



Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data

F. Silvestro¹, S. Gabellani¹, R. Rudari¹, F. Delogu¹, P. Laiolo¹, and G. Boni^{1,2}

¹CIMA Research Foundation, Savona, Italy

²DIBRIS, University of Genova, Genova, Italy

Correspondence to: F. Silvestro (francesco.silvestro@cimafoundation.org)

Received: 14 May 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 13 June 2014

Revised: 12 March 2015 – Accepted: 13 March 2015 – Published: 16 April 2015

Abstract. During the last decade the opportunity and usefulness of using remote-sensing data in hydrology, hydrometeorology and geomorphology has become even more evident and clear. Satellite-based products often allow for the advantage of observing hydrologic variables in a distributed way, offering a different view with respect to traditional observations that can help with understanding and modeling the hydrological cycle. Moreover, remote-sensing data are fundamental in scarce data environments. The use of satellite-derived digital elevation models (DEMs), which are now globally available at 30 m resolution (e.g., from Shuttle Radar Topographic Mission, SRTM), have become standard practice in hydrologic model implementation, but other types of satellite-derived data are still underutilized. As a consequence there is the need for developing and testing techniques that allow the opportunities given by remote-sensing data to be exploited, parameterizing hydrological models and improving their calibration.

In this work, Meteosat Second Generation land-surface temperature (LST) estimates and surface soil moisture (SSM), available from European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) H-SAF, are used together with streamflow observations (S. N.) to calibrate the Continuum hydrological model that computes such state variables in a prognostic mode. The first part of the work aims at proving that satellite observations can be exploited to reduce uncertainties in parameter calibration by reducing the parameter equifinality that can become an issue in forecast mode. In the second part, four parameter estimation strategies are implemented and tested in a comparative mode: (i) a multi-objective approach that includes both satellite and ground observations which is an attempt to use dif-

ferent sources of data to add constraints to the parameters; (ii and iii) two approaches solely based on remotely sensed data that reproduce the case of a scarce data environment where streamflow observations are not available; (iv) a standard calibration based on streamflow observations used as a benchmark for the others.

Two Italian catchments are used as a test bed to verify the model capability in reproducing long-term (multi-year) simulations.

The results of the analysis evidence that, as a result of the model structure and the nature itself of the catchment hydrologic processes, some model parameters are only weakly dependent on discharge observations, and prove the usefulness of using data from both ground stations and satellites to additionally constrain the parameters in the calibration process and reduce the number of equifinal solutions.

1 Introduction

The estimation of parameters in hydrological models is still a challenge in hydrology. Much work has been devoted to determining the best calibration strategy (Yapo et al., 1998; Madsen, 2000; Kim et al., 2007; Singh and Bardossy, 2012; Xu et al., 2013) with some trying to evaluate the uncertainties associated with the parameter estimation process (Beven and Binley, 1992; Vrugt et al., 2003; Carpenter and Georgakakos, 2006; Zappa et al., 2011). This issue has become even more complex with the increasing use of continuous and distributed hydrological models. This trend led to a significant increase of the number of parameters that need calibration. A large number of parameters allow for good per-

formance in the calibration phase, but this can lead to a large number of equifinal parameter sets (Beven and Binley, 1992) sometimes hampering the forecast ability of the models.

Traditionally, the calibration of hydrological models requires the appropriate series of historical data, particularly of streamflow data, which are not easily available everywhere in the world; this condition aggravates the equifinality problem, which is higher the lower the observation capacity. Such issues raised the attention of the scientific community, becoming the focus of coordinated scientific initiatives (e.g., prediction in ungauged basin (PUB), a science initiative of the International Association of Hydrological Sciences, which was developed in the period 2003–2012). In a world where the data sharing capacity is increasing, it seems that the problem of data shortage for hydrologic model calibration is not going to disappear; the level gauge stations can be at a limited number of locations in a catchment and in some areas are very rare; additionally, in some areas the access to river-discharge information has been declining since the 1980s (Vörösmarty et al., 2001).

As a consequence, the use of remote sensing for direct streamflow measurements has received increased attention lately although, even if promising in some cases, it faces various technological, physical and scale limits. The more straightforward approaches use statistical relationships between remotely sensed river widths and in situ measurements (Brakenridge et al., 2005; Pavelsky, 2014) making them suitable for the extension of existing historical data, but unusable for ungauged sites. The limits are mainly due to the fact that accurate estimates of streamflow require the availability of several hydraulic parameters (width, depth, slope, channel morphology), which are difficult to derive entirely from remote sensing. Simplified models that make use of some of these parameters introduce uncertainties that limit their applicability; moreover, the detection of changes in hydraulic parameters has to do with the spatiotemporal resolution of the satellite sensors. These models (see Bjerklie et al. (2003) for a comprehensive review) are therefore not suitable for detecting changes in discharge for medium to small-scale basins (Brakenridge et al., 2012).

It is therefore compulsory to look at other possibilities offered by satellite sensors. Nowadays, the remote sensing of other meteorological, hydrological and ecological variables is more reliable and widely available at the global scale. Satellite products such as precipitation, shortwave and longwave radiation, atmospheric profiles, vegetation parameters, land-surface temperature (LST), evapotranspiration (ET) and digital elevation models (DEMs) are now operational and widely used in meteorological and hydrological modeling.

Experiments to understand the accuracy of these products are quite popular (see e.g., Bitew and Gebremichael, 2011; Brocca et al., 2011a; Crow et al., 2012; Göttsche and Hulle, 2012; Yu et al., 2012; Murray et al., 2013; Zhang et al., 2013). This kind of data is by now available for a very high percentage of the earth's surface, and covers most of the areas

where the density of ground stations is poor. This leads to a panorama in which estimating the parameters of a hydrological model by using only satellite information is a real possibility (Silvestro et al., 2013). However, the ability to calibrate a model using satellite data, even in combination with traditional in situ data, is still a challenging topic. Scientific work in this field goes in many directions: Rhoads et al. (2001) used satellite-derived LST to validate a land-surface model, Caparrini et al. (2004) and Sini et al. (2008) assimilated remotely sensed measurements into a land-surface model to estimate the surface turbulent fluxes, Brocca et al. (2011a) analyzed different remotely sensed soil humidity estimations with the perspective of using them in hydrological modeling, White and Lewis (2011) used satellite imagery to monitor the dynamics of wetlands of the Australian Great Artesian basin and Khan et al. (2011) have recently proposed a procedure to calibrate a fully distributed hydrological model using satellite-derived flood maps.

The objective of this work is to analyze the calibration skill of a distributed continuous hydrologic model by augmenting the model constraints with satellite-retrieved data. As a first analysis, in the context of a classical uncertainty analysis (Beven and Binley, 1992; Shen et al., 2012), it is shown that using satellite data together with ground station observations can reduce both parameter uncertainty and equifinality. The uncertainty analysis is also used here to define parameter sensitivity to different types of observation, with the goal of underpinning the importance of having a plurality of observations that might influence different models parameters in different ways.

After that, three simple calibration methods were applied in order to exploit the advantages of utilizing multi-sensor observations. The first method belongs to the family of the multi-objective calibration (M. O.) approaches (Efstratiadis and Koutsoyiannis, 2010) and simultaneously tries to exploit satellite and streamflow data. The second and third methods are two attempts at using only satellite data without any streamflow measurements, simulating the case of a basin in a scarce data environment. These last experiments are conceived along the lines of Silvestro et al. (2013), but with a more comprehensive approach that exploits both LST and SSM estimates from satellites. The results of the presented methods are then compared with those obtained using a fourth standard calibration methodology based on streamflow data.

The hydrological model used in the study is Continuum (Silvestro et al., 2013). It is a distributed continuous model conceived to satisfy the principle of parsimony in parameterization (Perrin et al., 2001; Coccia et al., 2009; Todini, 2009; Efstratiadis and Koutsoyiannis, 2010) and to be balanced between a good representation of physical processes and the simplicity of the schematizations and implementation.

The article is organized as follows: Sect. 2 provides a short overview of the Continuum hydrological model, the description of the data set used, the parameter uncertainty analysis

Table 1. The parameters of Continuum model that need calibration at basin scale are reported with a brief description.

Parameter	Physical process parameterized
u_h [s^{-1}]	Flow motion in hillslopes
u_c [$m^{0.5} s^{-1}$]	Friction in channels
c_f [–]	Infiltration capacity at saturation
c_t [–]	Mean field capacity
R_f [–]	Anisotropy between the vertical and horizontal saturated conductivity and soil porosity
$V_{W_{max}}$ [mm]	Maximum storage capacity of the aquifer

$$V_{fc} = c_t V_{max}, \quad (1)$$

where V_{max} is the maximum capacity of the soil to storage water in the root zone.

The infiltration capacity parameter c_f controls the velocity of sub-surface flow (i.e., it is related to saturated hydraulic conductivity), defining the asymptotic minimum infiltration rate for saturated soils f_1 with the following equation

$$f_1 = c_f f_0, \quad (2)$$

where f_0 is the maximum infiltration rate for completely dry soil.

The parameters c_t and c_f regulate the dynamics of saturation at cell scale. Since both f_0 and V_{max} are distributed parameters estimated as functions of curve number (Gabellani et al., 2008), the pattern of f_1 and V_{fc} is spatially modulated by the pattern of curve number maps (Silvestro et al., 2013) which are a synthetic representation of the local soil properties.

The parameters $V_{W_{max}}$ and R_f govern the deep flow and the water-table dynamic (Silvestro et al., 2013). $V_{W_{max}}$ represents the absolute maximum water content of the aquifer on the whole investigated area; the maximum water content on each cell is estimated based on $V_{W_{max}}$ and on the slope (Saulnier et al., 1997). R_f is a multiplicative factor in the Darcy equation used to estimate the flux per unit area between two contiguous cells and mainly takes care of differentiating the saturated vertical and horizontal conductivity. These two parameters have a reduced influence compared to the other four parameters because of the slow temporal dynamic of the water table. The sensitivity to R_f increases with the total basin drainage area when the effect of the interaction between the water table and the vadose zone becomes crucial in the formation of the recession curve between the rainfall events (Silvestro et al., 2013).

Continuum accounts for LST as an explicit state variable and allows for the estimation of the soil moisture in the root zone as the saturation degree (SD), which is defined here by the ratio of the actual soil water content $V(t)$ and the maximum storage capacity V_{max} (see Fig. 1):

$$SD = \frac{V(t)}{V_{max}}. \quad (3)$$

Both of these variables are represented at DEM spatial resolution.

