Journal of Hydrology 539 (2016) 141–161

Contents lists available at ScienceDirect

Journal of Hydrology

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

Evaluation of TOPLATS on three Mediterranean catchments

Javier Loizu *, Jesús Álvarez-Mozos, Javier Casalí, Mikel Goñi

Department of Projects and Rural Engineering, Public University of Navarre, Pamplona, Navarre, Spain

ARTICLE INFO

Article history:
Received 20 November 2015
Received in revised form 18 March 2016
Accepted 9 May 2016
Available online 24 May 2016
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Yasuto Tachikawa, Associate Editor

Keywords:
TOPLATS
Sensitivity analysis
Morris and Sobol methods
Calibration and validation strategies
Optimization algorithm
Hydrological modeling time scale

SUMMARY

Physically based hydrological models are complex tools that provide a complete description of the different processes occurring on a catchment. The TOPMODEL-based Land–Atmosphere Transfer Scheme (TOPLATS) simulates water and energy balances at different time steps, in both lumped and distributed modes. In order to gain insight on the behavior of TOPLATS and its applicability in different conditions a detailed evaluation needs to be carried out. This study aimed to develop a complete evaluation of TOPLATS including: (1) a detailed review of previous research works using this model; (2) a sensitivity analysis (SA) of the model with two contrasted methods (Morris and Sobol) of different complexity; (3) a 4-step calibration strategy based on a multi-start Powell optimization algorithm; and (4) an analysis of the influence of simulation time step (hourly vs. daily). The model was applied on three catchments of varying size (La Tejeria, Cidacos and Arga), located in Navarre (Northern Spain), and characterized by different levels of Mediterranean climate influence. Both Morris and Sobol methods showed very similar results that identified Brooks–Corey Pore Size distribution Index (B), Bubbling pressure (ψ_c) and Hydraulic conductivity decay (f) as the three overall most influential parameters in TOPLATS. After calibration and validation, adequate streamflow simulations were obtained in the two wettest catchments, but the driest (Cidacos) gave poor results in validation, due to the large climatic variability between calibration and validation periods. To overcome this issue, an alternative random and discontinuous method of cal/val period selection was implemented, improving model results.

1. Introduction

The intense development in the field of hydrological simulation offers researchers worldwide dozens of models capable of simulating streamflow and other processes, at different time and spatial scales (e.g., Burnash et al., 1973; Chiew and McMahon, 2002; Brocca et al., 2011). Although they can easily be applied on different conditions (in terms of climate, catchment size or time-step), achieving the best simulation results depends largely on the users' knowledge of model structure and available tools to maximize the accuracy of the results (Khakbaz et al., 2012). Thus, achieving optimal calibrated and validated streamflow values requires, first, detailed sensitivity analyses to provide the modeler with objective criteria to identify the parameters to include on the calibration procedure and next, calibration and validation strategies to find the parameter values that optimize model results (Van Werkhoven et al., 2009). Model performance and optimal parameter values will depend then largely on: (1) catchment size, (2) rainfall pattern and climate conditions, (3) modeling time-scale, and the suitability of model structure to all of them (Demaria et al., 2007).

Sensitivity Analysis (SA) techniques can identify influential parameters, i.e. those whose uncertainty reduction will have the most significant impact on improving model performance (Gan et al., 2014) and provide model users with useful information to reduce calibration dimensionality (Garambois et al., 2013). If some insensitive parameters are identified through SA, they can be fixed reasonably at given values over their variation range. Thus, reducing calibration computational cost without decreasing model performance.

Sun et al. (2012) classified SA methods into three types: (1) local, (2) screening and (3) global, depending on the way parameters were perturbed. Local methods quantify the percentage change of outputs due to the change of model inputs relative to their baseline (nominal) values (Tang et al., 2007). These methods, also referred to as One-at-A-Time (OAT), evaluate the response of output variables to fractional changes in one single input parameter and are therefore less efficient on complex models. Even on models where parameters are independent, the combination of single-parameter influences can make local methods to fail on

* Corresponding author at: Public University of Navarre, Department of Projects and Rural Engineering, Laboratorio de Hidráulica, Edificio Los Olivos, Campus Arrosadia, 31006 Pamplona-Iruña, Spain. Tel.: +34 948168962.
E-mail addresses: javier.loizu@unavarra.es, jloizu@ono.com (J. Loizu).

http://dx.doi.org/10.1016/j.jhydrol.2016.05.025
0022-1694/

capturing model behavior due to non-linearity of model response (Norton, 2009). Screening methods also analyze the model response to a change in the inputs by varying one parameter at a time, but they provide a global sensitivity measure, since different elementary effects (EE) for each parameter are calculated and averaged (Campolongo et al., 2011). They are commonly applied to cases where a large number of parameters needs to be analyzed, or to computationally expensive models where more demanding quantitative techniques might lead to extended simulation times. Finally, global methods, vary simultaneously all studied parameters within their defined parameter space, thus providing information on both individual sensitivity and parameter interaction degrees. Global methods look at the entire input parameters distribution, using specifically designed Monte Carlo sampling techniques of various levels of sophistication, but their application to computationally demanding models might be constrained due to the large number of model runs required (Song et al., 2015). Global methods are recognized as appropriate for hydrological modeling, as they have to evaluate nonlinear processes and high parameter and data uncertainty due to spatial heterogeneity (Spear et al., 1994). Global methods include the following groups (Tang et al., 2007): (1) Regional SA (Young, 1978), (2) Bayesian SA (Oakley and O'Hagan, 2004), (3) regression based approaches (Spear et al., 1994), and (4) variance decomposition methods (Saltelli et al., 2000). Screening and global SA methods include two steps: first, a strategy is used to sample the parameter space (i.e. Design of experiment, DoE) and next a numerical measure is used to quantify the impacts of sampled parameters on model output (Wagener and Kollat, 2007).

Once the most sensitive parameters of a model have been identified through SA procedures, they need to be calibrated, i.e. estimated through an inverse method so that observed and predicted output values are in agreement (Zhang et al., 2009). Therefore, successful application of any hydrological model depends on how accurately the model is calibrated (Duan et al., 1992). Although model calibration used to be a labor intense task that depended largely on modeler knowledge and experience, nowadays computers allow automatic calibration techniques. These are commonly optimization algorithms that search for a set of parameters values that minimize the model prediction error relative to available measured data for the system being modeled (Tolson and Shoemaker, 2007). Gupta et al. (1998) pointed out that automatic calibration success depends largely on three aspects: (1) adequate calibration data (mainly in terms of data length and climate variability contained), (2) the objective function (maximum likelihood functions for measuring the "closeness" of the model and the data), and (3) the selected optimization algorithm. However, some studies reported difficulties in finding unique (global) optimum parameter values due to parameter nonuniqueness or equifinality, parameter correlation, or other limitations (Duan et al., 1992).

Calibration of hydrological models for areas with irregular rainfall patterns, such as Mediterranean ones, implies an extra effort in terms of model adaptability and data availability (Loaiza-Usuga and Pauwels, 2008). Several authors (Gan and Biftu, 1996; Li et al., 2010; Perrin et al., 2007) noted that arid catchments are generally more difficult to model than humid ones due to the complexity and variability of hydrological processes there. This can be related to model's response to intense rainfall events and to large inter-annual rainfall variability. Conventional continuous calibration and validation period selection (i.e., selection of a calibration period of n years, followed by a validation period of m years) may be a limitation when large differences on climate variables are found among both periods. Thus, alternative (random and discontinuous) period selection methods that lead to a similar calibration and validation climatological conditions and to a minimum of high flows included on the calibration period are worth being explored (Kim and Kaluarachchi, 2009). As stated by (Sorooshian and Gupta, 1983) it is not the length of the data series used but the information contained in it and the efficiency with which that information is extracted that are important. Random sampling approaches are expected to overcome different difficulties, which could include: (1) data availability discontinuity (i.e. Kim and Kaluarachchi, 2009), (2) lack of data series long enough to achieve proper calibration and validation results, or (3) large climate variability between calibration and validation periods.

Optimization algorithms used on hydrological model calibration are divided into local (Tolson and Shoemaker, 2007) and global search methods (Duan et al., 1993). One of the first optimization algorithms was proposed by Powell (1964), and was applied for the first time to hydrological modeling by Kobayashi and Maruyama (1976). This algorithm is a local, derivative-free method where one parameter value is changed at-a-time. Chen et al. (2005) applied a modified multi-start version of the Powell method for model calibration, which is also implemented on this study.

Hydrological models cover a range of variability in terms of parameter complexity, running time-scale, conceptual structure and spatial distribution design (lumped and distributed). According to these characteristics, they may offer better results under certain terrain or climate conditions. Among them, there has been a significant development of catchment models based on the TOPMODEL concept (Beven and Kirkby, 1979). From this initial conceptualization, Famiglietti and Wood (1994), started the development of a full hydrological catchment model that incorporated a separate computation of water and energy balances. This model was called TOPMODEL-based Land–Atmosphere Transfer Scheme (TOPLATS).

TOPLATS can be run at any user-specified time step, from daily (Bormann et al., 2007) to hourly (Loaiza-Usuga and Pauwels, 2008), or even on less than a minute time-step (Seuffert et al., 2002). While this permits the model to be applied for an extensive range of purposes, it can also affect model performance, especially in terms of runoff and soil moisture processes simulation. It has been applied on a wide range of locations worldwide but TOPLATS simulations on Mediterranean catchments was only reported in Loaiza-Usuga and Pauwels (2008) and in Loaiza-Usuga and Poch (2009). The complexity of TOPLATS makes it necessary to use efficient SA methods to get a better understanding of its behavior. To the authors' knowledge, no comprehensive SA of TOPLATS has been performed and published so far. Thus, a detailed SA of the different hydrological processes calculated by TOPLATS could be a worthwhile contribution to improve the understanding and to facilitate the calibration of this model.

This study aims to evaluate TOPLATS as a streamflow simulation tool in Mediterranean catchments. This evaluation includes a detailed SA of TOPLATS model to identify influential parameters that should be included on a subsequent calibration/validation (CAL/VAL) approach, so that optimum streamflow simulation is achieved. This is done for three catchments of different sizes located on an area of Mediterranean climate, and considering different modeling time-steps. This broad objective expands to achieve the following specific objectives: (1) to provide a detailed review of previous works carried out with TOPLATS, specifically those related with model parameterization and calibration, (2) to develop a sensitivity analysis of selected parameters on: surface runoff, baseflow, evapotranspiration, soil moisture patterns and streamflow simulation (discriminating between peaks, average and low flows), (3) to compare two SA methods of different complexity and computational requirements, (4) to evaluate the performance of an optimization algorithm for model calibration at different time-scale simulations (daily and hourly), (5) to appraise the influence of continuous or random period selection for

calibration and validation purposes, and (6) to evaluate model streamflow simulation performance on Mediterranean catchments of varying size and climate conditions (at daily and hourly time-step).

2. Study sites

This study is focused on three catchments located in the province of Navarre (Northern Spain): La Tejeria, Cidacos and Arga (Fig. 1). They all belong to the Ebro river basin (86,000 km^2), one of the major Spanish rivers, which flows into the Mediterranean Sea. Catchments were selected with the objective of covering a range of catchment sizes, climate conditions, and contrasting topography. Navarre is divided into Atlantic (10% of the area) and Mediterranean basins (90%). In the latter, characterized by a sub-humid Mediterranean climate, rainfall decreases notably as one heads south. Study sites and their measuring stations location are shown in Fig. 1.

2.1. La Tejeria catchment

La Tejeria is a micro-catchment, part of the Agricultural Experimental Catchment Network developed and maintained by the Government of Navarre (http://cuencasagrarias.navarra.es/index.cfm). It has a total extension of 1.59 km^2, with a gauging station and an automatic meteorological station (10 min basis) installed at the outlet. Elevation ranges from 496 to 649 m. The main flow channel has a length of 1.9 km, and the average slope within the catchment is 12%. Climate conditions in this catchment are intermediate between the other two catchments. In La Tejeria, average rainfall during the 2000–2012 period was 744 mm/year, although the annual rainfall can vary significantly from year to year. During the study period, mean daily temperature was 12.5 °C and relative humidity 73%. The watershed is underlined by marls and sandstones of continental facies. The prevailing soil class is *Typic Calcixerepts* (NRCS-USDA, 2014), covering 41% of the watershed and located on eroded hillslopes. These soils are relatively shallow (0.5–1.0 m deep) and the upper horizon has a clayey-silty texture (Casalí et al., 2008). Catchment's major land uses are: winter cereals, sunflower, and fallow land (total 92%), riparian vegetation (7%) and urban areas (1%). Further details on the catchment's instrumentation, soil type, land use and hydrological behavior can be found in Casalí et al. (2008). Previous research works on this catchment focused on soil moisture retrievals (Alvarez-Mozos et al., 2006) and soil erosion evaluation (Casalí et al., 2008).

2.2. Cidacos catchment

Cidacos river flows into the Aragón river, one of the main tributaries of Ebro. At Olite hydro station Cidacos has a catchment area of 258 km^2. Elevation ranges from 380 m to 1156 m. Main river's length is approximately 25 km, and catchment's mean slope is 18% (35% of the catchment presents flat-gentle slope areas (<10%) and 19% of the catchment is characterized by steep slopes (>30%). Climate in the area is defined as mild-Mediterranean, with high rainfall variability that caused dramatic oscillations in the annual discharge between 0.4 and 39 hm^3, with an average value of 19.65 hm^3 in the historic observed series (since 1989). The

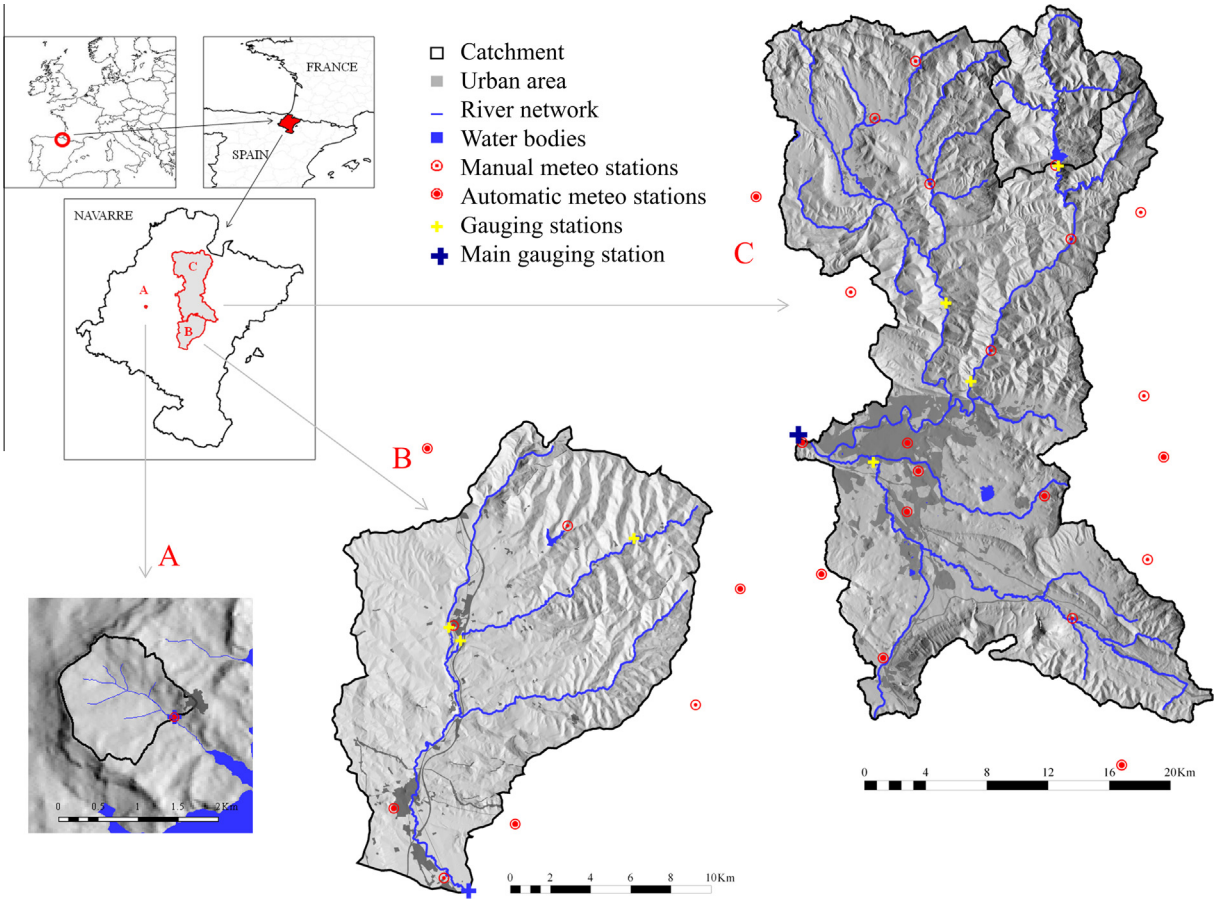

Fig. 1. Location, topography, hydrological features and instrumentation of the three studied catchments: (A) La Tejeria, (B) Cidacos, and (C) Arga.

annual mean precipitation was 639 mm. The catchment is equipped with 8 weather stations: 4 automatic stations (10 min basis) and 4 manual (daily). Cidacos has higher rainfall rates on the North-Eastern mountainous area (809 mm/year), which contrast with the low 453 mm/year measured in the South. Daily mean temperature is higher than in adjacent Arga catchment (12.6 °C), while relative humidity is lower (69%). Predominant soil type is *Typic Calcixerepts*, but *Typic Xerorthents* are also found in the North-Eastern mountainous area, and *Typic Xerofluvents* are predominant on the river network paths (NRCS-USDA, 2014). Predominant textures are clay-loam on accumulation hillslopes and silty-clay and loam on eroded hillslopes. Approximately half of the catchment is used for cereal cultivation (47%, including some minor irrigated areas), while the other half is divided into forest (27%) and dispersed bushes (22%) land covers. Urban areas account for the remaining 4%.

