S.P. (2007). Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrol. Earth Syst. Sci.,* 11, 1481-1500. http://dx.doi.org/10.5194/hess-11-1481-2007

Williams, J.R., Nicks, A.D., and Arnold, J.G. (1985). Simulator for Water Resources in Rural Basins, *J. Hydraul. Eng. ASCE,* 111(6), 970-986. http://dx.doi.org/10.1061/(ASCE)0733-9429(1985)111:6(970)

Winchell, M., Srinivasan, R., Di Luzio, M., and Arnold, J.G. (2013). *ArcSWAT Interface For SWAT 2012: User's Guide,* Texas Agricultural Experiment Station (Texas) and USDA Agricultural Research Service (Texas), Temple (Texas), March 2013.

Wishmeier, W.H., and Smith, D.D. (1978). *Predicting rainfall erosion losses,* USDA Agricultural Research Service (USDA-ARS) Handbook 537.

Wold, H. (1966). *Nonlinear estimation by iterative least squares procedures,* In Research Papers in Statistics, Wiley, New York.

Wu, S., Li, J., and Huang, G.H. (2008). A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. *Appl. Geogr.,* 28(3), 210-223. http://dx.doi.org/10.1016/j.apgeog.2008.02.006

Yao, C., McCool, D.K., and Elliot, W.J. (2010). *DEM Resolution Effects on Hillslope Length andSteepness Estimates for Erosion Modeling Conference Paper.* Conference: 2010 Pitts- burgh, Pennsylvania, June 20 ~ 23, 2010. http://dx.doi.org/10.13031/ 2013.32045

Zambrano-Bigiarini, M. (2013). *HydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time*

series, R package version 0.3-8. http://CRAN.R-project.org/package=hydroGOF.

Zhang, H., Yang, Q., Li, R., Liu, Q., Moore, D., He, P., Ritsema, C.J., and Geissen, V. (2013). Extension of a GIS procedure for calculating the RUSLE equation LS factor. *Comput. Geosci.,* 52, 177-188. http://dx.doi.org/10.1016/j.cageo.2012.09.027

Zhang, P., Liu, R., Bao, Y., Wang, J., Yu, W., and Shen., Z. (2014). Uncertainty of SWAT model at different DEM resolutions in a large mountainous watershed. *Water Res.,* 53, 132-144. http://dx.doi.org/10.1016/j.watres.2014.01.018