The snow-melting process was not considered in Silvestro et al. (2013); since multi-year simulations are carried out in this work, a simple snow-melting model has been introduced and described in Appendix A.

2.2 Data set

The first test case is the Orba basin that is located in the Apennine part of the Piemonte region (Italy). It has a total area of approximately 800 km² and is a tributary of the Tanaro River (Fig. 2).

The Piemonte and Liguria region's meteorological networks monitor the basin. Data from rain gauges, thermometers, hygrometers, shortwave radiometers and anemometers are available with a temporal resolution of 1 h. Two stage-gauging stations are working with maintained stage-discharge rating curves; the two stations are located quite far from each other along the river: Tiglieto in a head catchment (drained area: 75 km²) and Casalcermelli near the basin outlet (drained area: 800 km²).

For this application, we extended the data set used in Silvestro et al. (2013). The chosen period starts on 1 June 2006 and ends on 31 December 2011. The first 5 months of 2006 are used as the model warm-up period.

The second test case is the Casentino basin (Fig. 2). It is a head catchment of the Arno River basin located in Tuscany. The watershed is located in the central Apennines with an elevation that ranges between 200 and 1600 m a.s.l. The mountainous part of the basin is mainly covered by forest, while cultivated fields or zones with low vegetation primarily makeup the flat areas. Urban areas cover a low percentage of the territory.

The two basins are only marginally impacted by snowfall and snow cover during winter.

The meteorological network of the Tuscany region provides rainfall, air temperature, air humidity, solar radiation and wind speed and direction with a temporal resolution of 1 h. Only one stage-gauging station (Subbiano) is working with a maintained stage-discharge rating curve; the gauge is located in the flat area of the basin at about 10 km from the confluence of the Casentino River along the Arno River (drained area 670 km²). The period of simulation ranges from 1 June 2005 to 31 December 2011. The first 5 months of 2005 are used to warm up the model.

In both cases the period has been chosen based on the data availability and in order to have reliable stage-discharge curves to estimate the observed streamflow.

The model temporal resolution is set to 1 h as the micro-meteorological observations; the surface flow needs a finer time step for computational stability reasons and it was fixed to 30 s.

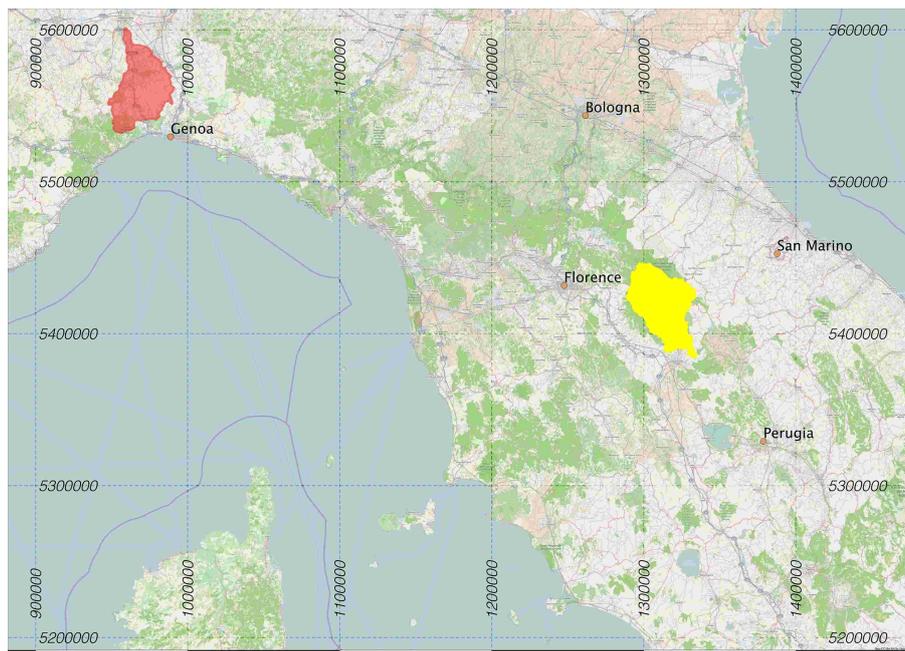


Figure 2. Location of the study areas. The Orba basin (in red) in the north-west of Italy and the Casentino basin (in yellow) in central Italy.

The remote-sensing data employed to implement the model and set additional constraints to the model parameters are

- the Istituto Geografico Militare (IGM) DEM used to extract the basin morphological parameters (<http://www.igmi.org/prodotti/datinumerici/datimatrix.php>);
- Land Surface Analysis Satellite Applications Facility (LSA-SAF) LST product retrieved from Meteosat Second Generation (MSG) observations (<http://landsaf.meteo.pt>);
- SM-OBS-1 surface soil moisture (SSM) retrieved from advanced scatterometer (ASCAT) (Wagner et al., 2013) and distributed within the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) program used as a benchmark to be compared with the model output (<http://hsaf.meteoam.it>);
- LSA-SAF LAI to parameterize the vegetation cover.

The DEM resolution is 0.0011° (about 100 m). The model spatial resolution is set equal to the DEM resolution. LST estimations are provided by LSA SAF of EUMETSAT (EUMETSAT, 2009). LST data have been available every 15 min with a spatial resolution of approximately 0.04° (about 4.5 km) since 2009. In order to compare model and satellite data, the approach followed by Silvestro et al. (2013) has been adopted. It allows LST obtained from the model to be projected to the same geometry of the satellite observations.

In order to carry out comparison at basin scale the mean of the pixel values has been used:

$$\overline{\text{LST}} = \frac{1}{N} \sum_{i=1}^N \text{LST}_i, \quad (4)$$

where N is the number of cells of the spatial grid, and LST_i the LST of i th cell. The comparison is carried out at those instants where both model and satellite data are available (1 h resolution) and if at least 50 % of satellite pixels that cover the basin have reliable data (e.g., in case of bad weather no-data values can be found in satellite product).

The H-SAF SM-OBS-1 product consists of European maps of large-scale SSM retrieved from ASCAT, the active microwave sensor, which flies on board two polar-orbiting Meteorological Operational (METOP) satellites. This product gives soil moisture estimates across different test sites in Europe, the Americas and Africa (Brocca et al., 2010; Albergel et al., 2012). EUMETSAT makes the product available, from 3 June 2009, in near real time with a spatial resolution of approximately 25 km and a revisit time of twice a day.

The SSM have been re-sampled to the model resolution using the nearest neighbor method. SM-OBS-1 (H07) data with quality flag (provided with the product) greater than 15 were discarded. Since the product is referred to the first centimeters of soil, the soil water index (SWI) method, developed by Wagner et al. (1999), was applied to SSM satellite data to obtain an estimate of the saturation degree in the root zone. This filter allows the SSM estimates to be related to the profile soil moisture content. It relies on the assumption that the varia-

tion in time of the average value of the soil moisture profile is linearly related to the difference between the surface and the profile values. In this study a simple recursive formulation of the method was used (Stroud, 1999; Albergel et al., 2008) and the characteristic time length T , which represents the timescale of soil moisture variation (Wagner et al., 1999), was considered equal to 10 days. Since in situ soil moisture measurements are not available and the soil properties are not known quantitatively with high detail, the parameter T was set to a priori value that was estimated, as an order of magnitude, using the definition of T from Wagner et al. (1999) and also from Parajka et al. (2006) based on the mean soil characteristics of the considered catchments as described in the model (the average potential soil moisture capacity of the considered basins is around 150–170 mm; assuming a porosity of 0.3, a pseudo diffusivity of 10 days would then translate into a wetting front celerity of around 50 mm per day which is a reasonable value for these soils). The SWI was then rescaled to model climatology using a minimum–maximum correction technique (Brocca et al., 2013). After the rescaling, the mean SWI at basin scale was computed as a simple average of the values of pixels that cover the basin and used for comparison with model SD. The comparison is carried out at those instants where both model and satellite data are available (twice a day) and if at least 50 % of satellite pixels that cover the basin have reliable data (quality flag greater than 15). LAI maps were produced with a temporal update of 15 days as averaged values of daily LSA-SAF maps at a spatial resolution of 0.04° (EUMETSAT, 2008) and gridded with the nearest neighbor method on model resolution.

2.3 Statistics and scores

A series of statistics and scores are used to carry out the uncertainty analysis, the calibration and the validation of the Continuum model. They are presented and described in this section. A different subsection is used for each of the three considered observable variables: streamflow, LST and SWI.

2.3.1 Streamflow

The Nash–Sutcliffe (NS) coefficient (Nash and Sutcliffe, 1970) was chosen as the main likelihood function since it is one of the most widely used measures to evaluate model performances in hydrology:

$$NS = 1 - \frac{\sum_{t=1}^{t_{\max}} (Q_m(t) - Q_o(t))^2}{\sum_{t=1}^{t_{\max}} (Q_m(t) - \bar{Q}_o)^2}, \quad (5)$$

where $Q_m(t)$ and $Q_o(t)$ are the modeled and observed streamflows at time t . \bar{Q}_o is the mean observed streamflow.

Five other scores were evaluated to assess the model performance (Madsen, 2000; Batholomes and Todini, 2005):

the Chiew–McMahon (CM) coefficient (Chiew and McMahon, 1994):

$$CM = 1 - \frac{\sum_{t=1}^{t_{\max}} (\sqrt{Q_m(t)} - \sqrt{Q_o(t)})^2}{\left(\sum_{t=1}^{t_{\max}} (\sqrt{Q_m(t)} - \sqrt{\bar{Q}_o})\right)^2}; \quad (6)$$

the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} (Q_m(t) - Q_o(t))^2}; \quad (7)$$

the correlation coefficient (CORR):

$$CORR = \frac{\sum_{t=1}^{t_{\max}} (Q_m(t) - \bar{Q}_m) \cdot \sum_{t=1}^{t_{\max}} (Q_o(t) - \bar{Q}_o)}{\sqrt{\left(\sum_{t=1}^{t_{\max}} (Q_m(t) - \bar{Q}_m)^2\right) \cdot \left(\sum_{t=1}^{t_{\max}} (Q_o(t) - \bar{Q}_o)^2\right)}}, \quad (8)$$

where \bar{Q}_m is the mean modeled streamflow;

the relative error (Rel. Err.):

$$Rel. Err. = \frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} \frac{|Q_m(t) - Q_o(t)|}{Q_o(t)}; \quad (9)$$

the peak flow relative error (PFRE):

$$PFRE = \frac{1}{N_{\text{peaks}}} \sum_{tp=1}^{N_{\text{peaks}}} \frac{Q_{pm}(tp) - Q_{po}(tp)}{Q_{po}(tp)}, \quad (10)$$

where $Q_{pm}(tp)$ and $Q_{po}(tp)$ are the modeled and observed peak flows. The peak flows are selected considering the values larger than a fixed threshold Q_{th} .