2.3. Arga catchment

Arga river is also one of the main tributaries of Ebro. Only its upper part has been used on this study, specifically the 810 km^2 defined by the gauging station at the municipality of Arazuri, being the main river channel 53 km long. Elevation descends from 1400 m to 400 m at the outlet, with a mean slope of 24%. Flat and gentle-slope areas (<10% slope) represent 29% of the catchment, and steep slopes (>30%) are found on 37% of the area. These slope values represent the steepest relief of the three studied sites. In this study, a 69 km^2 subcatchment feeding a reservoir (Eugi) in the Northern boundary of the catchment has been subtracted and both rainfall and streamflow measured at the reservoir's outlet were removed from the analysis, thus an effective catchment area of 741 km^2 was finally used. Historical streamflow records indicate an average annual contribution of 423 hm^3 at Arazuri station. Arga catchment is heavily instrumented with 20 measuring stations, but somehow limited in temporal resolution since most of them (12) are manual stations working on daily basis. The catchment's average annual rainfall measured for the 2000–2012 period was 956 mm. Average daily temperature and relative humidity were 11.7 °C and 76% respectively. In the southern half of the catchment predominant soils belong to *Aquic* and *Typic Xerorthent* groups of USDA's Soil Taxonomy (NRCS-USDA, 2014), and present a silty-clay-loam texture. Geologically this southern part is underlined by clay marls and Pamplona gray marls. In the Northern area of the catchment the prevailing soil classes are distributed according to the landscape's position. While steeper areas are mainly occupied by *Typic Xerorthents*, the valley plain by *Typic Haploxerepts*. Soils are fine and more than 1 m deep except for those in the eroded hillslope that are shallow. The predominant land covers in the catchment are forest (46%, mostly in the Northern part), rainfed cereal crops (33%), bushes (12%) and urban areas (10%, including the city of Pamplona).

3. Methodology

3.1. TOPLATS hydrological model

3.1.1. Model description

TOPLATS was developed as a water and energy balance model to be used at local and catchment scales. For that, a simple soil–vegetation–atmosphere transfer scheme (SVAT) was implemented onto a TOPMODEL framework (Famiglietti and Wood, 1994). This was afterwards improved by Peters-Lidard et al. (1997) to correct deficiencies in the representation of energy fluxes (i.e. soil evaporation and ground heat flux). Also, some additional modifications were carried out by Pauwels and Wood (1999), who adapted it to high latitude areas, and Crow et al. (2005) who tested an expansion of the model to a soil water balance of four layers, and separated soil and canopy contributions to evapotranspiration. The basic concept underlying the model is that shallow groundwater gradients, estimated from the local topography through a Topographic Index (TI) (Sivapalan et al., 1987), set up spatial patterns of soil moisture. Those patterns are considered key factors of simulation control on: (1) storm events: influencing infiltration and runoff generation and (2) inter-storm events: being responsible for evaporation and drainage patterns. TOPLATS incorporates a soil vegetation atmosphere transfer scheme (SVAT) to represent local scale vertical water fluxes within the catchment scale TOPMODEL approach.

TOPLATS can be run in either a fully distributed mode or in a statistical mode (Seuffert et al., 2002). In the fully distributed mode the catchment is subdivided into a grid of regular size cells, where each of those model units (cells) has its own specific: soil–vegetation parameterization, soil-topographic index value (TI) and meteorological input data (Pauwels et al., 2002). The land–atmosphere scheme is then applied to each cell. The second mode (statistical) has been used in this study. In the statistical mode, TI is represented through its statistical probability distribution given a fixed bin-size (Fig. 2), thus reducing the computational demand. This mode was developed under the similarity concept, that is, locations with the same TI and soil type, are assumed to respond similarly (Pauwels et al., 2002), performing a semi-distributed catchment representation.

In TOPLATS, the soil column is divided into two layers (Fig. 2): a thin surface zone (SZ) and the deeper transmission zone (TZ). Furthermore, the land surface is partitioned into bare and vegetated areas. Separate water balances (Fig. 2) are formulated for the different water reservoirs: the surface zone, the transmission zone, the water table and the canopy. Infiltration is calculated as the minimum of the soil infiltration capacity (Milly, 1986) and net precipitation. The exchange of soil water between the upper and lower layers is calculated assuming diffusive flux (Peters-Lidard et al., 1997), where diffusivity is given as a function of Brooks and Corey (1964) parameters. Evaporation is calculated with a soil resistance formulation (Passerat De Silans et al., 1986) as the minimum of a soil controlled and an atmospherically controlled evaporation rate. Similarly, canopy transpiration is calculated as the minimum of a plant and an atmospherically controlled transpiration rate, where the canopy resistance to transpiration (Jarvis, 1976) is a function of a minimum: (1) stomatal resistance, (2) LAI, and (3) stress factors (i.e. solar radiation, vapor pressure deficit, air temperature and soil moisture) (Jacquemin and Noilhan, 1990). Plant growth is not directly simulated by TOPLATS, but the seasonal development of plant properties is described by user defined time-step updates of plant parameters, i.e., leaf area index, plant height and stomatal resistance (Bormann, 2006a).

TOPLATS has been applied with different objectives, on a broad range of time and space-scale conditions (Table 1). Some researchers used the model to analyze the effects of land use changes and soil classification uncertainty at the catchment scale (Bormann et al., 2007; Crow and Wood, 2002; Loaiza-Usuga and Poch, 2009; Loosvelt et al., 2015, 2014a; Viney et al., 2005). The spatial resolution of input data, and its influence on different model outputs, such as water balances and flow components, were also investigated (Wood et al., 1988; Endreny et al., 2000; Bormann, 2006a, 2006b). TOPLATS has also been used for discrete observations up-scaling (Crow et al., 2005), local weather prediction (Seuffert et al., 2002) or crop growth analysis (Pauwels et al., 2007). Different types of remotely sensed information have been integrated with TOPLATS trough data assimilation (DA) procedures for improving streamflow simulation (Pauwels et al., 2002, 2001), soil moisture simulation (Crow et al., 2001; Houser et al., 1998; Lucau-Danila et al., 2005) or latent heat fluxes estimation (Crow and Wood, 2003).

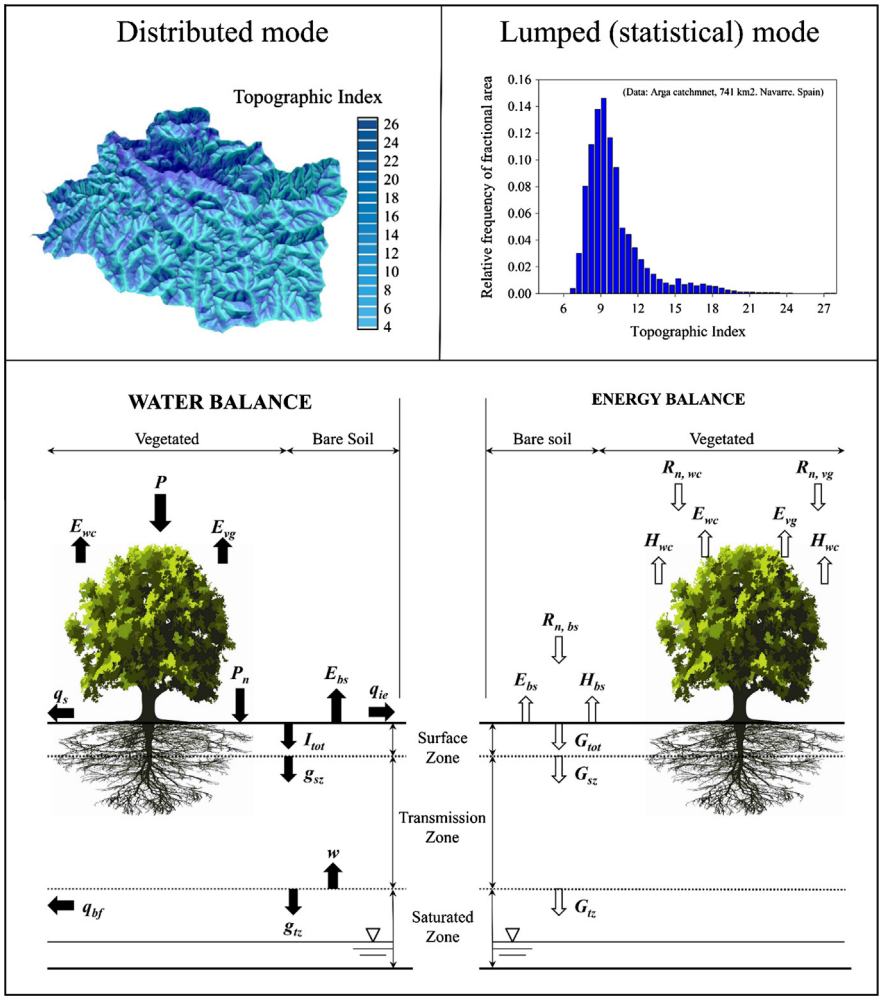

Fig. 2. TOPLATS model schematic representation: lumped and distributed Topographic Index representation, and SVAT balances (water and energy) of vegetated soil (*vg*), wet canopy (*wc*) and bare soils (*bs*). Water balance budgets: precipitation (P), net precipitation (P_n), evaporation (E), infiltration excess runoff (q_{ie}), saturation excess runoff (q_s) baseflow (q_{bf}), infiltration (I), drainage (g) and capillary rise (w). Energy balance budgets: net radiation (R_n), latent heat flux (H) and soil heat flux (G).

3.1.2. Previous sensitivity analysis and calibration studies

The model's behavior and performance has also been explored, for instance, to evaluate its sensitivity to soil parameters (Loosvelt et al., 2014b, 2011) or to analyze different calibration strategies (Goegebeur and Pauwels, 2007; Loaiza-Usuga and Pauwels, 2008; Pauwels et al., 2009). Table 1 presents a compilation of the most relevant works carried out with TOPLATS where parameter values were investigated with different objectives and degrees of complexity. On this study, parameters selected for TOPLATS sensitivity analysis and calibration were chosen principally according to information extracted from these works.

Calibration related researches (Table 1) were in some cases focused on obtaining optimal streamflow values at catchment outlet (Crow and Wood, 2002), while some others aimed to obtain accurate soil moisture simulations (Goegebeur and Pauwels, 2007). Pauwels et al. (2009) pointed out that the three most important soil parameters in the determination of the soil moisture content were the saturated hydraulic conductivity (K_s), the pore size distribution index (B), and the Bubbling pressure (ψ_c).

Bormann et al. (2007) proposed a calibration procedure based on first, reproducing the long-term water balance tuning vegetation parameters (stomatal resistances) and secondly, on optimizing the model efficiency by fitting the baseflow recession: adjusting baseflow at complete saturation (Q_0) and the Hydraulic conductivity decay (f) parameters. Seuffert et al. (2002) and Bormann (2006b) also worked on the manual calibration of Q_0, f and stomatal resistance (ST_r), while Goegebeur and Pauwels (2007) evaluated PEST and Extended Kalman Filter methods performance in K_s and B calibration. In Loosvelt et al. (2015) automatic calibration was performed based on a weight-adaptive recursive parameter estimation method, which is in concept a multi-start calibration approach.

3.1.3. Model set-up

TOPLATS requires the following seven climate variables as input: temperature (°C), relative humidity (%), atmospheric pressure (mm), wind speed (m/s), rainfall rate (m/s), longwave downward radiation (W/m^2) and shortwave downward radiation (W/m^2). All of them, except for longwave radiation (LR), were obtained from direct measurements at catchments' meteo stations; whereas LR was estimated from relative humidity (RH) and temperature (T) (Prata, 1996). Each station data was assigned to its corresponding percentage of catchment area, based on Thisessen polygons distribution. TOPLATS statistical mode was run on both daily and hourly time-scale. For hourly simulations, daily-measuring stations data were hourly distributed according to nearby hourly measurements. On the TOPLATS statistical mode, different land uses can be specified, but soil is considered homogenous. In La Tejeria catchment, five land uses were considered (winter cereals, sunflower, fallow land, riparian vegetation and urban areas). Land use input data

Table 1
Most relevant TOPLATS parameter calibration (*C*), Evaluation (*E*) and Uncertainty (*U*) analyses published in the past. The type of study, observed variables of interest, catchment or area size and parameters studied are given. Parameter abbreviations are as follows: Leaf Area Index (*LAI*), stomatal resistance (ST_r), crop height (*h*), Brooks–Corey Pore Size distribution Index (*B*), Bubbling pressure (ψ_c), saturated soil moisture (θ_s), residual soil moisture (θ_r), surface saturated hydraulic conductivity (K_s), first soil resistance parameter (r_{min}), subsurface flow at complete saturation (Q_D), Hydraulic conductivity decay (*f*), and initial water table (WT_i).

References	Type of study	Observed variable	Catchment or area size	Evaluated TOPLATS parameters													
				Vegetation			Soil							Baseflow		Water table	
				LAI	ST_r	*h*	*B*	ψ_c	θ_s	θ_r	K_s	r_{min}	Q_D	*f*	WT_i
Crow and Wood (2002)	Surface energy fluxes prediction	Soil moisture/streamflow/heat fluxes	575,000 km^2	E	E	E						E			
Seuffert et al. (2002)	Model coupling	Streamflow	2000/750/65 km^2										C	C	
Crow and Wood (2003)	Data assimilation	Latent heat flux	–	C									C	C	C
Crow et al. (2005)	Measurements upscaling	Soil moisture	6,400 km^2		E							E	E	E	
Lucau-Danila et al. (2005)	Parameter retrieval from SAR LAI	SAR data/soil moisture/plant height/(area)	18 agricultural fields	E		E					C		C	C	C
Bormann (2006b)	Spatial data resolution	Streamflow	134/81/63 km^2		C								C	C	
Goegebeur and Pauwels (2007)	Calibration	Soil moisture	2.26 km^2(area)				C				C				
Bormann et al. (2007)	Land use change	Streamflow	693 km^2		C		E	E	E	E			C	C	
Loaiza-Usuga and Pauwels (2008)	Calibration	Soil moisture	222 km^2(area)	E		E	C	C			C		E	E	
Pauwels et al. (2009)	Calibration	Soil moisture	250 km^2 (area)				C	C			C				
Loosvelt et al. (2011)	Soil hydraulic parameter (SHP) uncertainty	Soil moisture	point-scale				U	U	U	U	U				
Loosvelt et al. (2014a)	Prediction uncertainty	Streamflow	91 km^2										C	C	C

for Cidacos and Arga included nine land use types in order to differentiate irrigated from non-irrigated crops, and to distinguish different forest types and their degree of soil coverage. For each type, the parameterization required included: root depth, leaf area index, albedo, emissivity, stomatal resistance, etc. Most of those values were adapted from Crow and Wood (2002), Crow et al. (2005), Houser et al. (1998), Pauwels and Wood (1999), and Peters-Lidard et al. (1997). Specifically, LAI values for the different forest-types and crops, were obtained from Loaiza-Usuga and Pauwels (2008), Loosvelt et al. (2014a), and Lucau-Danila et al. (2005). Vegetation parameters were updated monthly, completing an annual cycle that applied for the whole simulation period (12 years). In La Tejeria, soil parameter information was obtained from available field observations (Casalí et al., 2008). Arga and Cidacos soil parameters were determined based on the soil texture class following (Rawls et al., 1982). A runoff-routing routine based on the unit hydrograph method proposed by the SCS – Soil Conservation Service (1972) was written in FORTRAN and added to TOPLATS.

3.2. Sensitivity analysis

A general description of SA methods is given in the introduction of this article, thus, this section extends only two relevant aspects: Design of Experiment sampling techniques (DoE), and the specific SA methods used in this study: Morris (1991) and Sobol (1993). Both Morris Method (MM) and Sobol Method (SM) follow approaches based on DoE techniques, which have a large influence on SA efficiency (Song et al., 2015).

Some SA methods require specific DoE sampling techniques designed ad hoc. This is the case of MOAT sampling for MM (Morris, 1991) and the sampling technique proposed by Saltelli (2002) (denoted as SOBOL) for SM (Gan et al., 2014). These sampling techniques are both based on simple random sampling, but different conditions need to be satisfied for their sample sizes (Song et al., 2015). A detailed description on both sampling methodologies can be found in Gan et al. (2014). Substantial differences in the number of required model runs (N_R) are found between both MOAT and SOBOL sampling methods. In the first case (Morris, 1991):

$$N_R = (k+1) * n \tag{1}$$

where *k* is the number of parameters whose sensitivity is evaluated and *n* is the number of samples.

On the other hand, Saltelli (2002) proposed two efficient approaches to reduce the computational cost of SM. The one applied on this study requires a larger number of model runs, but provides more consistent results:

$$N_R = [(2 * k) + 2] * n \tag{2}$$

3.2.1. Morris method

The screening-type Morris Method (Morris, 1991) takes advantage of elementary effects (EE) computed at evenly spaced values of each parameter over its entire range. The final effect is then calculated as the average of a set of partial effects. Therefore, it provides a global sensitivity measure with lower computational cost compared to most global methods. In this way, local sensitivities are integrated to a global sensitivity measure and the presence or absence of nonlinearities or correlation interactions with other parameters can also be identified (Van Griensven et al., 2006). Screening methods, such as MM, aim to identify the subset of non-influent factors in a model using a small number of model runs (Campolongo et al., 2011).

MM has a low computational cost, is simple to implement and its results are easy to interpret (Saltelli et al., 2000) but individual

interaction between parameters cannot be detected, since MM only calculates the overall interaction of a parameter with the rest (Saltelli et al., 2000). It measures qualitatively relative sensitivity by ranking input parameters in order of sensitivity but cannot quantify in absolute terms how much one parameter is more important than another (Campolongo and Saltelli, 1997).

Considering that $y(X)$ is the objective function (goodness-of-fit of the model or model output of interest) and $X = (X_1, \ldots, X_k)$ is the parameter set:

$$y(X) = f(X_1, \ldots, X_k) \tag{3}$$

in MM, $y(X)$ is calculated for each parameter set, where parameter values (X) are changed OAT. The difference resulting of both $y(X)$ values is divided by the variation of perturbed parameter values (Δ) to obtain the elementary effect of each parameter $d_i(X)$. Elementary effects are calculated as follows:

$$d_i(X) = \frac{[y(X_1, X_2, \ldots, X_{i-1}, X_i + \Delta, X_{i+1}, \ldots, X_k) - y(X)]}{\Delta} \tag{4}$$

This calculation process is repeated until the defined number of samples (n) are completed for each parameter. Finally, the mean (μ) and standard deviation (σ) values of d_i for the n samples are used as sensitivity indices, where μ and σ indicate the influence of each parameter on the objective function (Shin et al., 2013). A high μ value indicating an important overall influence of the parameters and a high value of σ (hereafter referred to as Morris interaction) meaning strong interactions with other parameters or the effect of nonlinearities (Shin et al., 2013). Instead of using μ, this study uses the mean of the absolute values of the n samples of d_i, denoted as μ^* (hereafter referred to as Morris sensitivity) to overcome the problem of the effects of opposite signs due to a models' non-monotonic characteristics (Campolongo et al., 2011). Further description of Morris method and its implementation for specific purposes can be found in Francos et al. (2003), van Griensven et al. (2006), Shin et al. (2013), and Wainwright et al. (2014).

3.2.2. Sobol method

SM is a variance decomposition global SA method where all parameters are varied simultaneously. One of its main advantages is that it is a model-free method, i.e. it can compute sensitivity indices regardless of the linearity, monotonicity (or other generic assumptions) on the underlying model (Baroni and Tarantola, 2014). In SM the variance of the model output is decomposed into fractions that result from either individual parameters or parameter interactions. The sensitivity of each parameter or parameter interaction is then assessed based on its contribution (measured as a percentage) to the total variance computed using a distribution of model responses (Zhang et al., 2013). Sobol sensitivity indices have been shown to be more effective than other approaches in capturing the interactions between a large number of variables for highly nonlinear models (Tang et al., 2007).

As in Eq. (3), considering that $y(X)$ is the objective function (goodness-of-fit of the model or model output of interest) and $X = (X_1, \ldots, X_k)$ is the parameter set, the total variance of function $y(X)$, $D(y)$, is decomposed into component variances from individual parameters (D_i) and their interactions ($D_{ij}, D_{ijk}, \ldots$):

$$D(y) = \sum_i D_i + \sum_{i<l} D_{ij} + \sum_{i<j<l} D_{ijl} + \cdots + D_{12\ldots k} \tag{5}$$

where D_i represents the average reduction in variance achieved if the parameter i was known (Ratto et al., 2001) and D_{ij} is the amount of variance due to the interaction between parameter X_i and X_j. The single parameter sensitivity (First-Order Sobol index), S_i, and parameter interaction (Second-Order index), S_{ij}, are then assessed based on their relative contribution to the total variance (D):

$$S_i = \frac{D_i}{D} \tag{6}$$

$$S_{ij} = \frac{D_{ij}}{D} \tag{7}$$

Also, a third index (Total-Order index), S_{Ti}, that measures the main effect of X_i and its interactions with all the other parameters (Zhang et al., 2013) can be calculated as:

$$S_{Ti} = 1 - \frac{D_{\sim i}}{D} \tag{8}$$

where $D_{\sim i}$ is the bottom marginal variance, that accounts for the amount of variance due to all of the parameters except for X_i (Massmann and Holzmann, 2012). Hereafter in this article, S_i is referred to as Sobol sensitivity and S_{Ti} as Sobol interaction. SM main characteristic is the use of two different sets of samples, generated by the same scheme and with the same number of elements. SM uses the first set to calculate the overall output mean and variance (i.e., the combined effects of all parameters) while the second sample is then used to resample each parameter, rather than setting each to a fixed value, for the calculation of total and individual variance contributions (van Werkhoven et al., 2008). D_i and $D_{\sim i}$ can be calculated as described in Massmann and Holzmann (2012) and Zhang et al. (2013). In the last years SM has been used frequently in different types of hydrological models (Gan et al., 2014; Massmann and Holzmann, 2012; Shin et al., 2013; van Werkhoven et al., 2008; Wainwright et al., 2014).

3.2.3. SA implementation

The sensitivity analysis was carried out using SimLab software (EC-JRC, 2008). The SA analysis included three steps: (1) sample generation (DoE), (2) model execution and (3) model's results analysis (referred to as Statistical post processor in SimLab). Seven parameters (i.e. $k = 7$) (Table 2) were selected for the SA based on previous studies (Table 1) and manual testing. MM required 80 runs, since $n = 10$ samples were considered. In contrast, SM required 4096 runs since $n = 256$ samples were used. As this study aimed to evaluate TOPLATS in all its complexity, model sensitivity was evaluated in terms of nine variables and efficiency measures, grouped in: (1) main hydrological processes (surface runoff, baseflow and evapotranspiration), (2) soil moisture behavior (mean and

Table 2
TOPLATS parameters included on Morris and Sobol sensitivity analysis. Type, symbol, units and parameter value ranges are given.

	Type	Parameter	Symbol	Units	Min limit	Max limit
1	Soil properties	Brooks–Corey Pore Size distribution Index	B	–	0.1	1.0
2	Soil properties	Bubbling pressure	ψ_c	m	0.1	1.0
3	Soil properties	Saturated soil moisture	θ_s	cm^3/cm^3	0.4	0.6
4	Soil properties	Surface saturated hydraulic conductivity	K_s	m/s^1	$1e^{-07}$	$1e^{-03}$
5	Soil properties	First soil resistance parameter	r_{min}	s/m^1	4000	80,000
6	TOPMODEL	Subsurface flow at complete saturation	Q_0	mm/day^1	5	175
7	TOPMODEL	Hydraulic conductivity decay	f	m^{-1}	1	14

standard deviation of surface zone) and (3) goodness-of-fit (efficiency) measures (detailed in Section 3.3).

3.3. Calibration–Validation procedure

3.3.1. Optimization algorithm

In this work, a Multi-Start approach of the Powell Method (MSPM) was used to calibrate the model. Powell Method (PM) (Powell, 1964) belongs to the category of local and derivative-free algorithms. PM is a conjugate directions method, which are capable of minimizing quadratic functions in a finite number of steps. Since a general nonlinear function can be approximated reasonably well by a quadratic function near its minimum, this type of conjugate directions method is expected to speed up the convergence of even general nonlinear objective functions (Rao, 2009). PM can find a global optimum when specifically tuned to a certain objective function, but in cases of sophisticated algorithm optimization (such as complex hydrological models), the selection of appropriate parameter starting points may affect algorithm performance substantially. Inadequate selection of initial parameter values may lead to convergence to inferior local optima or even to numerical dispersion (Paik et al., 2005). All in all, PM has been frequently applied on different hydrological calibration procedures and models (Chen et al., 2005; Geem and Roper, 2010; Yang et al., 2011; Zhang and Lindström, 1997).

In our study, to avoid the algorithm ending at local minima constraint, within each parameter's search range different values were selected as initialization values for the search. Dispersed values covering the full feasible parameter range were selected. The algorithm was free-available and programmed in FORTRAN (Press et al., 1988), the same programming language that was used to develop TOPLATS which facilitated the integration of both tools.

3.3.2. Goodness-of-fit

In this study, four measures of agreement were used for SA, calibration and validation. The first was the relative volumetric error (Pbias), which represents the percent volume difference between simulated and observed streamflow fluxes. As such, negative biases correspond to model under-estimation and positive biases to over-estimation (Khakbaz et al., 2012). It is calculated as follows:

$$Pbias(\%) = 1 - \frac{\sum_{i=1}^{n}(Q_{sim}(i) - Q_{obs}(i))}{\sum_{i=1}^{n} Q_{obs}(i)} * 100 \quad (9)$$

The second measure of agreement used was the Nash & Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), which measures the fraction of the observed streamflow variance explained by the model, calculated as the relative magnitude of the residual variance (noise) to the observed variance (information). Its optimal value is 1.0 and values should be larger than 0.0 to indicate a "minimally acceptable" performance (Yapo et al., 1996). NSE is calculated as follows, with exponent value $k = 2$.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{obs}(i) - Q_{sim}(i))^{k}}{\sum_{i=1}^{n}(Q_{obs}(i) - \overline{Q_{obs}})^{k}} \quad (10)$$

The last two measures used (NSE_1 and $NSE_{0.5}$) are simple modifications of the original NSE equation, designed for the analysis of specific streamflow ranges. In NSE_1, $k = 1$ and in $NSE_{0.5}$, $k = 0.5$. In NSE, peak flows simulation accuracy is prioritized, whereas NSE_1 focuses on average flows and $NSE_{0.5}$ on low flows.

3.3.3. CAL/VAL implementation

Previously, different studies discussed the advantages of different random period selection techniques (RND) for model calibration and validation (Brath et al., 2004; Kim and Kaluarachchi, 2009; Senarath et al., 2000). Some of them evaluated data length requirements when RND approach is applied, compared to the conventional (CON) approach, i.e. Kim and Kaluarachchi (2009) concluded that randomly sampled data required a shorter calibration length (36 months) than continuous data (120 months) to reach good model performance. Similarly, Perrin et al. (2007) tested the random approach with two daily rainfall-runoff models showing that, in general, even just 350 calibration days sampled out of a longer data set including dry and wet conditions could be sufficient to obtain robust estimates of model parameters. On the conventional calibration approach, Yapo et al. (1996) estimated a minimum of 96 months for adequate model calibration and also noted that parameter uncertainty was reduced when wettest data records were used.

This study, following referenced studies recommendations on CON and RND period selection, considered a 12 years period, and included two calibration and validation (CAL/VAL) period selection strategies in order to evaluate the influence of climatic particularities on the CAL/VAL results. First, a "conventional" strategy (CON) was followed, where the first hydrological year was used for warming-up (2000–2001), the next six years for calibration (2001–2007), and the remaining five years for validation (2007–2012). Secondly, a "random" strategy (RND) was proposed and evaluated. On this second approach, each month's data was randomly assigned to either CAL or VAL period according to two conditions: (1) the mean streamflow of CAL and VAL series could not differ more than 10% and, (2) their standard deviations could not differ more than 15%. This second strategy was adopted to circumvent the imbalance between CAL and VAL periods resulting from irregular climatic conditions typical of Mediterranean areas. In total, 72 months were randomly assigned to CAL and 60 months to VAL, maintaining the same proportion as in the conventional strategy.

A four steps calibration (Fig. 3) was then performed exactly in the same way for both approaches (CON and RND): (1) manual calibration of the global water balance by adjusting Leaf Area Index (LAI) and Initial water table depth (WT_i), and (2) MSPM of six parameters (Table 2) on 8 different start points. On the latter, the algorithm was setup to maximize NSE_1. Optimization of NSE_1 was assumed to improve model efficiency while maintaining water balance on optimal performance. In a final step, out of the 8 sets of optimal parameter (OP) values obtained, the one offering best results (lower Pbias and higher NSE) was taken and used as a new initial set. Then, a new pair of refinement optimizations (steps 3 and 4) were carried out: run number 9 was set up again with NSE_1 as objective function (with the objective of finding possible better results with parameter values close to selected OP) and run number 10 using NSE as objective to prioritize high flows simulation improvement.

4. Results and discussion

4.1. Sensitivity analysis

Plots presented in this section (Figs. 4–6) show results from MM (left) and SM (right) analysis of daily TOPLATS simulations. Output variables are shown in the *y*-axis for the three catchments and TOPLATS parameters in the *x*-axis. Circle size represents individual parameter sensitivity (μ^* in MM and S_i in SM) while blue color ranges identifies total-order sensitivity results (σ in MM and S_{Ti} in SM), which includes parameter interaction.