NS, CM and CORR indicate good matching between model and observations when they are close to 1, and RMSE, Rel. Err. and PFRE when they tend to 0.

2.3.2 Land-surface temperature

The bias between modeled and satellite-derived LST was considered as a skill score:

$$BIAS = \frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} \left| \overline{LST_m(t)} - \overline{LST_s(t)} \right|, \quad (11)$$

where $\overline{LST_m(t)}$ and $\overline{LST_s(t)}$ are the modeled and satellite LST averaged at basin scale at time t .

We used the BIAS in order to check the capability of the model to reproduce the mean LST on the selected period, more than it would for the overall shape of the time series.

2.3.3 Soil water index

We considered NS as a score to evaluate the performances of the model to reproduce the SWI derived by the satellite

observations. In this case, considering SD and SWI to be directly comparable, (SD ≡ SWI) NS is defined as

$$NS = 1 - \sum_{t=1}^{t_{\max}} \frac{(\overline{SWI_m(t)} - \overline{SWI_s(t)})^2}{(\overline{SWI_m(t)} + \overline{SWI_s(t)})^2}, \quad (12)$$

where $\overline{SWI_m(t)}$ and $\overline{SWI_s(t)}$ are the modeled and satellite SWI averaged at basin scale at time t .

$\overline{SWI_s}$ is the satellite SWI averaged in space and time.

2.4 Experimental setup

2.4.1 Uncertainty analysis

The issue of model parameter uncertainty and sensitivity has been one of the main themes of scientific discussions over the last 30 years. Many authors faced the problem following different approaches (see e.g., Beven and Binley, 1992; Liu et al., 2005; Carpenter and Georgakakos, 2006; Zappa et al., 2011; Rakovec et al., 2014), but it is widely accepted and recognized that parameter uncertainty is inevitable and that an optimal set of parameters rarely exists that allows for the best performance of the model in every condition; generally, there are multiple sets of parameters able to give similar results and which are therefore equivalent if the final aim is identified, i.e., the so-called equifinality (Savenije, 2001).

In this work, we did not carry out a full predictive uncertainty analysis, but we analyzed the parameter uncertainty based on equifinal realizations obtained by a Monte Carlo experiment; hence to achieve this some concepts of the generalized likelihood uncertainty estimation (GLUE) method (Beven and Binley, 1992) are used, similarly to what was done by Zappa et al. (2011) and Shen et al. (2012). Finally, we made reference to the work of Liu et al. (2005) in order to estimate the probability of parameter couples conditioned to the observations.

The concepts of the GLUE approach are applied using objective functions (scores) based on streamflow, LST and SWI in order to analyze how these variables, which are modeled by Continuum and measured through ground-based or remote measurement systems, are related to the model parameters. The main objective of the analysis is to study the dependence of each single parameter from the observed variables in order to pinpoint the importance of the different observations in determining the parameter set performing best in reproducing the full set of observations.

The uncertainty analysis has been done in the Orba basin by considering the four most sensitive parameters of the model as in Liu et al. (2005) (c_t , c_f , u_c , u_h ; for Continuum, see Silvestro et al., 2013, for details).

First, the analysis based on streamflow is done using NS as likelihood function (Eq. 5). The sampling space of the four parameters was defined by combining the literature (Beven and Binley, 1992; Liu et al., 2005; Zappa et al., 2011; Shen et

Table 2. Range of variability of the parameters used in the calibration and uncertainty analysis.

Parameter	Unit	Min	Max
c_t	[-]	0.15	0.65
c_f	[-]	0.015	0.1
u_c	$m^{0.5} s^{-1}$	15	55
u_h	1/s	0.0002	0.0015

al., 2012) with the results of the preliminary sensitivity analysis done by Silvestro et al. (2013) and considerations on the role and physical meaning of the parameters themselves. In Table 2 the range of variability of the parameters is reported.

The other two parameters were set based on physically reasonable values (due to the morphology and the soil type of the basins) assuming that there is no additional information about them. $V_{W_{\max}}$ is set equal to 2000 mm, and R_f is set equal to 1, which indicates a weak anisotropy between vertical and horizontal saturated conductivity. These two parameters, which represent deep soil processes, are only weakly related to the processes that influence LST and SWI observations and hence they are unlikely to be influenced by the chosen calibration strategy.

The analysis was done by simulating the four parameters c_t , c_f , u_c , u_h , and generating a set of 3000 streamflow simulations for the sub-period 16 August 2006 to 30 September 2006. The parameters have been extracted from a multi-uniform distribution bound in the domain of the parameters. The chosen sub-period includes various streamflow regimes.

As a further method to deepen the parameter uncertainty assessment the original data were transformed into a Gaussian space and ranked in increasing order once standardized (see Liu et al., 2005, for details). Observed and modeled data are then related as follows:

$$\eta_o = \eta_s + \xi, \quad (13)$$

where η_o and η_s are the normalized vectors of observed and modeled streamflow with 0 mean and unit variance, and ξ is the error vector. The likelihood function L_j for the j th parameter set after I_{\max} simulation steps can be expressed as (Xu et al., 2013)

$$L_j = \exp\left(-\frac{1}{2} \cdot \sum_{i=1}^{I_{\max}} \xi_{i,j}^2\right). \quad (14)$$

This function, when properly scaled, can be considered as the posterior parameter probability density.

Similarly to what has been done with streamflow data, by comparing modeled and satellite-derived LST and SWI (mean at basin scale) it is possible to carry out the uncertainty analysis to understand how these two variables are related to the model parameters.

In the case of LST the considered analysis period is August–September 2009. The BIAS between modeled and satellite-derived LST was considered as a skill score.

The procedure was then applied to the ASCAT SSM data after their transformation in SWI (Wagner et al., 1999). The considered period is August–October 2011. The model SD and the satellite SWI maps have been averaged at basin scale and the resulting time series have been used to compute the NS that, as for the hydrograph, has been considered as a score to evaluate the performances of the model in reproducing the SWI derived by the satellite observations.

2.4.2 Parameters estimation methodologies

To investigate the different ability of traditional streamflow-based and satellite-based calibration methods in identifying soil parameters is one of the objectives of this study. For this purpose, three satellite-driven calibration methodologies have been designed and benchmarked to a standard streamflow-based methodology for the detection of the most sensitive and impacting parameters of the model (u_c , u_h , c_t , c_f).

Calibration based on streamflow observations (S. N.): the benchmark

A methodology based on the maximization of the Nash–Sutcliffe coefficient between observed and modeled streamflow time series has been considered as a benchmark in order to compare the methods described in the next two subsections with a more standard approach; hence, hereafter we will call this method S. N.

Multi-objective calibration (M. O.)

The M. O. approach has been designed in order to exploit the use of all the information available for the calibration process that is generally represented by different observed variables. The method is based on the setup of a multiple objective function such that

$$\text{Min}\{F_1(\theta), F_2(\theta), \dots, F_n(\theta)\} \text{ with } \theta \in \Theta. \quad (15)$$

θ is restricted to the feasible parameter space Θ (Madsen, 2000; Kim et al., 2007; Efstratiadis and Koutsoyiannis, 2010).

This calibration approach is based, in our case, on the comparisons of (ground or satellite) observed vs simulated streamflow, LST and SWI.

The multi-objective function is designed through the following single objective function:

$$F_1 = \left[\sum_{t=1}^{t_{\max}} \frac{(Q_m(t) - Q_o(t))^2}{(Q_m(t) - < Q_o >)^2} \right], \quad (16)$$

$$F_2 = \left[\sum_{t=1}^{t_{\max}} |LST_m(t) - LST_s(t)| \right], \quad (17)$$

$$F_3 = \left[\sum_{t=1}^{t_{\max}} \frac{|Q_m(t) - Q_o(t)|}{Q_o} \right]_{Q_o > Q_T}, \quad (18)$$

$$F_4 = \left[\sum_{t=1}^{t_{\max}} \frac{(SWI_s(t) - SD_m(t))^2}{(SWI_s(t) - < SD_m >)^2} \right], \quad (19)$$

where Q is the streamflow; LST the mean LST at basin scale; SWI and SD are the mean of soil water index and SD at basin scale; subscripts m, o and s indicate model, gauge observations and satellite estimation, respectively; t is the time; and Q_T is a discharge threshold. Different periods for LST, SWI and Q could be considered in order to choose the most suitable time window in terms of availability of data and good representativeness of the variable dynamic.

Following Madsen (2000), the n (where $n=4$) contributors F_i have been combined in the following way:

$$F_{\text{adj}} = \left[(F_1 + A_1)^2 + \dots + (F_n + A_n)^2 \right]^{0.5}, \quad (20)$$

where the transformation factor A_i is calculated as

$$A_i = \text{MAX}(F_{j,\text{min}}; j = 1, 2, \dots, n) - F_{i,\text{min}}. \quad (21)$$

F_{adj} (Eq. 20) consists in four terms and A_i, \dots, A_n have the role of balancing the weights of the different objectives that can have, as in the presented application, different ranges and units.

The term F_1 depends on the streamflow and it is the complement to 1 of the Nash–Sutcliffe coefficient; the term F_2 depends on LST and it is the mean of the absolute errors (absolute BIAS) calculated at each time step; F_3 is a relative error estimated on streamflow values larger than a threshold and it is useful to reproduce flow peaks, and F_4 depends on the soil humidity and it is the complement to 1 of the Nash–Sutcliffe coefficient. The different terms have also been chosen to be consistent with the results shown in the uncertainty analysis. All the components tend to 0 when simulated and observed variables coincide, so that the calibration process consists in the minimization of the function F_{adj} . The resulting parameters set is representative of a balance point of the multi-dimensional Pareto front due to the different components of the multi-objective function (Madsen, 2000, 2003).