4.1.1. Hydrological processes

SA results indicated that f had the overall largest influence on hydrological processes sensitivity. It was the main parameter

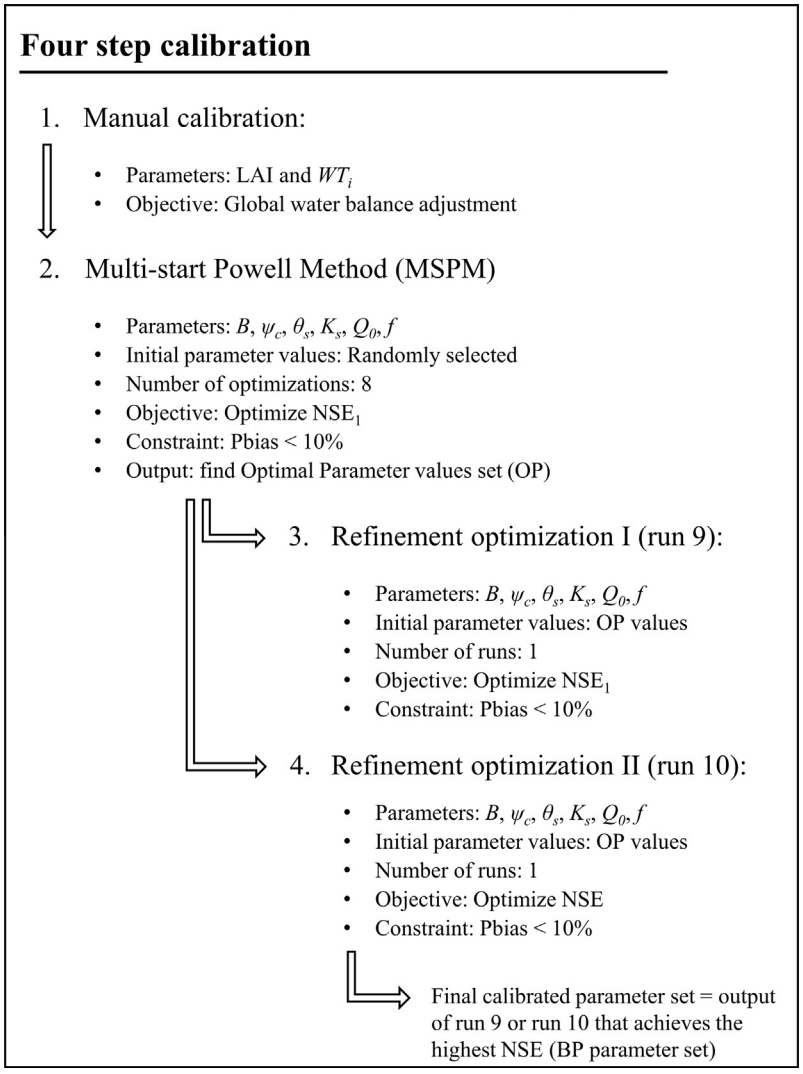

Fig. 3. Four step calibration procedure scheme.

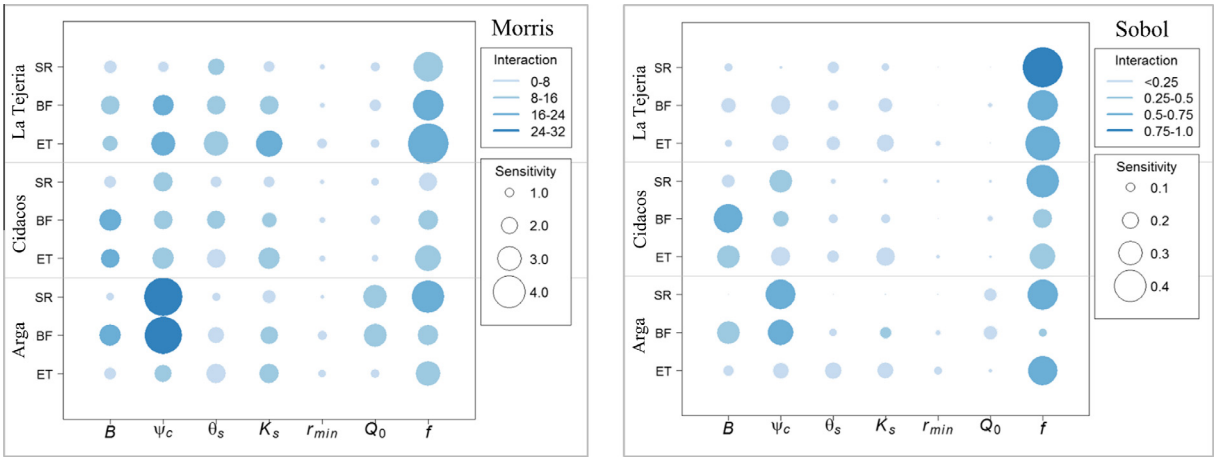

Fig. 4. Sensitivity analysis results of main hydrological processes (surface runoff, baseflow and evapotranspiration) using Morris (left) and Sobol (right) methods. Parameter sensitivity (μ^* and S_i) and interaction (σ and S_{Ti}) are shown. MM results expressed in mm/day and SM results expressed as a decimal of the total variance.

responsible in La Tejeria processes, and shared responsibility with B in Cidacos and ψ_c in Arga.

Both methods, MM and SM identified parameter f as the one having the largest influence on surface runoff (*SR* in Fig. 4) generation in La Tejeria. Parameter f was also the most influential in Cidacos, where the second most sensitive was ψ_c. MM and SM also agreed in Arga, identifying f and ψ_c as the most influential parameters, but here also Q_0 seemed to have some relevance too

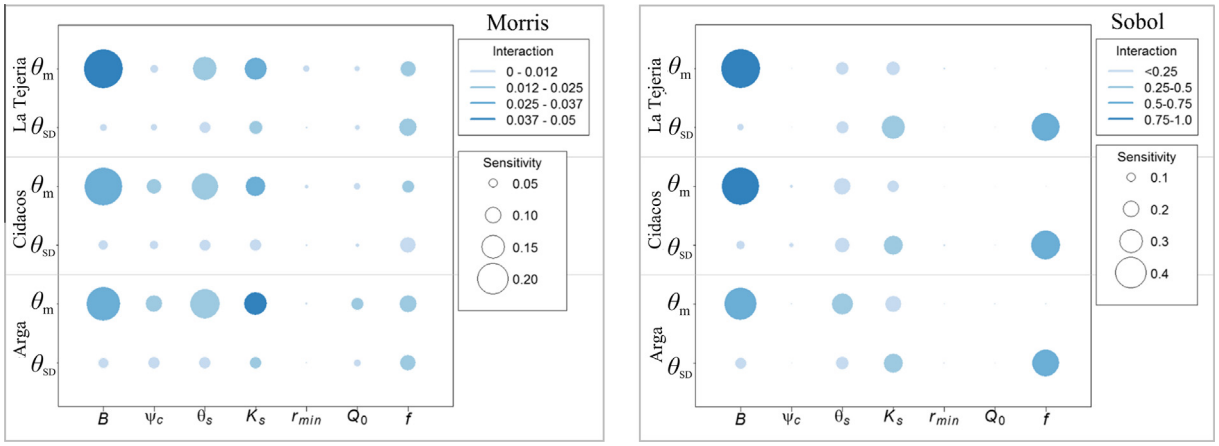

Fig. 5. Sensitivity analysis results of Surface Zone (SZ, 5 cm) soil moisture processes (mean value θ_m, and standard deviation, θ_{SD}) using Morris (left) and Sobol (right) methods. Parameter sensitivity (μ^* and S_i) and interaction (σ and S_{Ti}) are shown. MM results expressed in mm^3 mm^{-3} and SM results expressed as a decimal of the total variance.

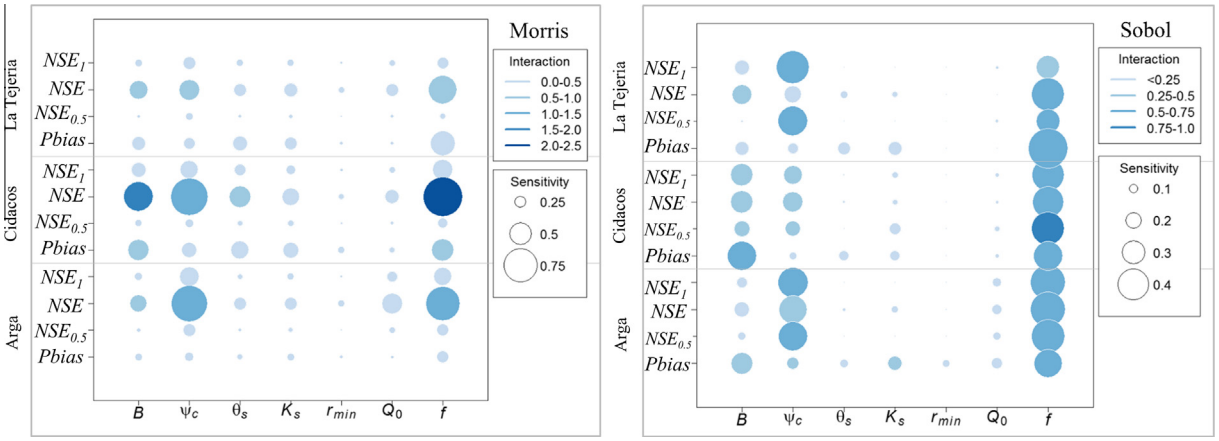

Fig. 6. Sensitivity analysis results of model efficiency (NSE_1, NSE, $NSE_{0.5}$ and Pbias) using Morris (left) and Sobol (right) methods. Parameter sensitivity (μ^* and S_i) and interaction (σ and S_{Ti}) are shown. MM results expressed in NSE coefficient value and SM results expressed as a decimal of the total variance.

according to MM. Some differences were found among catchments in terms of baseflow (*BF* in Fig. 4) sensitivity, as different parameters appeared as the most influential: *f* in La Tejeria, *B* in Cidacos and ψ_c in Arga. Strictly speaking ψ_c shared its relevance with *B* on Arga according to SM. All in all, these 3 parameters were the most significant in all catchments. Results obtained by SM depicted stronger differences between parameters in terms of baseflow sensitivity, whereas MM did not show significant differences among most parameters. Evapotranspiration (*ET* in Fig. 4) outputs were clearly sensitive to *f* in La Tejeria and Arga, while in Cidacos up to four parameters shared a similar influence degree: *B*, ψ_c, *f* and K_s. In terms of parameter interaction for the three processes, ψ_c and *f* had larger interaction levels. This was particularly notorious for ψ_c in Arga, according to MM, and for *f* in all catchments according to SM.

When comparing the results obtained with MM and SM, in most cases both methods came up with the same set of most influential parameters, but some results showed different patterns. When MM gave four (out of 7) parameters a similar degree of sensitivity in many cases, SM was able to identify more clearly the most influential pair of parameters. This discrepancy on secondary level parameters identification was clearly observed when comparing parameters *B* and ψ_c with parameters θ_s and K_s. Despite having similar individual sensitivities in MM, there were relevant differences among them on SM. Thus, SM was able to differentiate more clearly secondary parameters, especially in *BF* and *ET* sensitivity analysis.

4.1.2. Soil Moisture

Soil moisture sensitivity analysis results did not show relevant differences between MM and SM despite their different computational requirements (Fig. 5). Both were capable of clearly identifying *B* as the main parameter affecting surface zone soil moisture mean value (θ_m) outputs. Regardless of some minor disparities on secondary parameters estimation, both agreed on the estimation of θ_s and K_s as the second and third parameter in terms of individual sensitivity. Despite differences in catchment size, θ_m sensitivity patterns for the three studied catchments were found to be very similar. Pauwels et al. (2009) concluded that K_s, *B* and ψ_c were the three most important soil parameters on the determination of soil moisture content. These results are similar to the ones obtained here, but differed on the role attributed to ψ_c and θ_s. On a SA following a completely different approach, Loosvelt et al. (2014b) found that TOPLATS soil moisture output was mostly influenced by the residual soil moisture content (θ_r), θ_s, ψ_c, *B* and K_s. In their study, Loosvelt et al. (2014b) were able to identify areas within the parameter range with different level of sensitivity, related to the shape of the soil moisture retention curve. They found the highest sensitivity for low values of *B*, ψ_c, θ_s, and for high θ_r values.

Regarding soil moisture dynamics (i.e. surface zone soil moisture standard deviation (θ_{SD}), Fig. 5) f was identified as the most influential, followed by K_s and θ_s. These results were clearer with SM, which seemed to take advantage of its more detailed setup, so primary, secondary and tertiary parameters could be more plainly distinguished from non-influential ones. No significant differences were found among catchments in terms of θ_{SD} sensitivity. As shown in Fig. 5, B and K_s presented the highest level of interaction with other parameters on model θ_m output, and f in soil moisture dynamics.

4.1.3. Model efficiency

In this section, sensitivity results of four efficiency measures are described. These results offered notable differences between catchments, where diverse climate regimes seemed to affect remarkably efficiency results (Fig. 6). Furthermore, differences between MM and SM were larger than in the previous analyses.

As a general inference of the SA of efficiency measurements, it can be stated that 3 parameters appeared to be responsible for the largest fraction of model sensitivity to medium, high and low flow simulation (Fig. 6): Brooks–Corey Pore Size distribution Index (B), Bubbling pressure (ψ_c) and Hydraulic conductivity decay (f). An important remark extracted from this analysis was the similarity of the wettest catchments (i.e. Arga and La Tejeria), where ψ_c was globally more influential than in Cidacos, where B gained importance. Few discrepancies were found among methods, but for instance when ψ_c was the second most influential parameter in NSE in Cidacos according to MM, that place was taken by B according to SM.

NSE_1, (i.e. efficiency for average flows) was mainly controlled by ψ_c and f in La Tejeria and Arga (Fig. 6). In Cidacos, in addition to those two parameters, MM and SM identified B as an important parameter as well. As in previous analyses, SM detected more clearly non-influential parameters (e.g. θ_s or r_{min}). Regarding efficiency of high flows simulation (i.e. NSE, Fig. 6) results could be summarized as follows: f being the key factor affecting NSE in La Tejeria, whereas B, ψ_c and f shared similar individual effects in Cidacos and ψ_c and f accounted for most of the sensitivity in Arga. Sensitivity to low flow simulation (i.e. $NSE_{0.5}$, Fig. 6) was mostly similar to NSE_1. Lastly, Pbias was primarily affected by f in La Tejeria and by B and f in Cidacos and Arga.

Regarding parameter interaction of the four efficiency measures, f was the parameter that showed larger interaction with the rest. In general, it could be observed (Fig. 6) that B, ψ_c, and f shared the highest levels of interaction. In La Tejeria and Arga those interactions affected mainly ψ_c and f. On the other hand, those interactions were found to be especially intense between B and f in Cidacos.

4.1.4. SA global remarks

Previous studies commented on weakness and advantages of MM and SM for hydrological models analysis (Shin et al., 2013). According to Campolongo and Saltelli (1997) there could be a discrepancy between the identification of parameters with both algorithms. However, in this study both methods generally agreed, and identified the same set of most influential parameters on surface runoff (f and ψ_c), baseflow (f, B and ψ_c) and evapotranspiration processes (f and four other secondary parameters, B, ψ_c, θ_s, K_s). They also showed agreement on finding the most sensitive parameter responsible of mean (B) and soil moisture variation values (f). However, some qualitative differences were found on the SA of efficiency.

Parameters identified here as the most influential have a clear physical meaning, but some of them, particularly soil parameters, participate on the calculations of several water fluxes, which complicates the understanding of their precise influence on each model output. In any case, parameters Q_0 and f are, together with the varying WTD, the controlling factors of the generated baseflow amount. As observed from the specific formulation detailed in Famiglietti and Wood (1994), while Q_0 is responsible of the magnitude of the generated baseflow, parameter f defines the shape of the evacuation flow, leading to faster (low f values) or slower (higher f) evacuation of subsurface flow from saturated areas in the catchment. Parameter f controls this way the availability of water in the soil for other processes, such as evapotranspiration, being thus the main responsible of water balance patterns in the catchments (including θ_{SD}).

As mentioned above, the interaction of parameters gets more complex when soil parameters are evaluated. In TOPLATS, the soil is modeled through the equations of Brooks and Corey (1964), that numerically calculate the soil moisture content θ_m depending on B, ψ_c, θ_r, θ_s and the matric head (ψ) (Famiglietti and Wood, 1994; Loosvelt et al., 2014b). As it was observed from the SA results, B was the most influential parameter on θ_m. This parameter is the exponent value regulating the θ_m equation in TOPLATS. Higher B values will thus lead to higher mean soil moisture conditions. Also, larger θ_s and ψ_c values will increase soil moisture mean content.