The introduction of remotely sensed LST and SWI in a multi-objective calibration is a new approach since objective functions are usually based on parameters related to observed hydrographs (e.g., total volume as in Yapo et al., 1998; Efstratiadis and Koutsoyiannis, 2010). This approach follows and improves the investigations carried out by other authors (Crow et al., 2003; Santanello et al., 2007; Koren et al., 2008; Flores et al., 2010; Montzka et al., 2011; Ridler et al., 2012; Corbari and Mancini, 2014; Sutanudjaja et al., 2014; Wanders et al., 2014) who attempted to combine, in the calibra-

Table 3. Characteristics of the synthetic rainfall events used for the estimation of the two surface parameters in the R. S. approach. The event length is set depending on the considered basins and sections. Sections with smaller upstream drainage area are used to estimate u_h while sections with larger area to estimate u_c .

Parameter	P_{cum} [mm]	Reference section	Area [km ²]	t_{lo} [h]	Duration [h]
u_h [s ⁻¹]	10, 20, ..., 60, 70	Tiglieto	75	4.5	4
u_c [m ^{0.5} s ⁻¹]	10, 20, ..., 60, 70	Casalcermeli	800	11.6	10
u_h [s ⁻¹]	10, 20, ..., 60, 70	Upstream (no gauge)	58	2.98	4
u_c [m ^{0.5} s ⁻¹]	10, 20, ..., 60, 70	Subbiano	670	8.4	9

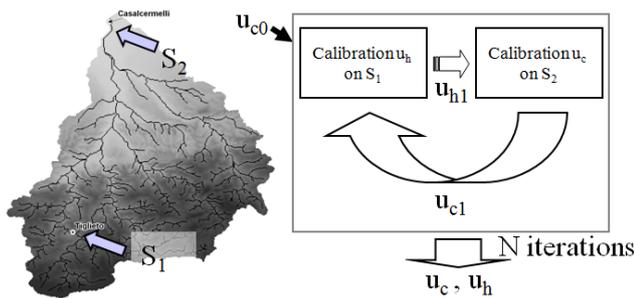


Figure 3. Scheme of calibration of the surface parameters in R. S. approach based on the reproduction of the basin lag time at two different closure sections using synthetic rainfall events. A first-guess u_c is used to calibrate u_h on section S_1 , this latter is used to calibrate u_c on S_2 , then the procedure is iterated for N times.

tion process, remote-sensed and in situ observations of variables other than streamflow. The target here is not applying and testing quite sophisticated or complex algorithms for calibration, like those described in Yapo et al. (1998) or Vrugt et al. (2003), rather to assess if the use of satellite observations leads to an advantage in model calibration with respect to its capability of simulating discharge values. Here, a very simple brute-force calibration approach was used.

Remote-sensing data calibration approach (R. S.)

When no streamflow data are available, we can still calibrate the model on satellite data, LST or SWI (derived by SSM), and on the morphologic characteristics of the basin extracted from the DEM. This methodology was presented in Silvestro et al. (2013) with respect to LST and it investigates the possibility of calibrating a subset of model parameters in an ungauged basin. Since it can be applied using LST or SWI we have two calibration methods to be tested.

The morphologic characteristics mainly influence the surface flow while LST and SWI are more related to sub-surface flow.

The estimation of the overland and channel flow parameters is carried out by using geomorphological information derived from the DEM. The methodology is described in Sil-

vestro et al. (2013), and we synthetically report its description in the following steps and in Fig. 3:

1. Step 1: identify a formulation to estimate the typical lag time (t_{lo} : temporal distance between the center of mass of the hydrograph and the center of mass of the mean hyetograph) of a basin based on its main morphologic characteristics. The soil was considered to be completely impermeable so that the sub-surface and deep flow parameters (c_t, c_f, R_f and V_{Wmax}) therefore become irrelevant.
2. Step 2: identify two sections along the streamline of the basin, one at the head of the basin and the other downstream. Estimate the time lag, t_{lo} based on the DEM and geographical information.
3. Step 3: generate a set of synthetic events with constant intensity in space and having a duration equal to the typical response time of the basin closed at the abovementioned sections (see Table 3).
4. Step 4: set a first estimate for the value of u_c and calibrate u_h for each value of P_{cum} referring to the upstream section, using the objective function to minimize:

$$Of = |t_{lo} - t_{lm}|, \tag{22}$$
 where t_{lo} is the t_l derived by the geomorphologic characteristics of the basin while t_{lm} is the t_l obtained from the model simulations. Calculate the average of the u_h values.
5. Step 5: fix u_h and calibrate u_c as in Steps 3 and 4, referring this time to the downstream section.
6. Step 6: iterate the process until it converges.

According to Silvestro et al. (2013), it is possible to separate the calibration of the two surface flow parameters. In the case of the head section with reduced paths in the channelized network, the influence of u_c is scarce; as a consequence, an average value of u_c can be set and the calibration can be done only for u_h . The value of u_c is then calibrated based on data from a downstream section with a longer channelized

Table 4. Summary of the periods considered for the calibration process for each variable. In the case of M. O. approach the three periods are merged.

Basin	Q	LST	SWI
Orba	Aug–Oct 2006	Jul–Oct 2009	Jul–Nov 2011
Casentino	Sep–Dec 2005	Jul–Oct 2009	Jul–Nov 2011

network. This procedure is iterated as shown in Fig. 3. Usually 3–4 iterations are sufficient for a good convergence of the process.

Once the surface flow parameters are estimated, the sub-surface soil parameters can be evaluated optimizing a proper score between satellite derived and modeled LST (Silvestro et al., 2013) or SWI. In this case we considered the BIAS and Nash–Sutcliffe coefficient at basin scale as scores for LST for SWI, respectively.

2.4.3 Calibration settings

In both test cases a calibration period has been chosen for each variable (Streamflow, LST, SWI). Using a same period for all the variables is not always the best option, in fact they describe components of the hydrological cycle that need to be sampled in different periods of time; moreover, in certain cases data are available on different periods that do not always overlap. Furthermore, it is interesting to understand what kind of results can be achieved reducing the length of calibration periods but augmenting the number of observed variables. Alternatively, in the case data are available, one could work on longer periods of time that ensure to catch the seasonality of the hydrological processes but with the disadvantage of lengthening the calculation time.

The four calibration strategies that have been compared are the following:

- The standard approach S. N. described in the “Calibration based on streamflow observations (S. N.): the benchmark” section.
- The M. O. strategy described in the “Multi-objective calibration (M. O.)” section (in this case the calibration periods of the different observations are merged).
- The R. S. approach described in the “Remote-sensing data calibration approach (R. S.)” section using the satellite LST as comparison data. Hereafter we will call this strategy R. S. (LST).
- The R. S. approach described in the “Remote-sensing data calibration approach (R. S.)” section using SWI estimation derived from satellite SSM as comparison data. Hereafter we will call this strategy R. S. (SWI).

The periods used for calibration are chosen in order to balance the following characteristics:

- have the presence of different streamflow regimes and soil moisture conditions;
- presence of extreme conditions: e.g., flood and drought periods;
- presence of reliable data, especially for SWI and LST comparison;
- having periods’ length manageable in terms of computational time.

In the case of the Orba basin the calibration was carried out considering the period July–October 2009 for LST comparison, July–November 2011 for SWI comparison and August–October 2006 for streamflow comparison (in this latter period, two intense events preceded by periods of droughts occurred; as a result, the model is forced to work under extreme conditions). The streamflow threshold used in the third component of the M. O. function is $Q_T = 200 \text{ m}^3 \text{ s}^{-1}$. Validation of multi-annual simulations were carried out using the parameters calibrated with the proposed methodologies. The validation period is from 1 January 2006 to 31 December 2011; the first 5 months were used for the model warm-up.

In the case of the Casentino basin the calibration was carried out considering the July–October 2009 for LST comparison, July–November 2011 for SWI comparison and September–December 2005 for streamflow comparison. The periods were chosen based on the same constraints presented for the Orba case study (see previous paragraph). The streamflow threshold used in the third component of the M. O. function was $Q_T = 200 \text{ m}^3 \text{ s}^{-1}$. The validation period is from 1 January 2005 to 31 December 2011; the first 5 months were used for the model warm-up.

Table 4 summarizes the calibration periods for the two test cases.

3 Results

3.1 Uncertainty analysis

Each score based on streamflow data, and presented in Sect. 2.3.1, can be influenced differently by different flow regimes and hydrograph characteristics; therefore, for each simulation the NS was plotted against the other scores (Zappa et al., 2011); the results are reported in Fig. 4 and the graphs show that in all cases there are sets of behavioral parameters (Beven and Binley, 1992) that give similarly good values of the scores, indicating good simulation of the observed streamflow series.

In Fig. 5, the dotted plots of the four parameters are reported. Each graph shows the NS value as a function of the parameter values. The variability of NS for a single parameter is quite high. In the case of the two surface parameters, u_c and u_h (upper subplots in the figure), a maximum for NS can

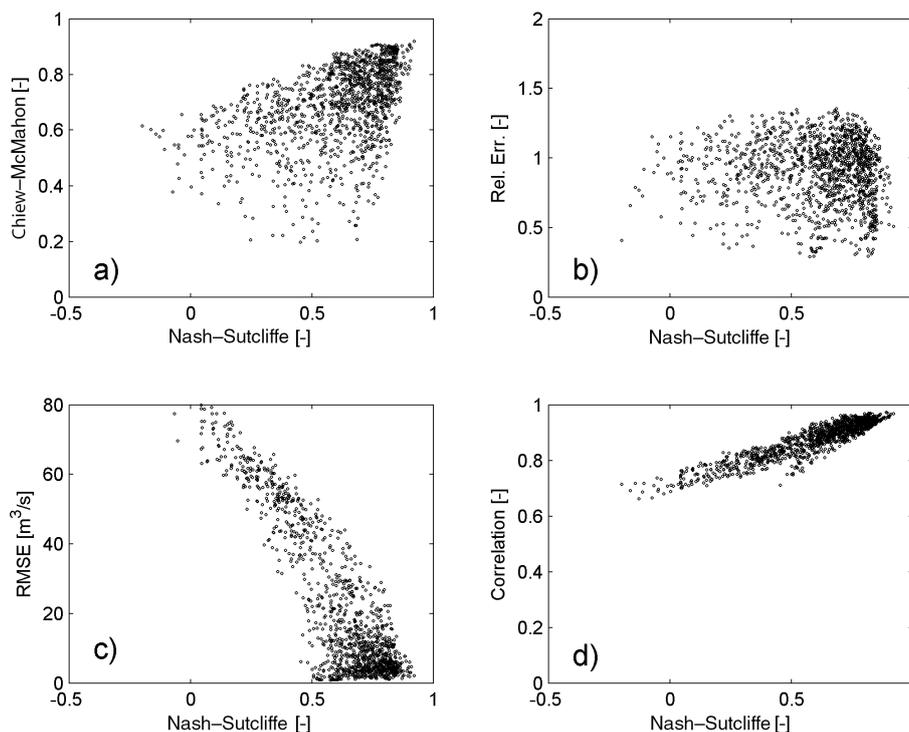


Figure 4. The Orba basin, August–September 2006. NS on the hydrographs (x axis) versus other statistics (y axis): Chiew–McMahon coefficient (CM), root mean square error (RMSE), the correlation coefficient (CORR) and the relative error (Rel. Err.). The graphs show that in all cases there are sets of parameters that give similarly good values of the compared scores.