Parameter ψ_c plays an important role on the partitioning between infiltration and runoff generation. This parameter's value defines a discontinuity in TOPLATS when the matric head value (ψ) reaches ψ_c. When this point is reached (usually during the wettest winter days), saturation excess runoff is activated and soil moisture and conductivity values are modified accordingly. In any case parameter interactions in TOPLATS are complex, as diffusive flux from SZ to RZ is also controlled by K_s, θ_s, θ_r, and B (Peters-Lidard et al., 1997) and the same parameters take part on the drainage calculations as well (Famiglietti and Wood, 1994).

Globally, differences among methods can be numerically summarized as follows: (1) in most of the cases evaluated in this study (81%) MM and SM identified the same parameter as the most influential, being this agreement particularly strong on soil moisture SA, but also on other main hydrological processes and on efficiency and Pbias; (2) in 67% of the cases tested both methods agreed at identifying the same pair of most influential parameters; and (3) in the identification of the three most influential parameters the agreement of both SA techniques was higher on soil moisture (83%), and hydrological processes (78%), but lower on efficiencies and Pbias (50%). This results are in agreement with Wainwright et al. (2014), who presented another comparison of MM and SM and concluded that both methods provided consistent parameter importance rankings when used on a reservoir-aquitard-aquifer model. However, as also investigated by Gan et al. (2014), clearer qualitative differences between secondary parameters were provided by SM.

4.2. Model Optimization

4.2.1. MSPM performance

TOPLATS calibration dimensionality was reduced to just 6 parameters thanks to SA analysis that identified parameter r_{min} as uninfluential to most efficiency and hydrological model outputs, so it was not included on the MSPM calibration. In order to allow for comparison of optimization algorithm performance on the different catchments, Table 3 shows NSE_1, NSE and Pbias median values (of the 10 optimization runs) at initial (IP) and final optimal points (OP). These results clearly identified Cidacos as the catchment where MSPM improved TOPLATS performance the most. IP efficiency values were very low in Cidacos, indicating that parameter combinations far from the optimal values had a stronger influence on efficiency results than in the other two catchments. In La Tejeria and Arga, most IP value sets, even the farthest from the final OP values, offered positive NSE values. For Cidacos, extremely large

Table 3
Improvement of efficiencies and Pbias reduction achieved by MSPM on La Tejeria, Cidacos and Arga catchments. Values presented on the table are median values of the 10 (8 MSPM and 2 refinement) optimization runs obtained with the Initial Parameter (IP) values and Optimal Parameter (OP) values.

Catchment	NSE_1		NSE		Pbias (%)	
	IP	OP	IP	OP	IP	OP
1. La Tejeria	0.40	0.68	0.27	0.82	15	3
2. Cidacos	−0.41	0.50	−1.51	0.57	112	4
3. Arga	0.37	0.49	0.38	0.57	5	3

Pbias (112%) values were obtained at IP points. This value dropped to 4% after optimization, showing that using NSE_1 as the optimization criteria also resulted in a well-balanced volume simulation. These strong Pbias reductions in Cidacos contrasted with the more modest Pbias improvements in the other two catchments (12% and 2% in La Tejeria and Arga, respectively) where IP Pbias values were already moderate or even low. Optimization results showed that TOPLATS offered a more stable behavior on wetter catchments (i.e. Arga), where even before calibration, results were not far from optimal ranges. On the other hand, efficiency results before optimization (IP) showed large variability in La Tejeria, and extremely large in Cidacos (not shown). Better results on wetter catchments were, somehow, expected due to discontinuities that may appear between saturated areas and part of their upstream grid cells. Thus, those cells do not contribute to subsurface flow during the drier seasons, especially on Cidacos catchment.

4.2.2. Optimal parameter values

Optimal (OP) and Best performing parameter values (BP) found by MSPM shown in Fig. 7. For simplification and to allow catchment behavior comparison, only CON approach values are shown in this section. Parameter value ranges to be explored by the algorithm were defined according to references in previous TOPLATS works (Goegebeur and Pauwels, 2007; Loaiza-Usuga and Pauwels, 2008; Peters-Lidard et al., 1997).

First parameter evaluated was B, whose BP values for La Tejeria and Arga were found at values close to one, while in Cidacos it was substantially lower (0.41). Regarding ψ_c, despite MSPM ended on some local minima in some of the optimization runs, in all cases the BP ψ_c value was within the 0.1–0.4 m range. In general, in the three catchments, highest efficiencies were found as a combination of high B and low ψ_c values (despite some optimizations ending at the opposite combination, low B and high ψ_c). Saturated soil moisture (θ_s) yielded different best performing values in La Tejeria, with a low optimal θ_s, and Cidacos and Arga, where optimal θ_s was over 0.55 m^3/m^3.

K_s, had its BP within $1e^{-4}$ and $1e^{-5}$ m/s for La Tejeria and Arga, but in Cidacos these values were higher (between $1e^{-3}$ and $1e^{-4}$ m/s) (Fig. 7). Differently from some other parameters (i.e. ψ_c or f), optimal K_s values were always located on a small range of the defined parameter search space. Subsurface flow at complete saturation (Q_0) optimal values increased as catchment area increased. Similarly to K_s, optimal Q_0 values (OP) were clearly identified on specific areas (value range). Finally, for the decay parameter (f), which is key for event recession flow adjustment, OP and BP values are presented. Parameter f has a strong parameter interaction that resulted in similar efficiency results with varied f values, but BP values were close to 8.0 in La Tejeria and Arga. In Cidacos lower f values were found, which is related to a more rapid decrease of baseflow in this catchment.

BP values obtained in this study are in accordance with parameter values reported in other studies, but significant differences were found on K_s values. Other TOPLATS studies, where parameter values were detailed include Loaiza-Usuga and Pauwels (2008), who fixed some parameter within the following ranges: f (0.01–1.2 m^{-1}) and Q_0 (2.4–61 mm/day), and calibrated three parameters: K_s, obtaining an optimal mean of $3.8e^{-6}$ m/s; B, with a mean calibrated value of 0.72; and ψ_c, with a set of optimal values with 0.7 m as mean. used, after calibration, $f = 2.5$ m^{-1}, and $Q_0 = 6.0$ mm/day on a 91 km^2 catchment. Ranges of calibrated parameter values reported in Pauwels et al. (2009) for different soil types were: B (0.47–0.65), ψ_c (0.35–0.45 m) and K_s ($2.8e^{-6}$–$3.5e^{-6}$ m/s). Parameter values presented in Loosvelt et al. (2011), for different soil types were in the following ranges: B (0.15–0.69), ψ_c (0.2–0.94 m) and K_s ($5.8e^{-7}$–$1.1e^{-5}$ m/s). As it has been extracted from our SA study, large parameter interaction allows obtaining similar optimal efficiencies with substantially different parameter values combination in TOPLATS.

4.3. Comparison of conventional (CON) and random (RND) calibration approach on CAL/VAL results

TOPLATS daily efficiencies and Pbias results after CAL and VAL for both CON and RND strategies are presented in Table 4. Simulations of La Tejeria offered similar NSE results when CON and RND strategies were applied, although best efficiency in calibration was obtained with CON (0.82). Similarly, validation NSE results were over 0.7 in this catchment for both strategies and Pbias results were low, below 5% in all cases. Results for Arga were similar, with little differences between CON and RND strategies. In Arga, the conventional approach provided more stable NSE results in CAL (0.63) and VAL (0.60), whereas variation was larger in RND approach (0.71/0.54). Total simulated volumes were close to the observed, with Pbias lower than 5% on calibration and 10% on validation.

In Cidacos, due to the irregularity of its rainfall, extreme events seemed to have a large influence on the NSE. CON calibration achieved a NSE_1 of 0.53 and a corresponding NSE of 0.61 with a low Pbias (2%). But, in this case, validation results were poor, due to an abnormal decrease in simulated water table depth (WTD) that caused a strong streamflow underestimation (77%). The validation period was notably drier than the calibration one, and this affected the results critically. The RND approach overcame this issue and improved results substantially in Cidacos reaching a NSE_1 of 0.54 and a NSE of 0.91 for the calibration period. Validation yielded then (RND) efficiency values of 0.39 and 0.25 (for NSE_1 and NSE, respectively) and a remarkable reduction in Pbias from 77% to 3%. BP parameter values identified by this second approach helped TOPLATS to perform in a more consistent way. The deficiency of this method was that it was not able to properly simulate extreme flows during validation, but this was principally caused by just 3 single events (Fig. 11), where TOPLATS underestimated streamflow notably.

TOPLATS streamflow simulation efficiency values obtained here are similar to those reported by Bormann et al. (2007), on a 693 km^2 wet catchment, with a daily NSE results of 0.66 and 0.61 for CAL and VAL respectively. Two parameters were manually calibrated on that study, reaching a simulation Pbias below 3%. Bormann et al. (2007) also noted the influence that hydroclimatic conditions had on model's CAL/VAL results. This author signaled that large differences on evapotranspiration, precipitation and discharge among CAL/VAL periods limited the possibility to achieve proper CAL/VAL results. The RND strategy applied here guarantees that enough wet weather data is included on the calibration period, leading to better model performance. This conclusion was also noted by Anctil et al. (2004). Also in accordance with our results in mediterranean catchments, Kim and Kaluarachchi (2009) found that the number of high-flow months included on the calibration data had a great influence on model efficiency.

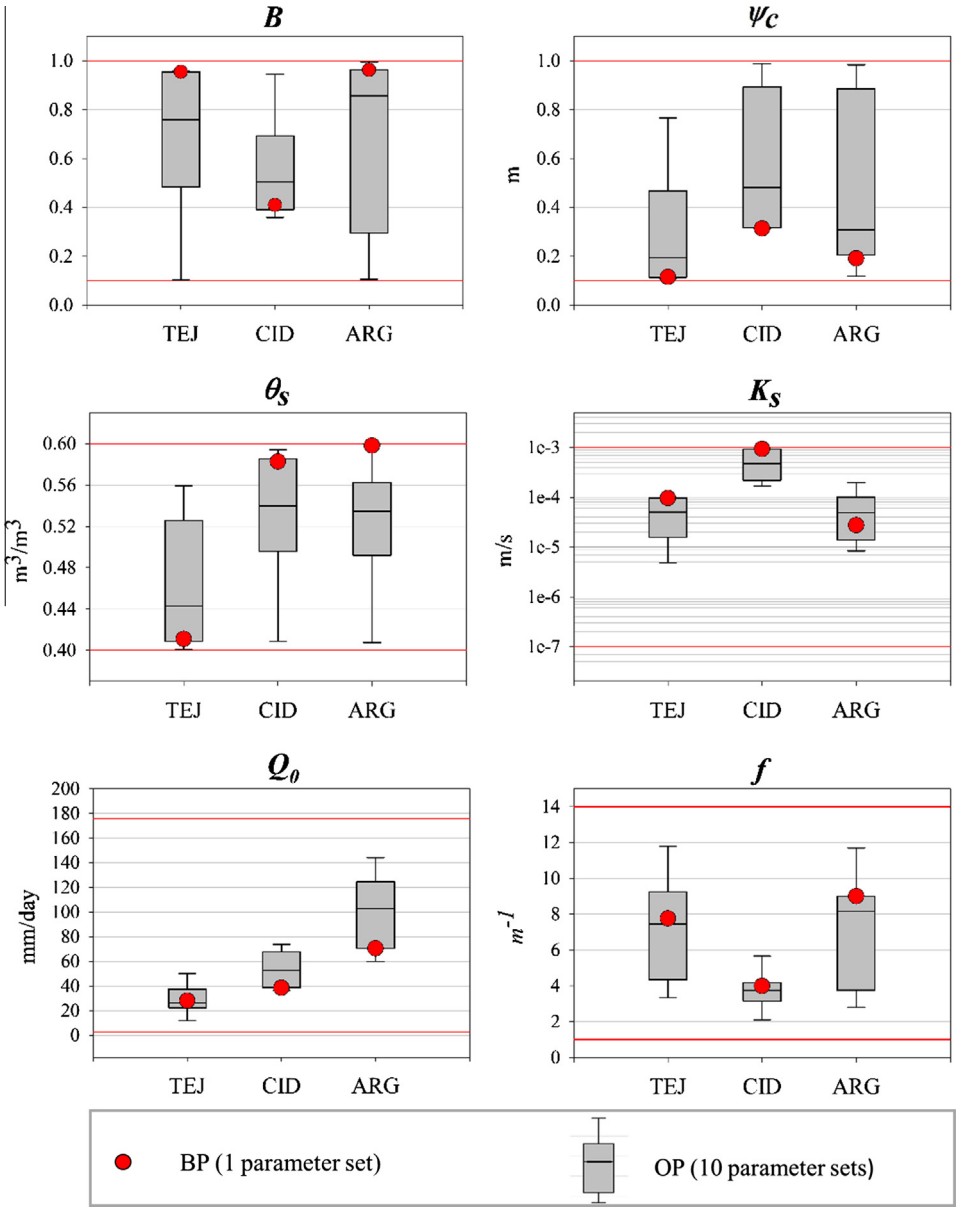

Fig. 7. TOPLATS Optimal (OP) and Best Performing (BP) parameter values obtained through MSPM model calibration of daily streamflow simulation with the conventional (CON) approach. Red lines indicate optimization parameter search range. Results for La Tejeria (TEJ), Cidacos (CID) and Arga (ARG) catchments. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Daily streamflow simulation efficiencies (NSE_1 and NSE) and bias results achieved after optimization.

Catchment	Calibration type	NSE_1		NSE		Pbias	
		CAL –	VAL –	CAL –	VAL –	CAL %	VAL %
1. La Tejeria	CON	0.71	0.65	0.82	0.72	2	2
	RND	0.64	0.69	0.77	0.73	−5	−4
2. Cidacos	CON	0.53	0.35	0.61	0.09	2	−77
	RND	0.54	0.39	0.91	0.25	8	−3
3. Arga	CON	0.50	0.49	0.63	0.60	−3	−9
	RND	0.53	0.46	0.71	0.54	2	−5

One of the points of interests of this study was to evaluate the performance of the model on different climate conditions. Difficulties faced in Cidacos calibration and validation also relate with Perrin et al. (2007) and Li et al. (2010), who concluded that stable parameter values proved more difficult to reach in dry catchments (in Cidacos unstable parameter values were obtained on CON calibration). Also Li et al. (2010) used random selection of CAL/VAL periods of different lengths, and concluded that humid catchments required shorter CAL/VAL periods than dry ones to obtain stable parameter values, coinciding with our results in Mediterranean catchments.