be identified, while for c_t and c_f the behavior of NS is quite homogeneous for all the values in the physically acceptable range. This indicates that u_c and u_h are closely linked to the streamflow simulation in the model regardless of the other parameter values, while the impact of c_f and c_t in the discharge follows more complex paths, and it is hard to identify such parameters by matching the streamflow time series alone. For values of u_c greater than $30\text{--}35\text{ m}^{0.5}\text{ s}^{-1}$ and values of u_h greater than $7\text{--}9\text{ s}^{-1}$, the values of the NS coefficient seem to be uniformly distributed over a large range, this indicates that different combinations of the two parameters can lead to very different performances. For lower values of u_c and u_h , the NS coefficient converges to high values highlighting a minor variability of the score and general better performances.

Simulations with NS lower than a fixed threshold ($\text{NS} = 0.4$) are considered non-behavioral according to Shen et al. (2012). By sorting the discharge time series according to NS values it is possible to evaluate the percentiles at each time step and show the uncertainty in terms of confidence intervals. In Fig. 6, the 10 and 90 % confidence limits are reported for two time windows across the main streamflow events, which occurred in the considered period. The results show that the observed streamflow lays at the 90 % limit; therefore, a parameter configuration that allows for the reproduction of flow observations exists at any time. Most of

the observed hydrographs, and specifically the peak flow, lay in the 10 % limits. Part of the receding curve is not included showing some limited ability of the model in the representation of the processes related to the drainage of the soil and aquifers.

The results of the analysis done using the likelihood function represented by Eq. (14) are presented in Figs. 7 and 8 where the probability density is plotted considering two parameters at a time. In this case, a more evident concentration of the likelihood function appears when compared to the dotty-plot representation of the NS score presented in Fig. 5. This is again valid, especially if the case of parameters u_c versus u_h is considered (Fig. 7). Nevertheless, various relative maximums of the likelihood function are present.

The uncertainty analysis using skill scores based on streamflow provides evidence of the presence of equifinal sets of model parameters, this behavior can be found in other continuous and distributed models; nevertheless, there is a reduced number of parameter sets that generate evidently better performances among all the possible configurations randomly generated. This raises the necessity of finding additional constraints to improve the estimation of the parameters. The focus was then placed on two meteo-hydrological variables whose observations are now widely available from remote-sensing techniques: LST and SWI.

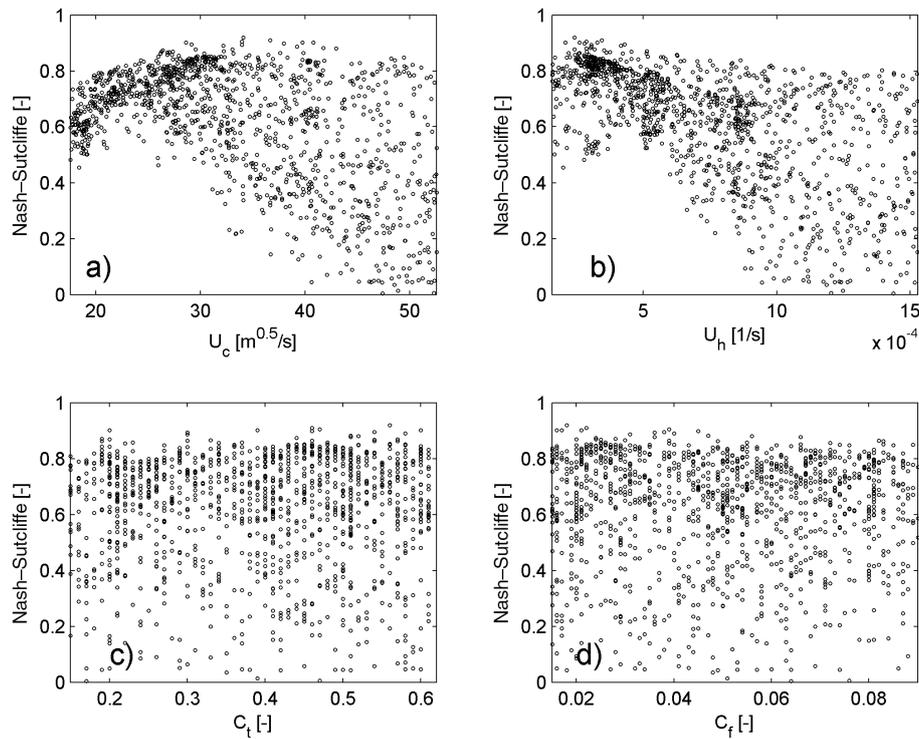


Figure 5. The Orba basin, August–September 2006. Dotty-plot representation of the NS on the hydrographs. Each subplot reports the parameter value on x axis and the value of NS statistic on y axis. The surface parameters, (a) and (b), have quite clear ranges where NS reaches maximum values independently from other parameters. This does not occur for sub-surface flow parameters, (c) and (d).

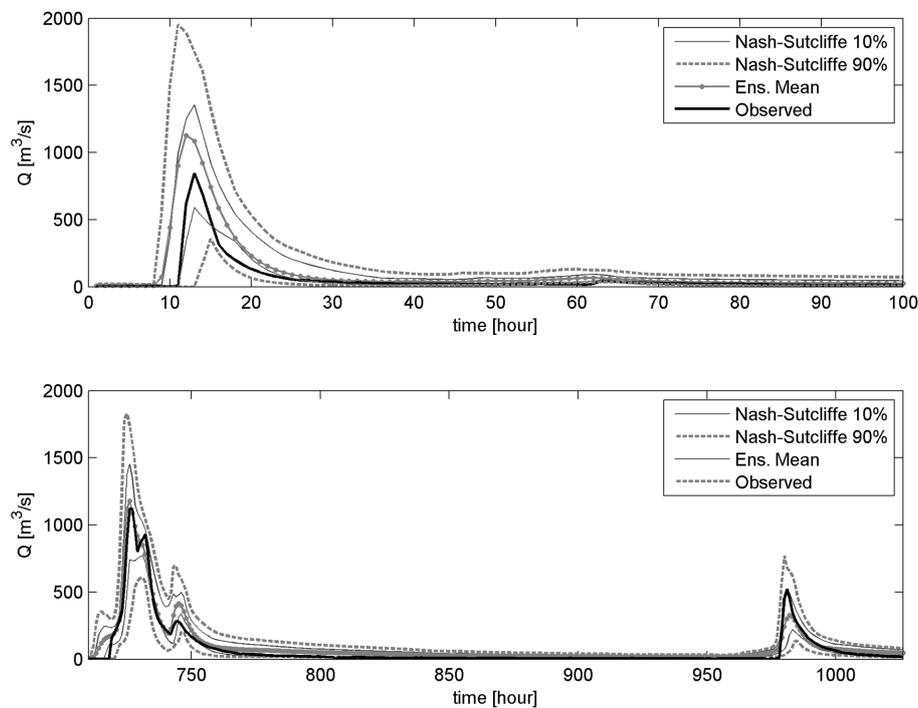


Figure 6. The Orba basin, August–September 2006. Confidence intervals compared with observed streamflow. Two representative periods are shown. x axis reports time while y axis reports streamflow. Observed streamflow is compared with the confidence intervals based on NS statistic and the ensemble mean.

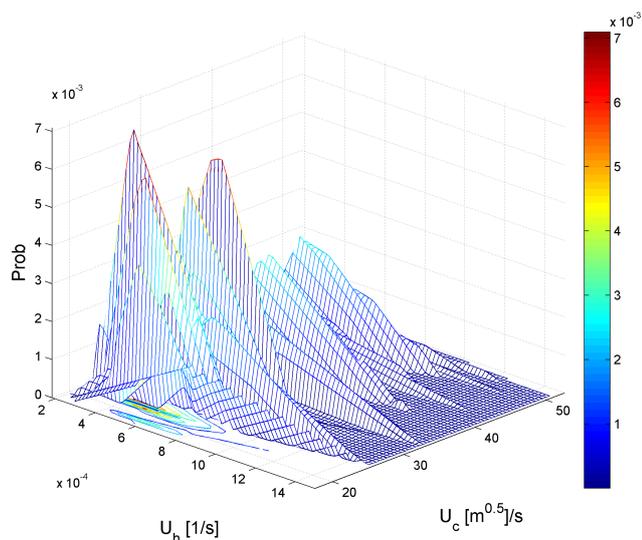


Figure 7. The Orba basin, August–September 2006. Example of a two by two sensitivity analysis: surface parameters, u_c versus u_h . On x and y axis the values of the two parameters are reported while on z axis the parameter likelihood is shown.

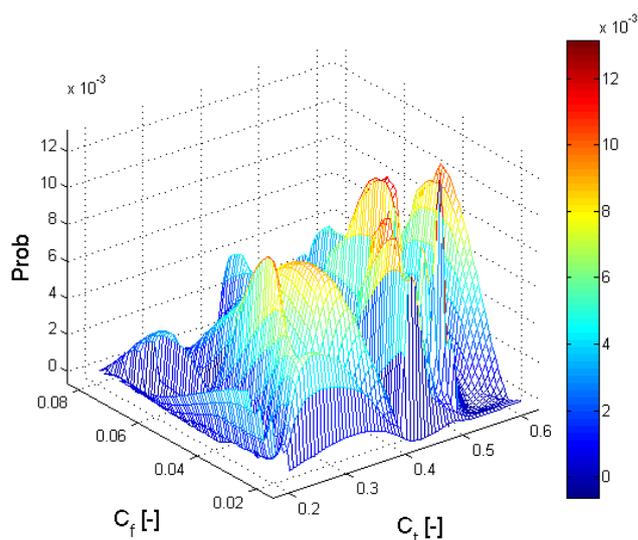


Figure 8. The Orba basin, August–September 2006. Example of a two by two sensitivity analysis: sub-surface parameters, c_t versus c_f . On x and y axis the values of the two parameters are reported while on z axis the parameter likelihood is shown.

By comparing modeled and satellite-derived LST (mean at basin scale), it is possible to build dot-plot representations similar to those presented in Fig. 5 but using the BIAS score (Eq. 11). Figure 9 shows that it is almost impossible to find a well defined or unique set of surface parameters that minimize BIAS on LST. The same considerations can be done for c_f . The c_t shows an evident trend: this is reasonable since this parameter strongly influences the time of permanence of water in the soil and the LST diurnal dynamics (Caparrini et al., 2004; Sini et al., 2008; Silvestro et al., 2013).

Finally, the time series obtained by averaging at basin scale the model SD and the satellite SWI maps have been used to build a dot-plot graph using NS described in Eq. (12) as score (Fig. 10). The maximum of NS lies in the range 0.45–0.55 of the parameter c_t ; and a weak, but quite evident independence of c_f arises with optimal values around 0.015–0.025 (close to the lower limit of the parameter range). In both cases the NS values are in a quite narrow range in correspondence of the aforementioned parameter ranges.

The values of the parameters individuated by the LST and SWI analysis are consistent with some of the best equifinal parameter combinations on the basis of the streamflow analysis.

3.2 Calibration and validation on multi-yearly simulations

The results of the uncertainty analysis presented in Sect. 3.1 show that it is possible to use ground and remote-sensing observations in order to reduce equifinality of the parameters defining the hydrological model. The analysis evidences

Table 5. The Orba basin parameters were calibrated following the different approaches: using the multi-objective function (M. O.), remote-sensing data and morphologic characteristics (R. S.), and the standard hydrographs comparison (S. N.). The table reports the values obtained by the 4 different calibration strategies.

Basin	Orba			
	S. N.	R. S. (LST)	R. S. (SWI)	M. O.
u_c [$m^{0.5} s^{-1}$]	29.0	29.42	29.42	29.92
u_h [1/s]	0.00052	0.000458	0.000458	0.00041
c_t [-]	0.52	0.56	0.58	0.40
c_f [-]	0.020	0.030	0.018	0.020

that is possible to identify relatively narrow ranges of parameters' values that optimize the scores based on different observations, each parameter being mainly linked to one or more observed variables and less to others. The uncertainty analysis also highlights that streamflow-based calibration offers very limited capability to detect soil parameters and that satellite-based observations deliver a complementary capability in this respect. To exploit this opportunity, the three parameters estimation methodologies presented in Sect. 2.4.2 that use satellite data have been tested and compared with the more standard calibration method based on streamflow data described as well in Sect. 2.4.2.

3.2.1 The Orba basin

The sets of parameters obtained by the four calibration strategies are reported in Table 5.

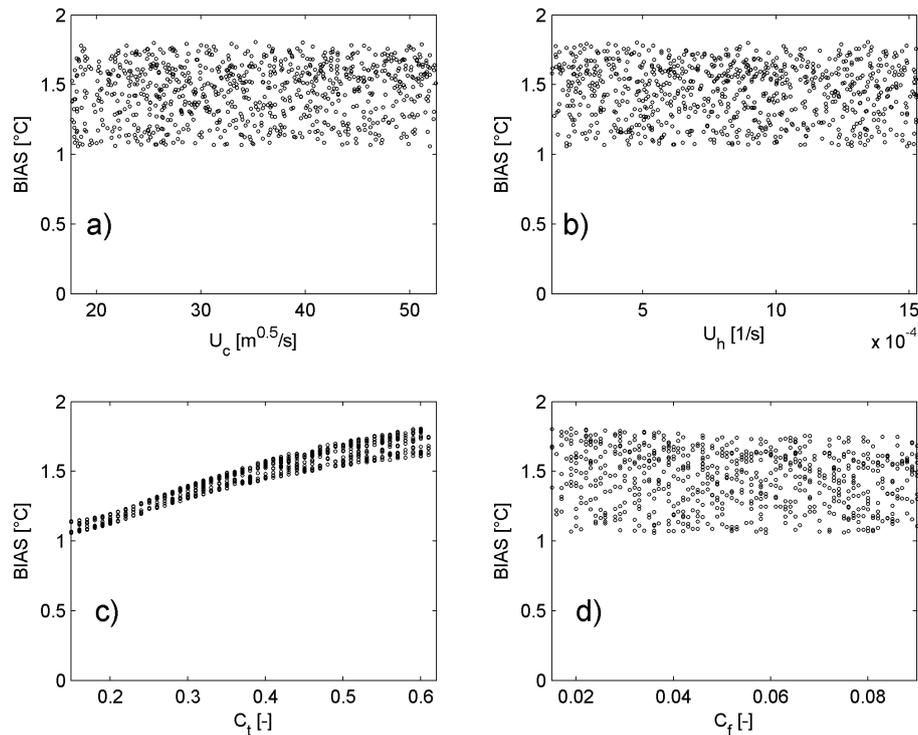


Figure 9. The Orba basin, August–September 2009. Dotty-plot representation of the BIAS on mean LST at basin scale. Each subplot reports the parameter value on x axis and the value of BIAS statistic on y axis. The parameter c_t (c) has a quite clear range where BIAS reaches minimum values independently from the other parameters. This does not occur for the other parameters (a, b, d).

Table 6. The Orba basin skill scores on hydrographs for the entire validation period (2006–2011) and for the different parameter sets. The scores value is calculated for the two sections with available streamflow observations. PFRE is calculated with $Q_{th} = 200 \text{ m}^3 \text{ s}^{-1}$ for Casalcermelli section and with $Q_{th} = 50 \text{ m}^3 \text{ s}^{-1}$ for Tiglieto section.

Basin	Section	Parameter set	N. S.	CM	RMSE	CORR	PFRE
Orba	Casalcermelli	M. O.	0.82	0.81	1.67	0.91	−0.162
		R. S. (LST)	0.81	0.83	1.41	0.90	−0.133
		R. S. (SWI)	0.82	0.82	1.35	0.90	−0.063
		S. N.	0.83	0.82	1.31	0.91	−0.095
	Tiglieto	M.O.	0.69	0.65	0.80	0.87	−0.251
		R. S. (LST)	0.67	0.62	0.78	0.83	−0.245
		R. S. (SWI)	0.66	0.63	0.76	0.81	−0.143
		S. N.	0.66	0.62	0.70	0.80	−0.213

The surface flow parameters obtained with the R. S. and M. O. calibration methodologies are similar, while the S. N. method produces slightly lower u_c and higher u_h values. In the case of sub-surface flow, the values are a little bit different for the three considered cases probably because they are more sensitive to the different adopted approaches (Efstratiadis and Koutsoyiannis, 2010).

In the case of R. S. (SWI), the calibrated c_t and c_f values confirm the results of Sect. 3.1, even if the dotty plot in Fig. 10 does not show a really strong independence of c_t and

c_f from the other parameters, but more a range of the two parameters that shows good simulations of SWI.

In the case of R. S. (LST), results are different in respect to the results of Sect. 3.1 (Fig. 9). The c_t optimal parameter has a different value with respect to the optimal range found in uncertainty analysis; this is probably related to the fact that the simulations show a complex inter-dependency between the parameter c_t and the parameter c_f both closely related to LST. This ends up increasing the equifinal parameter sets when only LST is used; on the other hand, c_t and c_f

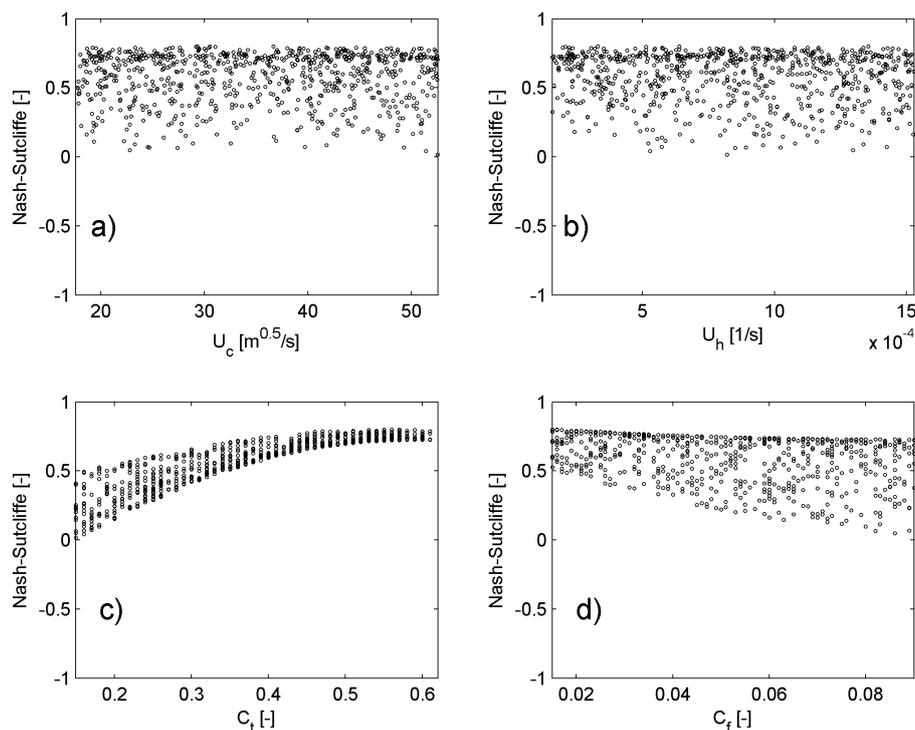


Figure 10. The Orba basin, August–October 2011. Dotty-plot representation of the NS on mean SWI at basin scale. Each subplot reports the parameter value on x axis and the value of NS statistic on y axis. The parameter c_t (c) has a quite clear range where NS reaches maximum values independently from the other parameters. This occurs even for parameter c_f although the independence is weaker (d). No independence is evident for the surface parameters (a, b).

influence in a complex way the different terms of the multi-objective function during the calibration process driving to the direction of a reduction of equifinality. When using the R. S. (LST) strategy, the benefit of exploiting LST data seems more related to the opportunity of doing a calibration in case of lack of streamflow data than in reducing equifinality as already noted in Silvestro et al. (2013).