4.4. Analysis of model outputs

4.4.1. Water on the soil profile

Fig. 8 presents daily simulated soil moisture (θ) behavior in relation with recorded rainfall. Following the best results detailed in previous 4.3 section, for La Tejeria and Arga the results shown in Figs. 8 and 9 are those obtained with CON calibration approach, while Cidacos's ones correspond to RND strategy. The plot represents the whole period studied (i.e. including CAL/VAL and

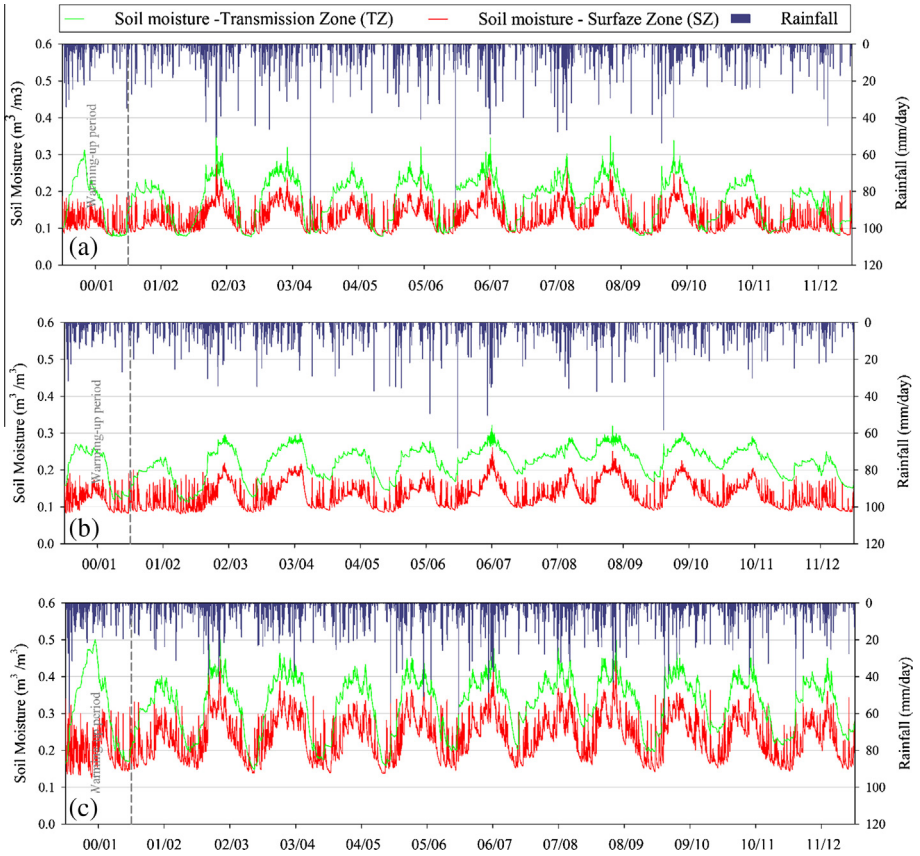

Fig. 8. Simulated soil moisture content (θ) (cm^3/cm^3) in the Surface Zone (SZ) and in the Transmission Zone (TZ) for (a) La Tejeria, (b) Cidacos and (c) Arga catchments.

warming-up periods). It can be observed that within each catchment, the deeper TZ kept higher soil moisture contents than the SZ for most part of the year. This difference was higher during winter periods and lower during summer when θ dropped in both layers.

Simulated mean soil moisture (θ_m) values in the SZ were 0.13 $cm^3\,cm^{-3}$ in La Tejeria, 0.13 $cm^3\,cm^{-3}$ in Cidacos and 0.24 $cm^3\,cm^{-3}$ in Arga. La Tejeria presented the smallest mean value difference between layers (5%), whereas that difference reached 9% in Cidacos and Arga. For the TZ θ_m values were 0.18 $cm^3\,cm^{-3}$ in La Tejeria, 0.22 $cm^3\,cm^{-3}$ in Cidacos and 0.33 $cm^3\,cm^{-3}$ in Arga. During summer periods, larger differences among layers were found in Cidacos. Differences between Arga (with the highest soil moisture mean values in both layers) and the other two catchments were especially remarkable on the SZ. In our study, obtained soil moisture values were, globally, low in comparison with other studies (Goegebeur and Pauwels, 2007; Loosvelt et al., 2011). Soil moisture fluctuation (θ_{SD}) was in all the catchments larger in TZ than in SZ. In both layers, largest fluctuations were found in Arga catchment. θ_{SD} values in La Tejeria were 0.03 $cm^3\,cm^{-3}$ (SZ) and 0.06 $cm^3\,cm^{-3}$ (TZ); in Cidacos 0.03 $cm^3\,cm^{-3}$ and 0.04 $cm^3\,cm^{-3}$; and in Arga 0.06 $cm^3\,cm^{-3}$ and 0.07 $cm^3\,cm^{-3}$.

Regarding the proportion of saturated catchment area (Fig. 9), it can be observed that intense and persistent rainfall events resulted in similar maximum percentages of saturated catchment in La Tejeria (43.9%), Cidacos (37.3%) and Arga (37.4%). However, mean and percentile analysis of saturation degrees showed large variation among catchments (not shown). In Arga, up to 24% of the simulation days presented more than 10% of its area as fully saturated. In La Tejeria, that saturation level (10%) was reached only during 6% of the days and in Cidacos in 3%. Mean value of saturated area (%) in La Tejeria, Cidacos and Arga were 1.5%, 1.7% and 6.7% respectively. Largest saturated areas were found in Arga, about four times the values obtained for the other 2 catchments. Highest saturation variability rates were also found in Arga simulation, with a standard deviation of 6%, while variation was 4.3% in La Tejeria and 3.3% in Cidacos.

Fig. 9 also depicts the water table depth, whose principal statistical measures (mean, WTD_m, and standard deviation, WTD_{SD}) are detailed as follows: WTD_m in La Tejeria was 0.71 m, in Cidacos 1.14 m and in Arga 0.63 m. Regarding water table variation, WTD_{SD} in La Tejeria was 0.27 m, in Cidacos 0.31 m and in Arga 0.19 m. As in Figs. 8 and 9 La Tejeria and Arga results are those obtained with CON calibration approach, while Cidacos's ones correspond to RND strategy. As it was commented for Table 4, and can be observed in Fig. 9, RND approach was able to maintain a stable behavior of WTD in Cidacos. Water table depth, especially its minimum values (i.e. closer to surface, mainly during persistent winter events) is a key factor in TOPLATS runoff generation. When WTD reaches the value of ψ_c (m), simulated runoff increases non-linearly, affecting remarkably simulation efficiency results and model performance. Cidacos, the driest catchment had the deepest WTD_m (1.14 m) and also the highest WTD_{SD}. Arga had the most stable WTD with a standard deviation of just 0.19 m.

4.4.2. Streamflow simulation

- Daily streamflow

Results presented hereafter (Figs. 10–12) for La Tejeria and Arga also correspond to the conventional CAL/VAL strategy, while Cidacos results are the ones obtained with RND. Streamflow results are presented in three types of plots: time series (Fig. 10), simulated vs observed scatter plots (Fig. 11) and streamflow duration curves (Fig. 12) to allow for detailed TOPLATS performance analysis. On the time-series plot (Fig. 10) it can be observed that almost no

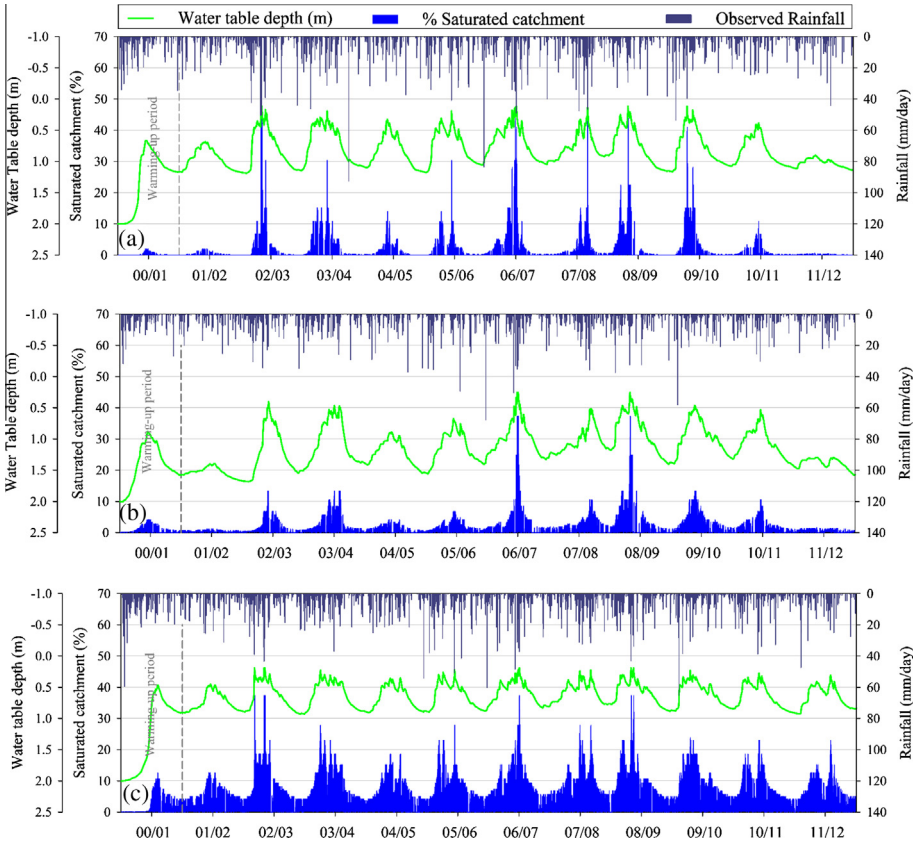

Fig. 9. TOPLATS daily simulated plot: (1) water table depth (m), and (2) Percentage of catchment area with fully saturated soil profile (%) for (a) La Tejeria, (b) Cidacos and (c) Arga catchments.

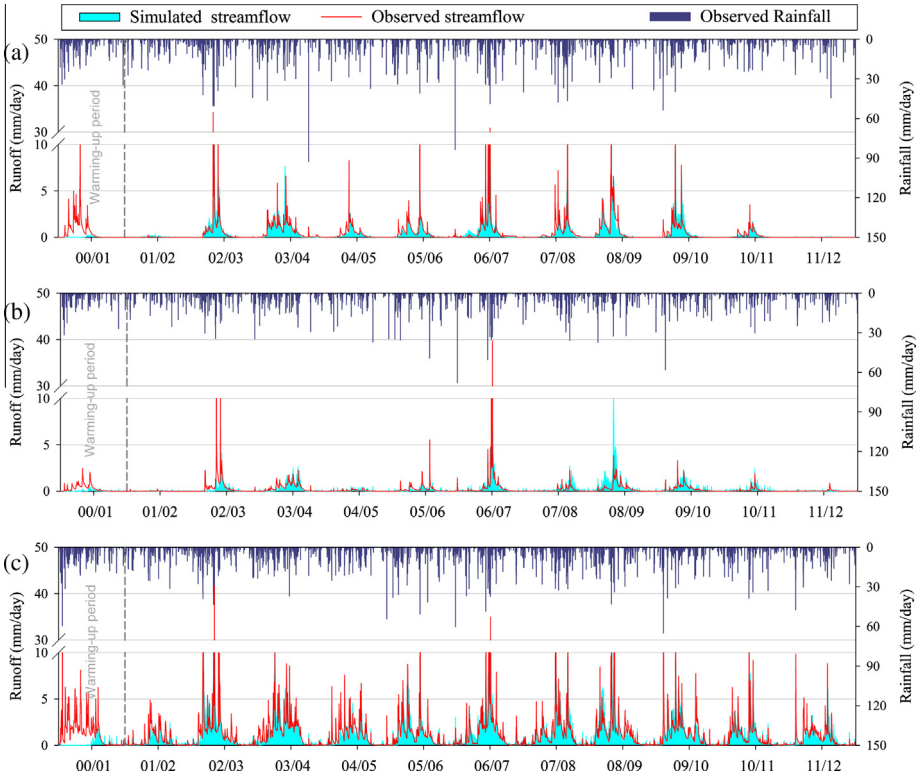

Fig. 10. Daily rainfall, observed streamflow and simulated streamflow in: (a) La Tejeria, (b) Cidacos and (c) Arga.

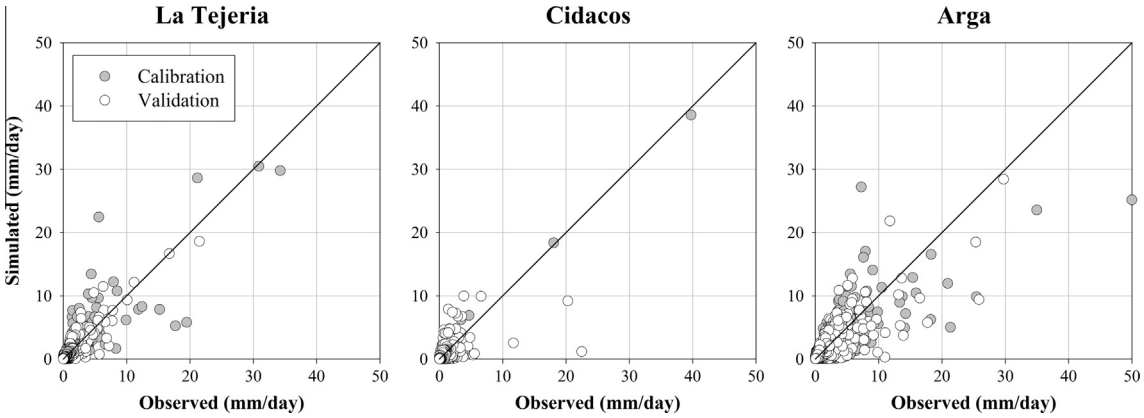

Fig. 11. Scatter plot of daily observed and simulated streamflow (mm/day) in: (a) La Tejeria, (b) Cidacos and (c) Arga.

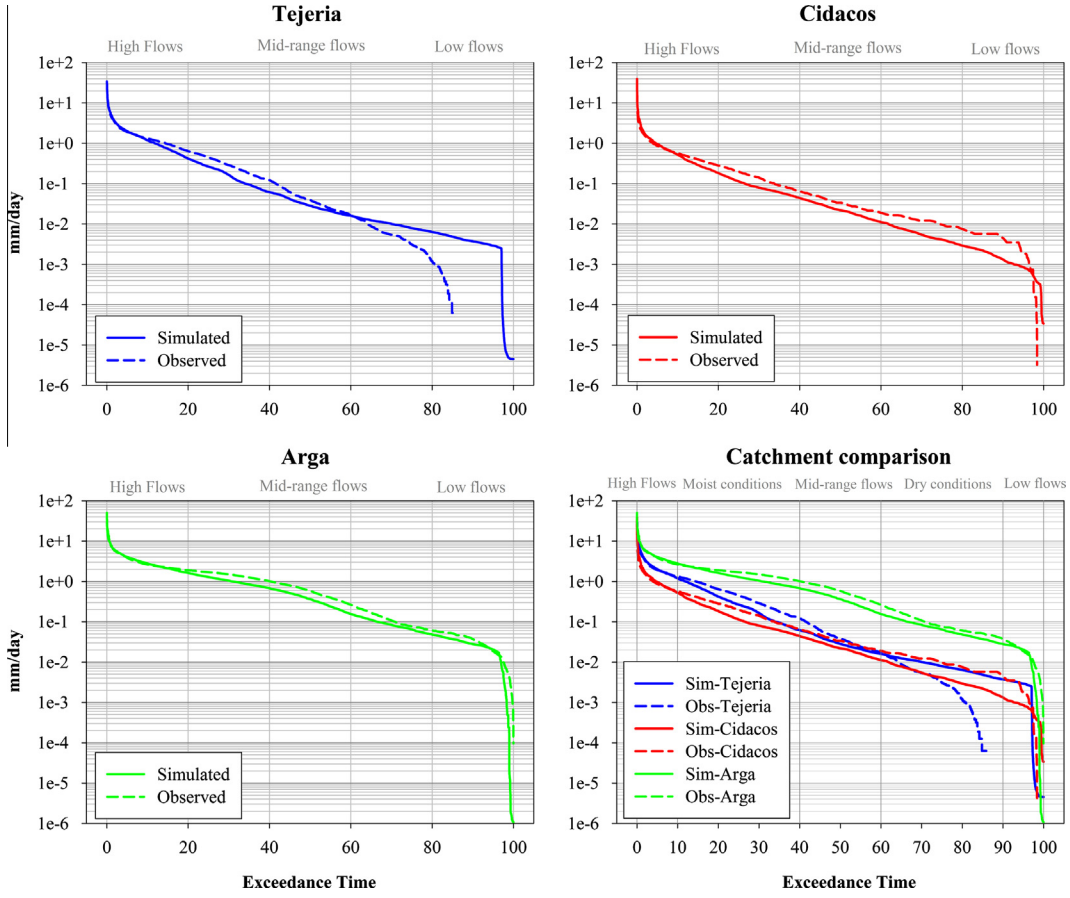

Fig. 12. Daily streamflow duration curves (observed vs simulated) in (a) La Tejeria, (b) Cidacos, (c) Arga and (d) catchments comparison.

runoff was simulated during the warming-up one-year period in all three catchments. This was due to the initial value defined for the water table depth (2 m). Shallower initial WTD values resulted in a more rapid model response but worse overall results for the remaining period. Large inter-annual variability (Fig. 10) was observed in terms of streamflow in la Tejeria and Cidacos. Despite this large variability, TOPLATS was able to respond adequately, particularly in La Tejeria. Streamflow behavior was much more stable in Arga, allowing better model response, but TOPLATS faced difficulties on Cidacos extreme storm events simulation.