Figure 11 reports values of the Nash–Sutcliffe coefficient (that depends only on the streamflow) versus values of the objective function of the M. O. approach. The parameter values that optimize the single score (S. N.) are not the same that optimize the M. O. function.

M. O. approach appears to be a good way to reduce equifinality. Looking at the dotty-plot representations in Sect. 3.1 built with streamflow, LST and SWI data, it is evident that even when graphs show an independence of a parameter from the others, this independence is not very pronounced. In other cases there is no evident independence. Combining the different objectives showed in the “Multi-objective calibration (M. O.)” section, we should eliminate those solutions (parameters combinations) that give good values for a certain metric (e.g., NS on streamflow) but not optimal values for another one (e.g., BIAS on LST), thus obtaining an overall better calibration.

Obviously, the choice of the single components of the M. O. function influences the way the various variables impact on the final results, but the applied methodology proposed by Madsen (2000) helps to normalize and balance the weights of the components.

In Table 6 the values of the scores for the validation period are reported, while Figs. 12 and 13 show the comparison between modeled and observed streamflow. The calibration period belongs to the validation period but it is considerably shorter. We thus decided to estimate the scores on the entire time window. In each figure some significant sub-periods are reported on small panels, while in the bottom panel the entire simulation period using a logarithmic scale is shown.

The values of the scores are good in all the cases. The Casalcermelli section performances are better than those for Tiglieto; this may be due to the fact that the first section corresponds to a larger drainage area and therefore the integration effects smooth the uncertainties of the rainfall fields. The M. O. approach leads to score values on the streamflow similar to the S. N. method in the validation period, while the R. S. approach produced poorer performance with respect to the other two approaches. In the case of PFRE, the two best sets of parameters are S. N. and R. S. (SWI). Notwithstanding, all the parameter sets led to good results in terms of the modeled hydrographs.

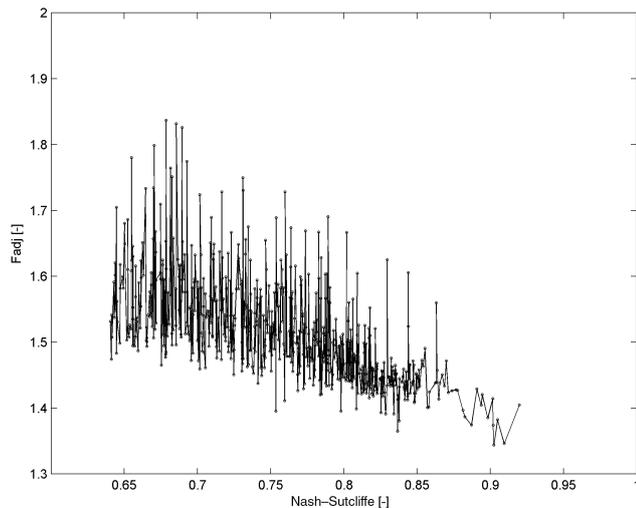


Figure 11. Nash–Sutcliffe coefficient of hydrographs (x axis) versus the multi-objective function (Fadj on y axis). The maximum of NS does not correspond with the minimum of the Fadj because this latter is influenced by the component that depends on LST and SWI. Nevertheless, in correspondence of the minimum value of Fadj (y axis) the NS value (x axis) is around 0.9, which means good performance in terms of NS score.

The M. O. calibration strategy leads to good performances in reproducing the observed streamflows despite the fact that these latter measurements are not the only ones used in the calibration process; there are good performances over long periods of simulation for both of the considered outlet sections. The peak flows and the time of the peak flows are generally well reproduced as well as the periods of flow recession and drought between the most relevant events. In general M. O. produces better results with respect to S. N. when it comes to correlation and CM; similarly, in the Tiglieto section where the uncertainty in rainfall input hampers the S. N. calibration strategy performance and the advantage of having more sources of observation is more evident.

The series of mean LST at basin scale were compared with LST satellite estimation. A similar comparison was made between the satellite SWI and modeled SD. Tables 7 and 8 show the values of the scores for the four considered parameter sets.

The classic calibration obtained with the maximization of the Nash–Sutcliffe coefficient of the streamflow (S. N.) allows good performances to be obtained in terms of streamflow simulation, but it produces higher values of LST BIAS, while the M. O. approach balances between the different components. The reproduction of SWI is quite good for both the M. O. and S. N. cases.

The accumulated discharge volume simulated by the model was compared with that derived from the streamflow observations in order to verify the behavior of the model in

Table 7. The Orba basin. Values of bias on LST series obtained with the different parameter sets. The basin scale mean has been considered.

Variable	Parameter set	BIAS
LST	M. O.	0.89
	R. S. (LST)	0.58
	R. S. (SWI)	0.92
	S. N.	1.02

Table 8. The Orba basin. Values of NS on SWI series obtained with the different parameter sets. The basin scale mean has been considered.

Variable	Parameter set	NS
SWI	M. O.	0.74
	R. S. (LST)	0.56
	R. S. (SWI)	0.80
	S. N.	0.77

Table 9. The Casentino basin parameters were calibrated following the different approaches: using the multi-objective function (M. O.), remote-sensing data and morphologic characteristics (R. S.), and the standard hydrographs comparison (S. N.). The table reports the values obtained by the four different calibration strategies.

Basin	Casentino			
	S. N.	R. S. (LST)	R. S. (SWI)	M. O.
u_c [$m^{0.5} s^{-1}$]	28.51	43.02	43.02	47.43
u_h [1/s]	0.00052	0.00047	0.00047	0.00043
c_t [–]	0.44	0.49	0.52	0.51
c_f [–]	0.018	0.032	0.018	0.029

terms of total runoff volumes. The results are reported in Fig. 14. The model reproduces with fine approximation the observed volumes; the error on the entire period is approximately -1.9 , -1.3 , -3.0 and -2.5 % for M. O., R. S. (LST), R. S. (SWI) and S. N., respectively. These errors are probably lower than the uncertainties introduced by the level-discharge transformation.

3.2.2 The Casentino basin

The sets of parameters obtained by the four calibration strategies are reported in Table 9.

In Table 10 the scores are reported, while Figs. 15 and 16 show the comparison between modeled and observed streamflow. The figures report on small panels different significant events, while in the bottom panel the entire simulation period is shown using a logarithmic scale.

Table 10. The Casentino basin skill scores on hydrographs on the entire validation period (2005–2011) for the different parameter sets. The scores value is calculated for the Subbiano section. PFRE is calculated with $Q_{th} = 200 \text{ m}^3 \text{ s}^{-1}$.

Basin	Section	Parameter set	NS	CM	RMSE	CORR	PFRE
Casentino	Subbiano	M. O.	0.80	0.77	2.37	0.89	-0.251
		R. S. (LST)	0.78	0.74	2.39	0.88	-0.311
		R. S. (SWI)	0.79	0.75	2.25	0.88	0.092
		S. N.	0.81	0.75	2.23	0.89	0.075

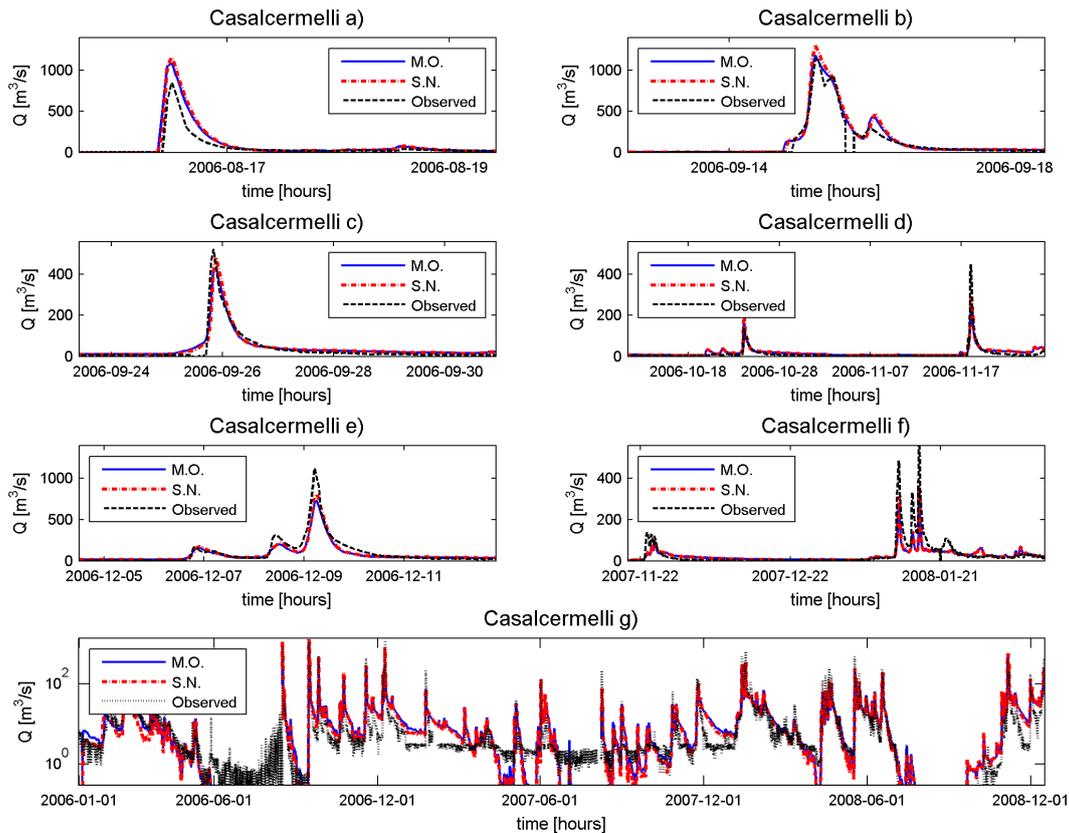


Figure 12. Observed streamflow for the Orba basin compared with simulations obtained using the M. O. and S. N. parameter sets for the period 2006 to 2008. Time is reported on x axis and streamflow on y axis. The bottom subpanel shows the entire period with log scale on y axis. The other subpanels show the main flood events with linear scale on y axis.

Table 11. The Casentino basin. Values of BIAS on LST series obtained with the different parameter sets. The mean basin scale has been considered.

Variable	Parameter set	BIAS
LST	M. O.	1.83
	R. S. (LST)	1.79
	R. S. (SWI)	1.87
	S. N.	1.88

Table 12. The Casentino basin. Values of NS on SWI series obtained with the different parameter sets. The mean basin scale has been considered.