Scatter plots (Fig. 11) illustrate very clearly how few streamflow peaks affected NSE results. In Cidacos, two high peak flows were very accurately simulated during the calibration period, and account large responsibility for the NSE = 0.91 efficiency obtained (Table 4). But on the other hand, main peak events during validation were poorly simulated (Cidacos validation: NSE = 0.25). NSE validation efficiency results were extremely affected by just 7 days data with large simulation miscalculations (out of a 4383 days simulation), while validation Pbias remained very low (3%). In La Tejeria and Arga there were no such differences in CAL and VAL events and NSE results were more balanced (Table 4).

Daily simulated streamflow was also evaluated according to detailed specific flow ranges (high flows, mid-range flows and low flows) in flow duration curves comparing simulated and observed flows (Fig. 12). Main divergences were found in low rate flows in La Tejeria, where TOPLATS clearly over-estimated

observed streamflows. This was apparent, since up to 14% of the studied days (mostly summer periods) no streamflow was recorded at all, but TOPLATS kept simulating small streamflow amounts. In any case, this overestimation of low flows, has a very limited impact on total volume and efficiency results. In La Tejeria, also a slight underestimation of modeled flows was found during catchment's moist conditions (10–50%). In Cidacos, largest discrepancies (underestimation) between simulated and observed values were found on the 40% of lower-flow days. The same pattern applied for most of the days, where TOPLATS underestimated in most of the flow ranges, yet the difference decreased towards high flows. During the driest 1.5% of the days, no streamflow was observed in Cidacos, so the model showed overestimation on that range (right tale of the curve, 97–100%). Arga presented the best agreement between simulated and observed flows in all streamflow level ranges but TOPLATS tended to underestimate mid-range flows. When all catchment results are plotted together, it is observed that Arga (wettest catchment) patterns clearly differentiate from the other two due to its highest rainfall conditions. On the highest streamflow data range, it could be observed that

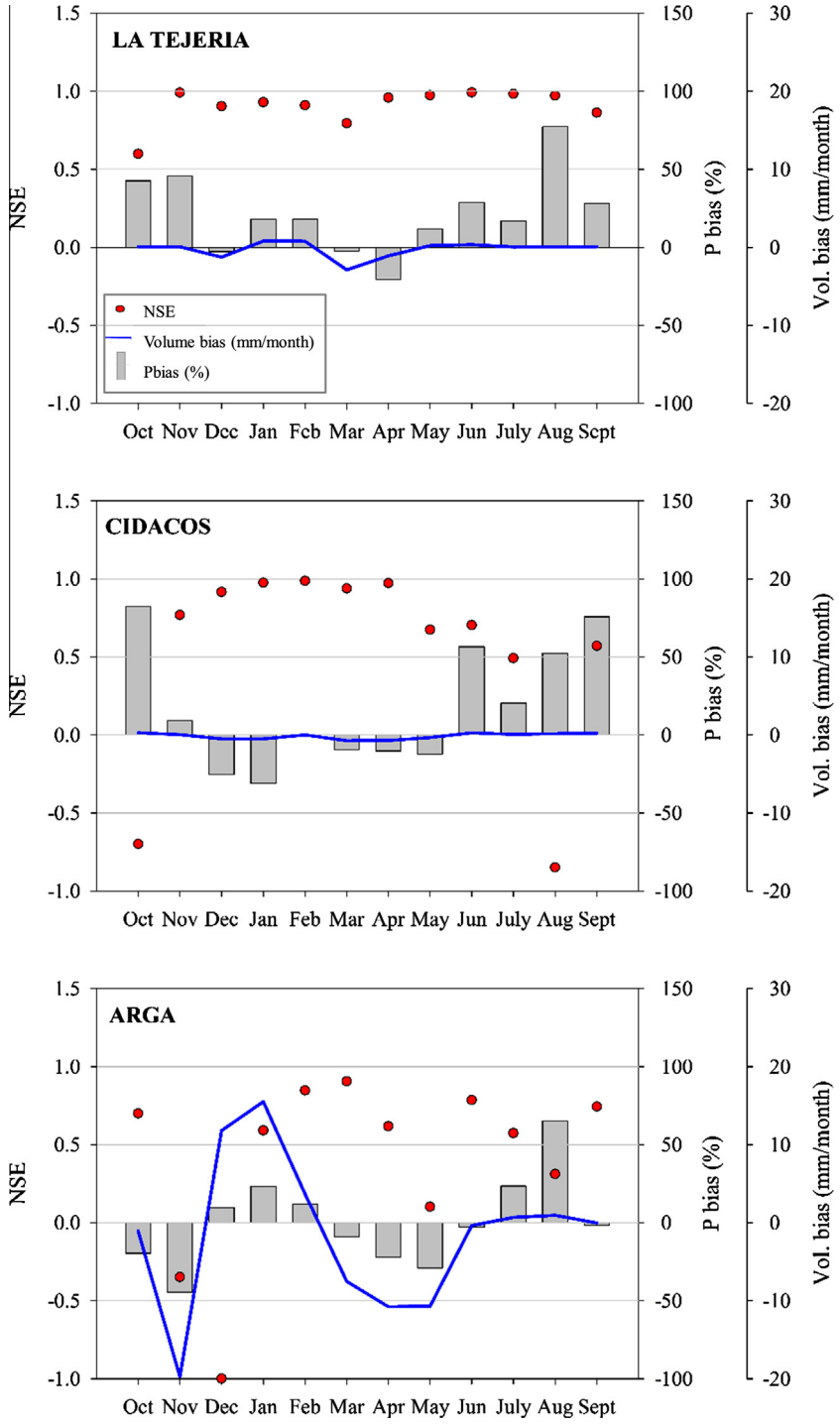

Fig. 13. Comparison of monthly observed and simulated streamflow results: (1) NSE (dots), (2) Pbias (%) (bar plot), and (3) Volume bias (mm) (line plot) results in La Tejeria, Cidacos and Arga.

Arga had the most intense (mm/day) runoff events and Cidacos the least. High flow rate range, as presented in the catchments' comparative plot of Fig. 12, include the 10% of the days with highest streamflow values (438 days).

- Monthly streamflow

Daily observed and simulated streamflow values were aggregated on monthly basis to obtain monthly results. These monthly results are useful to understand and evaluate model performance at a seasonal time-scale, and for different modeling purposes (e.g. for hydrological resource management). Three types of monthly streamflow simulation results are shown in Fig. 13: (1) monthly NSE, (2) monthly Pbias (%) and (3) Volume bias (mm/month).

Results obtained in La Tejeria showed a high 0.94 overall NSE value. In nine out of twelve months, results were over 0.9 in terms of NSE. The lowest efficiency was found in October (0.6), at the beginning of the hydrological year. The largest Pbias was found during the simulation of summer and autumn months, when intense rainfall events occur while TOPLATS simulated water table is at its deeper values, and saturated catchment area values are low. Since Pbias is expressed as percentage, its largest values were obtained in August, when streamflow was largely overestimated (77%). However, this high Pbias value corresponded with a very minor errors in volume units (mm/month), as commented for daily results (Fig. 12). Largest Volume bias in La Tejeria were obtained in the March–April (spring) period, where TOPLATS underestimated monthly streamflow. That streamflow underestimation in March months implied a lower (0.80) NSE.

NSE monthly median was 0.74 in Cidacos. In this catchment, two different periods could be distinguished in terms of monthly efficiency results. First, from December to April, TOPLATS offered a very accurate performance, reaching NSE values over 0.9 for each month of that winter-spring five months period. On the contrary, the model was unable to properly simulate low flows typical in August, September and October. In Cidacos, the model clearly tended to overestimate streamflow for dry periods. Mean monthly overestimation from June to October was 57% (probably also related to river extractions for irrigation purposes). In terms of streamflow volumetric error, largest discrepancies, but still rather low, were found in March and April, when the model underestimated flow volumes. That systematic underestimation was similarly observed for the whole winter-spring period.

Out of the three evaluated catchments, TOPLATS performed worst in Arga, in terms of monthly efficiency results. Monthly median NSE in Arga was 0.60. This catchment showed the largest NSE variability between months, offering its best NSE results in February (0.85) and March (0.9). Results indicated a clear pattern of modeled overestimation during winter months (December–January–February) and underestimation in autumn (October–November) and spring (March–April–May).

When daily and monthly NSE were compared, it was observed that best results in both cases were obtained in La Tejeria. In Cidacos, monthly results outperformed clearly daily NSE, while the opposite occurred in Arga, where very low efficiencies, mainly in November and December caused low monthly NSE. Bormann (2006b) also performed TOPLATS applications in several catchments in Germany, ranging from 63 to 134 km^2. Daily NSE results presented, varied from 0.59 to 0.73 on the calibration period and from 0.52 to 0.69 on validation years. Efficiency results were found to increase on that study, reaching 0.76–0.85 on weekly analysis and 0.82–0.90 on monthly evaluation.

4.5. Hourly vs daily streamflow simulation

Hourly simulation of the catchments was also similarly performed for the same 12 years (2000–2012) period, with a total of 96,432 h simulation. At this time-scale, only the CON calibration period method was applied, as no abnormal behavior of the model was found in any of the catchments. Similarly to daily analysis, the first year was considered as warming-up. Hourly streamflow simulated values (mm/h) versus observed data (mm/h) are first presented (Fig. 14) in a scatter plot. While in La Tejeria no clear pattern of over or under-estimation could be identified, in Cidacos and Arga it was observed that extreme events were underestimated by TOPLATS. As observed in Fig. 14, for most of the time-series, Cidacos had the lowest streamflow values (most points below 0.5 mm/h), but also the highest intensity events were observed in that catchment.

Best NSE_1 results obtained for hourly calibration period (Table 5) were similar to their corresponding daily results (Table 4) in Cidacos (0.48 vs 0.53) and Arga (0.51 vs 0.50). On the contrary, in La Tejeria, NSE_1 was significantly lower for hourly calibration (0.56

Table 5
Hourly streamflow simulation efficiency (NSE_1 and NSE) and bias results achieved after optimization with the conventional (CON) calibration approach.

Catchment	NSE_1		NSE		Pbias	
	CAL –	VAL –	CAL –	VAL –	CAL %	VAL %
1. La Tejeria	0.56	0.58	0.55	0.66	8	5
2. Cidacos	0.48	0.51	0.41	0.77	−6	8
3. Arga	0.51	0.52	0.64	0.66	1	4

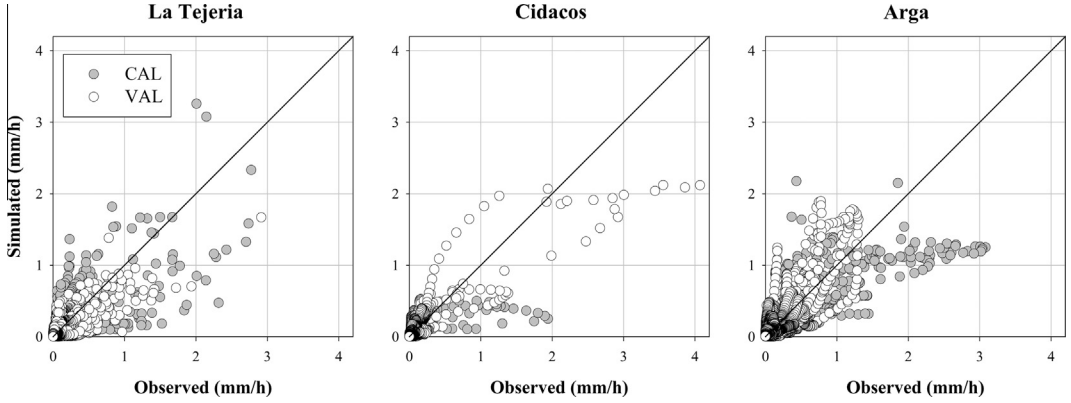

Fig. 14. Scatter plot of hourly observed and simulated streamflow (mm/h) in: (a) La Tejeria, (b) Cidacos and (c) Arga.

vs 0.71). For the validation period, NSE_1 hourly results were notably better in Cidacos (0.51 vs 0.35) and slightly better Arga (0.52 vs 0.49), but hourly results performed worse in La Tejeria (0.58 vs 0.65).

Analysis of NSE performance concluded that in La Tejeria daily simulations outperformed substantially hourly simulations. This results can be explained by the small size of La Tejeria, with very short runoff evacuation times (hours), which may not be properly captured by the model when run at hourly time step, even though the model was added a specific routing capability. In Cidacos, the NSE value of 0.77 obtained in validation when run hourly contrasted with the poor value obtained with daily time-step (0.09). NSE results in Arga were quite stable from calibration to validation periods at both running time-scales: the difference in daily simulation was −0.03 (calibration NSE was 0.63 and validation 0.6) and +0.02 in hourly results (Table 5). Total simulation NSE (CAL + VAL) in Arga was 0.62 for daily data and 0.65 for hourly. Hourly simulation with TOPLATS was also performed by Loosvelt et al. (2015) on a 91 km^2 catchment in Belgium, reporting NSE = 0.33 for CAL and 0.44 for VAL.

Application of MSPM in combination with NSE_1 as objective function proved to achieve excellent Pbias reduction (all volume errors lower than 8%, Table 5) allowing TOPLATS to be used as a consistent water resources management tool for continuous simulation both at the daily and hourly scale.

5. Conclusions

The evaluation of TOPLATS performed in this research provides model users with useful guidelines to: (1) identify the model parameters having the largest influence on main hydrological processes and streamflow simulation efficiency; (2) perform an efficient calibration of the model through an automatic calibration algorithm applied with a multi-start approach; (3) gain insight on the performance, advantages and limitations of the model when applied at daily and hourly time-scales and (4) be aware of the importance of climatological variability to be contained on the calibration period in Mediterranean catchments.

Sensitivity analyses performed with Morris and Sobol methods yielded very similar results and concluded that 3 parameters: Brooks–Corey Pore Size distribution Index (B), Bubbling pressure (ψ_c) and Hydraulic conductivity decay (f) had the overall largest influence on the hydrological processes (surface runoff, baseflow, evapotranspiration and surface zone soil moisture dynamics) and streamflow simulation efficiency. Thus, their inclusion in any TOPLATS calibration is recommended. Morris method gave similar results to Sobol, with lower computational requirements, which makes it a reliable and suitable method to be applied on complex and largely parameterized physically based models similar to TOPLATS.

Regardless of catchment size or climate influence, mean surface soil moisture was found to be controlled by B, and its dynamics by f. Streamflow generation on wet catchments (Arga and La Tejeria) seemed to be mainly influenced by ψ_c, whereas the drier catchment (Cidacos) was largely affected by B.

Model calibration achieved a substantial efficiency improvement in the three evaluated catchments, and reduced Pbias to values below 10% (in most cases below 5%). Calibration improved both average flows and high peaks simulation results, in both calibration and validation periods.

Climatic variations between calibration and validation periods compromise model performance and parameter stability, particularly in arid zones. A random and discontinuous period selection strategy was applied in this study to overcome this issue, which outperformed model calibration with the conventional calibration and validation period selection strategy. This random approach is thus recommended when large climate variability is found between calibration and validation periods, and particularly when run on daily basis.

In terms of streamflow simulation efficiency, when TOPLATS was run in hourly and daily basis, catchment size had an influence on the results. Mean Nash & Sutcliffe efficiency of calibration and validation periods on the experimental micro-catchment studied (La Tejeria) was higher in hourly basis (NSE 0.77) than in daily basis (NSE 0.60). The opposite occurred on the largest catchment (Arga) where hourly efficiency (0.65) outperformed daily results (NSE 0.62). On the intermediate size catchment (Cidacos), efficiency results were very similar (NSE 0.58 and 0.59). Surface soil moisture behavior (characterized by rapid variations) seemed not to be properly simulated at large time scales, affecting model behavior notoriously.

Monthly efficiency results showed that TOPLATS performed optimally in the two smallest catchments, but some systematic deviation was found on the largest catchment, uncovering a clear overestimation pattern during winter and underestimation in autumn and spring. This seasonal deviations could be compensated by implementing data assimilation techniques and using observed streamflow, remote sensing or in-situ soil moisture data, to update model variables (i.e. soil moisture, water balances) in order to correct that systematic deviated pattern.

Acknowledgments

This study was partially funded by the Spanish Ministry of Science and Innovation (project CGL2011-24336) and by the Public University of Navarre through a pre-doctorate research scholarships to the first author.

References

Alvarez-Mozos, J., Casali, J., Gonzalez-Audicana, M., Verhoest, N.E.C., 2006. Assessment of the operational applicability of RADARSAT-1 data for surface soil moisture estimation. IEEE Trans. Geosci. Remote Sens. 44, 913–924.

Anctil, F., Perrin, C., Andréassian, V., 2004. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. Environ. Model. Softw. 19, 357–368.

Baroni, G., Tarantola, S., 2014. A general probabilistic framework for uncertainty and global sensitivity analysis of deterministic models: a hydrological case study. Environ. Model. Softw. 51, 26–34.

Beven, K.J., Kirkby, M.J., 1979. Physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. Sci. Hydrol. 24, 43–69.

Bormann, H., 2006a. Effects of grid size and aggregation on regional scale landuse scenario calculations using SVAT schemes. Adv. Geosci. 9, 45–52.

Bormann, H., 2006b. Impact of spatial data resolution on simulated catchment water balances and model performance of the multi-scale TOPLATS model. Hydrol. Earth Syst. Sci. 10, 165–179.

Bormann, H., Breuer, L., Gräff, T., Huisman, J.A., 2007. Analysing the effects of soil properties changes associated with land use changes on the simulated water balance: a comparison of three hydrological catchment models for scenario analysis. Ecol. Model. 209, 29–40.

Brath, A., Montanari, A., Toth, E., 2004. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. J. Hydrol. 291, 232–253.

Brocca, L., Melone, F., Moramarco, T., 2011. Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. Hydrol. Process. 25, 2801–2813.

Brooks, R.H., Corey, A.T., 1964. Hydraulic properties of porous media and their relation to drainage design. Trans. ASABE 7, 26–0028.

Burnash, R.J., Ferral, R.L., Mc Guire, R.A., 1973. A generalized streamflow simulation system conceptual modeling for digital computers.

Campolongo, F., Saltelli, A., 1997. Sensitivity analysis of an environmental model: an application of different analysis methods. Reliab. Eng. Syst. Safe. 57, 49–69.

Campolongo, F., Saltelli, A., Cariboni, J., 2011. From screening to quantitative sensitivity analysis. A unified approach. Comput. Phys. Commun. 182, 978–988.

Casalí, J., Gastesi, R., Álvarez-Mozos, J., De Santisteban, L.M., de Lersundi, J.D.V., Giménez, R., Larrañaga, A., Goñi, M., Agirre, U., Campo, M.A., López, J.J., Donézar, M., 2008. Runoff, erosion, and water quality of agricultural watersheds in central Navarre (Spain). Agric. Water Manage. 95, 1111–1128.

Chen, R.S., Pi, L.C., Hsieh, C.C., 2005. Application of parameter optimization method for calibrating tank model. J. Am. Water Resour. Assoc. 41, 389–402.

Chiew, F.H.S., McMahon, T.A., 2002. Modelling the impacts of climate change on Australian streamflow. Hydrol. Process. 16, 1235–1245.
Crow, W.T., Drusch, M., Wood, E.F., 2001. An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval. IEEE Trans. Geosci. Remote Sens. 39, 1622–1631.
Crow, W.T., Ryu, D., Famiglietti, J.S., 2005. Upscaling of field-scale soil moisture measurements using distributed land surface modeling. Adv. Water Resour. 28, 1–14.
Crow, W.T., Wood, E.F., 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. Adv. Water Resour. 26, 137–149.
Crow, W.T., Wood, E.F., 2002. The value of coarse-scale soil moisture observations for regional surface energy balance modeling. J. Hydrometeorol. 3, 467–482.
Demaria, E.M., Nijssen, B., Wagener, T., 2007. Monte Carlo sensitivity analysis of land surface parameters using the variable infiltration capacity model. J. Geophys. Res. 112, D11113.
Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28, 1015–1031.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76, 501–521.
EC-JRC, 2008. SimLab 2.2 Reference Manual Rep.
Endreny, T.A., Wood, E.F., Lettenmaier, D.P., 2000. Satellite-derived digital elevation model accuracy: hydrological modelling requirements. Hydrol. Process. 14, 177–194.
Famiglietti, J.S., Wood, E.F., 1994. Multiscale modeling of spatially variable water and energy balance processes. Water Resour. Res. 30, 3061–3078.
Francos, A., Elorza, F.J., Bouraoui, F., Bidoglio, G., Galbiati, L., 2003. Sensitivity analysis of distributed environmental simulation models: understanding the model behaviour in hydrological studies at the catchment scale. Reliab. Eng. Syst. Safe. 79, 205–218.
Gan, T.Y., Biftu, G.F., 1996. Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. Water Resour. Res. 32, 3513–3524.
Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C., Di, Z., 2014. A comprehensive evaluation of various sensitivity analysis methods: a case study with a hydrological model. Environ. Model. Softw. 51, 269–285.
Garambois, P.A., Roux, H., Larnier, K., Castaings, W., Dartus, D., 2013. Characterization of process-oriented hydrologic model behavior with temporal sensitivity analysis for flash floods in Mediterranean catchments. Hydrol. Earth Syst. Sci. 17, 2305–2322.
Geem, Z.W., Roper, W.E., 2010. Various continuous harmony search algorithms for web-based hydrologic parameter optimisation. Int. J. Math. Model. Numer. Optim. 1, 213–226.
Goegebeur, M., Pauwels, V.R.N., 2007. Improvement of the PEST parameter estimation algorithm through Extended Kalman filtering. J. Hydrol. 337, 436–451.
Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. Water Resour. Res. 34, 751–763.
Houser, P.R., Shuttleworth, W.J., Famiglietti, J.S., Gupta, H.V., Syed, K.H., Goodrich, D. C., 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resour. Res. 34, 3405–3420.
Jacquemin, B., Noilhan, J., 1990. Sensitivity study and validation of a land surface parameterization using the HAPEX-MOBILHY data set. Bound.-Lay. Meteorol. 52, 93–134.
Jarvis, P.G., 1976. The interpretation of the variation in leaf water potential and stomatal conductance found in canopies in the field. Philos. Trans. Roy. Soc. Lond. B Biol. Sci. 273, 593–610.
Khakbaz, B., Imam, B., Hsu, K., Sorooshian, S., 2012. From lumped to distributed via semi-distributed: calibration strategies for semi-distributed hydrologic models. J. Hydrol. 418–419, 61–77.
Kim, U., Kaluarachchi, J.J., 2009. Hydrologic model calibration using discontinuous data: an example from the upper Blue Nile River Basin of Ethiopia. Hydrol. Process. 23, 3705–3717.
Kobayashi, S., Maruyama, T., 1976. Search for the coefficients of the reservoir model with the Powell's conjugate direction method. Trans. Jpn. Soc. Irrig. Drain. Reclam. Eng. 65, 42–47.
Li, C.Z., Wang, H., Liu, J., Yan, D.H., Yu, F.L., Zhang, L., 2010. Effect of calibration data series length on performance and optimal parameters of hydrological model. Water Sci. Eng. 3, 378–393.
Loaiza-Usuga, J.C., Pauwels, V.R.N., 2008. Calibration and multiple data set-based validation of a land surface model in a mountainous Mediterranean study area. J. Hydrol. 356, 223–233.
Loaiza-Usuga, J.C., Poch, R.M., 2009. Evaluation of soil water balance components under different land uses in a mediterranean mountain catchment (Catalan pre-pyrenees NE Spain). Zeitschrift fur Geomorphol. 53, 519–537.
Loosvelt, L., De Baets, B., Pauwels, V.R.N., Verhoest, N.E.C., 2014a. Assessing hydrologic prediction uncertainty resulting from soft land cover classification. J. Hydrol. 517, 411–424.
Loosvelt, L., Pauwels, V.R.N., Cornelis, W.M., De Lannoy, G.J.M., Verhoest, N.E.C., 2011. Impact of soil hydraulic parameter uncertainty on soil moisture modeling. Water Resour. Res. 47, n/a–n/a.
Loosvelt, L., Pauwels, V.R.N., Verhoest, N.E.C., 2015. On the significance of crop-type information for the simulation of catchment hydrology. Hydrol. Process. 29, 915–926.
Loosvelt, L., Vernieuwe, H., Pauwels, V.R.N., De Baets, B., Verhoest, N.E.C., 2014b. Local sensitivity analysis for compositional data with application to soil texture in hydrologic modelling. Hydrol. Earth Syst. Sci. 17, 461–478.
Lucau-Danila, C., Callens, M., Defourny, P., Verhoest, N.E.C., Pauwels, V.R.N., 2005. Vegetation parameter retrieval from SAR data using near-surface soil moisture estimates derived from a hydrological model. In: Owe, M., D'Urso, G. (Eds.), Proceedings of SPIE – The International Society for Optical Engineering, pp. 597603–597603–12.
Massmann, C., Holzmann, H., 2012. Analysis of the behavior of a rainfall-runoff model using three global sensitivity analysis methods evaluated at different temporal scales. J. Hydrol. 475, 97–110.
Milly, P.C.D., 1986. Event-based simulation model of moisture and energy fluxes at bare soil surface. Water Resour. Res. 22, 1680–1692.
Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics, 161–174.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. J. Hydrol. 10, 282–290.
Norton, J.P., 2009. Selection of Morris trajectories for initial sensitivity analysis. In: IFAC Proceedings Volumes (IFAC-PapersOnline), pp. 670–674.
NRCS-USDA, 2014. 12th Edition Keys to Soil Taxonomy.
Oakley, J.E., O'Hagan, A., 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach. J. Roy. Stat. Soc. Ser. B (Stat. Methodol.) 66, 751–769.
Paik, K., Kim, J.H., Kim, H.S., Lee, D.R., 2005. A conceptual rainfall-runoff model considering seasonal variation. Hydrol. Process. 19, 3837–3850.
Passerat De Silans, A., Vauclin, M., Bruckler, L., Bertuzzi, P., Brunet, Y., 1986. Numerical modeling of water and heat flows in unsaturated soils under atmospheric excitation. Comparison with field data. In: Heat Transfer, Proceedings of the International Heat Transfer Conference, pp. 2629–2634.
Pauwels, V.R.N., Balenzano, A., Satalino, G., Skriver, H., Verhoest, N.E.C., Mattia, F., 2009. Optimization of soil hydraulic model parameters using synthetic aperture radar data: an integrated multidisciplinary approach. IEEE Trans. Geosci. Remote Sens. 47, 455–467.
Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., 2001. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. J. Hydrol. 251, 88–102.
Pauwels, V.R.N., Hoeben, R., Verhoest, N.E.C., De Troch, F.P., Troch, P.A., 2002. Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. Hydrol. Process. 16, 995–1013.
Pauwels, V.R.N., Verhoest, N.E.C., De Lannoy, G.J.M., Guissard, V., Lucau, C., Defourny, P., 2007. Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. Water Resour. Res. 43, n/a–n/a.
Pauwels, V.R.N., Wood, E.F., 1999. A soil-vegetation-atmosphere transfer scheme for the modeling of water and energy balance processes in high latitudes 2. Application and validation. J. Geophys. Res. D: Atmos. 104, 27823–27839.
Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C., Mathevet, T., 2007. Impact of limited streamflow data on the efficiency and the parameters of rainfall-runoff models. Hydrol. Sci. J. 52, 131–151.
Peters-Lidard, C.D., Zion, M.S., Wood, E.F., 1997. A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes. J. Geophys. Res. D: Atmos. 102, 4303–4324.
Powell, M.J.D., 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J. 7, 155–162.
Prata, A.J., 1996. A new long-wave formula for estimating downward clear-sky radiation at the surface. Q. J. R. Meteorol. Soc. 122, 1127–1151.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1988. Numerical Recipes in Fortran 77: The Art of Scientific Computing. Cambridge University Press.
Rao, S.S., 2009. Engineering Optimization: Theory and Practice, fourth ed. John Wiley and Sons Ltd, New Jersey.
Ratto, M., Tarantola, S., Saltelli, A., 2001. Sensitivity analysis in model calibration: GSA-GLUE approach. Comput. Phys. Commun. 136, 212–224.
Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of soil water properties. Trans. Am. Soc. Agric. Eng. 25, 1316–1320, 1328.
Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun. 145, 280–297.
Saltelli, A., Tarantola, S., Campolongo, F., 2000. Sensitivity analysis as an ingredient of modeling. Stat. Sci. 15, 377–395.
SCS – Soil Conservation Service, 1972. National Engineering Handbook, section 4. Washington D.C.
Senarath, S.U.S., Ogden, F.L., Downer, C.W., Sharif, H.O., 2000. On the calibration and verification of two-dimensional, distributed, Hortonian, continuous watershed models. Water Resour. Res. 36, 1495–1510.
Seuffert, G., Gross, P., Simmer, C., Wood, E.F., 2002. The influence of hydrologic modeling on the predicted local weather: two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. J. Hydrometeorol. 3, 505–523.
Shin, M.J., Guillaume, J.H.A., Croke, B.F.W., Jakeman, A.J., 2013. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. J. Hydrol. 503, 135–152.
Sivapalan, M., Beven, K., Wood, E.F., 1987. On hydrologic similarity. 2. A scaled model of storm runoff production. Water Resour. Res. 23, 2266–2278.

Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–417.
Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Xu, C., 2015. Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. J. Hydrol. 523, 739–757.
Sorooshian, S., Gupta, V.K., 1983. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. Water Resour. Res. 19, 260–268.
Spear, R.C., Grieb, T.M., Shang, N., 1994. Parameter uncertainty and interaction in complex environmental models. Water Resour. Res. 30, 3159–3170.
Sun, X.Y., Newham, L.T.H., Croke, B.F.W., Norton, J.P., 2012. Three complementary methods for sensitivity analysis of a water quality model. Environ. Model. Softw. 37, 19–29.
Tang, Y., Reed, P., Wagener, T., Van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrol. Earth Syst. Sci. 11, 793–817.
Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour. Res. 43, n/a–n/a.
Van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. J. Hydrol. 324, 10–23.
Van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. Adv. Water Resour. 32, 1154–1169.
van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2008. Characterization of watershed model behavior across a hydroclimatic gradient. Water Resour. Res. 44, n/a–n/a.
Viney, N.R., Croke, B.F.W., Breuer, L., Bormann, H., Bronstert, A., Frede, H., Gräff, T., Hubrechts, L., Huisman, J.A., Jakeman, A.J., Kite, G.W., Lanini, J., Leavesley, G., Lettenmaier, D.P., Lindström, G., Seibert, J., Sivapalan, M., Willems, P., 2005. Ensemble modelling of the hydrological impacts of land use change. pp. 2967–2973.
Wagener, T., Kollat, J., 2007. Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. Environ. Model. Softw. 22, 1021–1033.
Wainwright, H.M., Finsterle, S., Jung, Y., Zhou, Q., Birkholzer, J.T., 2014. Making sense of global sensitivity analyses. Comput. Geosci. 65, 94.
Wood, E.F., Sivapalan, M., Beven, K., Band, L., 1988. Effects of spatial variability and scale with implications to hydrologic modeling. J. Hydrol. 102, 29–47.
Yang, G.J., Zhao, C.J., Huang, W.J., Wang, J.H., 2011. Extension of the Hapke bidirectional reflectance model to retrieve soil water content. Hydrol. Earth Syst. Sci. 15, 2317–2326.
Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. J. Hydrol. 181, 23–48.
Young, P.C., 1978. Modeling, identification and control in environmental systems. In: Vansteenkiste, G.C.N.H. (Ed.), Amsterdam, pp. 103–135.
Zhang, C., Chu, J., Fu, G., 2013. Sobol''s sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. J. Hydrol. 480, 58–68.
Zhang, X., Lindström, G., 1997. Development of an automatic calibration scheme for the HBV hydrological model. Hydrol. Process. 11, 1671–1682.
Zhang, X., Srinivasan, R., Zhao, K., Van Liew, M., 2009. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. Hydrol. Process. 23, 430–441.