Variable	Parameter set	NS
SWI	M. O.	0.80
	R. S. (LST)	0.63
	R. S. (SWI)	0.83
	S. N.	0.73

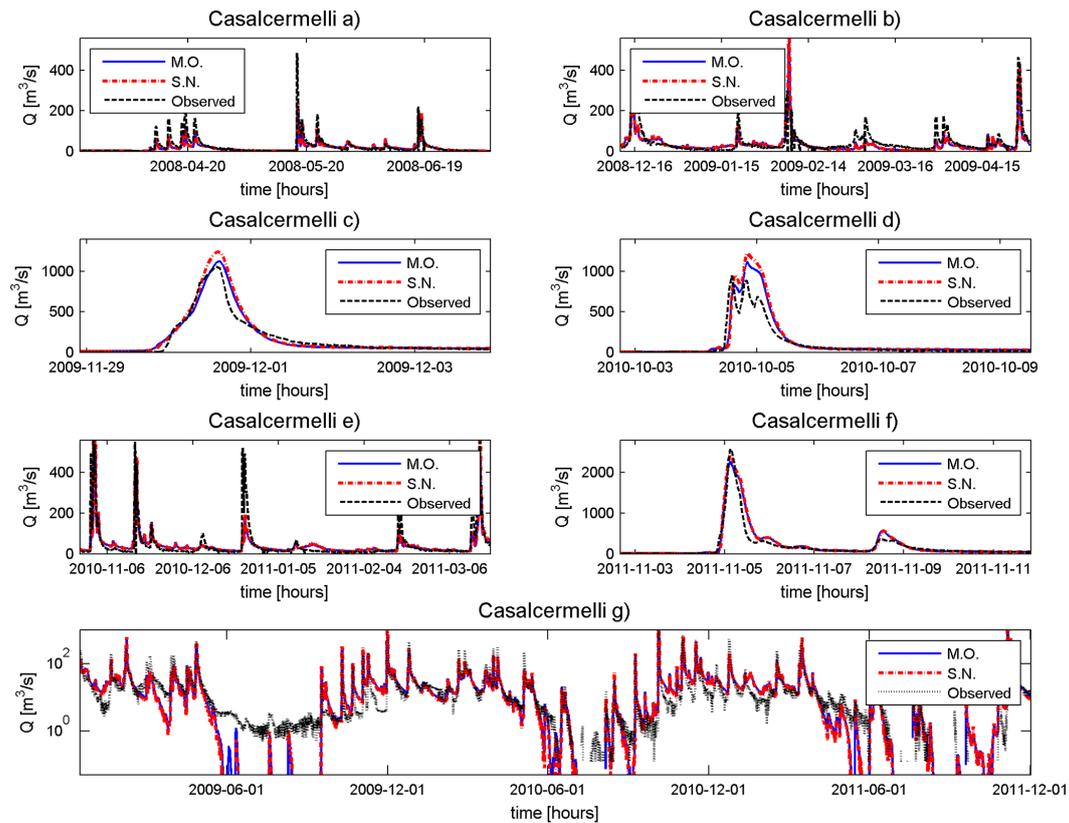


Figure 13. Observed streamflow for the Orba basin compared with simulations obtained using the M. O. and S. N. parameter sets for the period 2009 to 2011. Time is reported on x axis and streamflow on y axis. The bottom subpanel shows the entire period with log scale on y axis. The other subpanels show the main flood events with linear scale on y axis.

Tables 11 and 12 show the values of the scores for the variables LST and SWI. The values of the scores are good in all the cases; they are better for M. O. with respect to S. N. for both LST and SWI variables.

The model delivers good performances over long periods for all the considered calibration strategies. As in the case of the Orba basin, M. O. finds a compromise parameter set that allows a good modeling of all the three variables involved to be obtained in the objective functions, while the two R. S.: method shows that is possible to obtain a reasonable calibration even in the case of an ungauged basins. The peak flows are better reproduced by the configurations S. N. and R. S. (SWI) which show PFRE values smaller than 0.1 (10%).

The accumulated volume over the 7 years of simulation is generally well simulated (Fig. 17); in this case, there is a larger difference between the total volumes obtained with the four different parameter sets, the errors are in fact of the order of 9, 8.8, 5.3 and 4 % for M. O., R. S. (LST), R. S. (SWI) and N. S., respectively. The errors on the total volume are larger than in the case of the Orba basin.

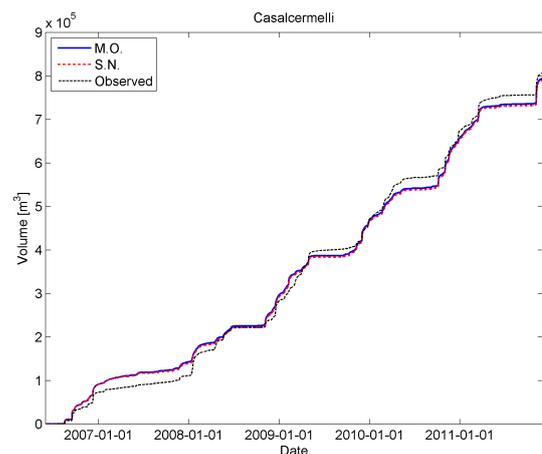


Figure 14. Accumulated runoff volumes for the Orba basin obtained with the M. O. and S. N. parameter sets for the period 2006–2011. Time is reported on x axis and total accumulated volume on y axis.

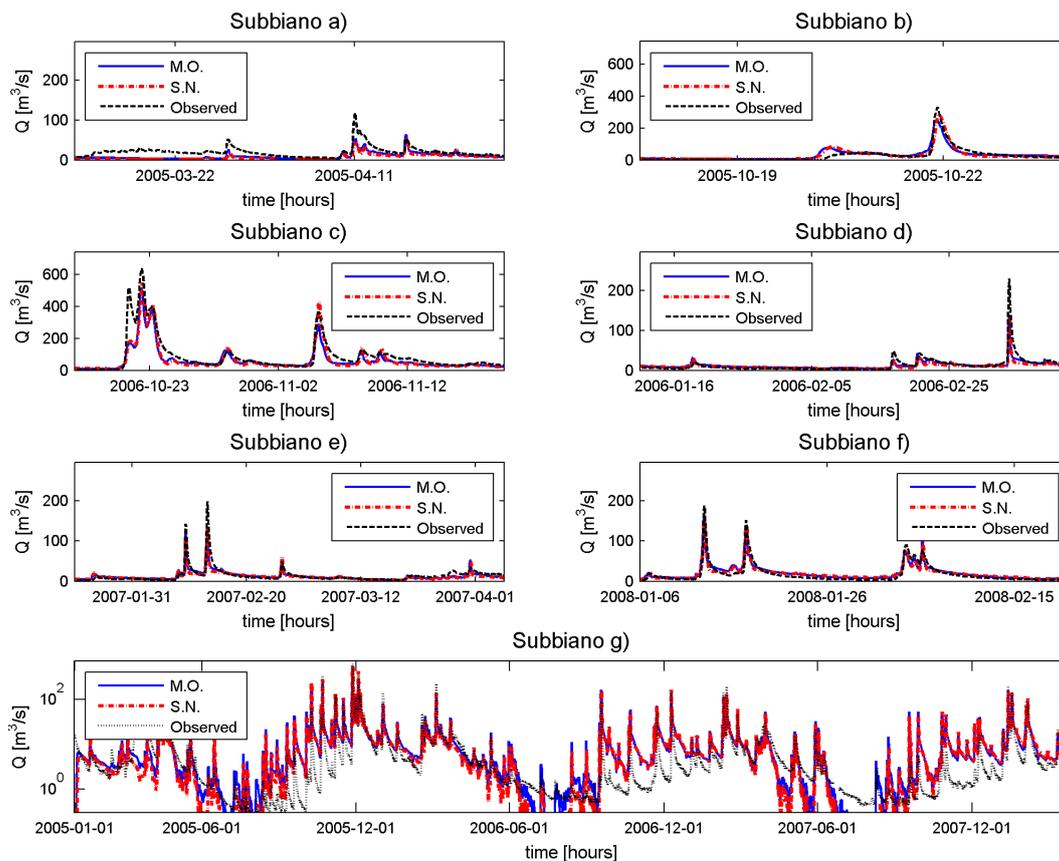


Figure 15. Observed streamflow for the Casentino basin compared with simulations obtained using the M. O. and S. N. parameter sets for the period 2005 to 2007. Time is reported on x axis and streamflow on y axis. The bottom subpanel shows the entire period with log scale on y axis. The other subpanels show the main flood events with linear scale on y axis.

4 Discussion and conclusions

This paper shows that satellite data are useful in reducing the uncertainty of the parameterization of a distributed hydrological model and that they can be used in calibration strategy to improve model representation of hydrological processes.

The uncertainty analysis (Zappa et al., 2011; Shen et al., 2012) of the most sensitive parameters shows that the equifinality can be reduced using LST and SWI satellite estimations. The independence of surface flow parameters seem to be linked in a clear way to streamflow, while soil parameters are more directly linked to LST and SWI.

Three methodologies to estimate a subset of the parameters of the model by exploiting remote sensing were applied. The first methodology consists in the minimization of a multi-objective function that depends on streamflow, LST and SWI and it is inspired by the work of Sutanudjaja et al. (2014) and Wanders et al. (2014). The second and third methodologies simulate the case when no streamflow data are available and the calibration is carried out based only on LST and SWI retrieved from satellite data and information derived from a DEM. A multi-year period validation was

done in terms of reproduction of both streamflow time series and total volume over the considered period. A comparison with a fourth standard calibration strategy based on streamflow data was also carried out.

The skill scores on streamflow show good performances when satellite data are involved in the calibration process (M. O. and R. S. methods), comparable with values obtained using only the streamflow in the objective function (S. N. method); even if the observed and simulated streamflow are in some cases quite different, the general trend is good and there are not large biases in terms of runoff volumes over long simulation periods. The largest errors seem to be more related to the uncertainties of the input rainfall fields rather than on the model parameterization. Moreover, the skill scores on LST and SWI have generally better values in the case of parameter sets derived by the multi-objective approach with respect to those obtained by the streamflow-based calibration strategy.

Both the results of calibration, especially the M. O. approach, and uncertainty analysis confirm that a way to reduce equifinality and to augment the parameter constraints is related to the increase of model state variables and model

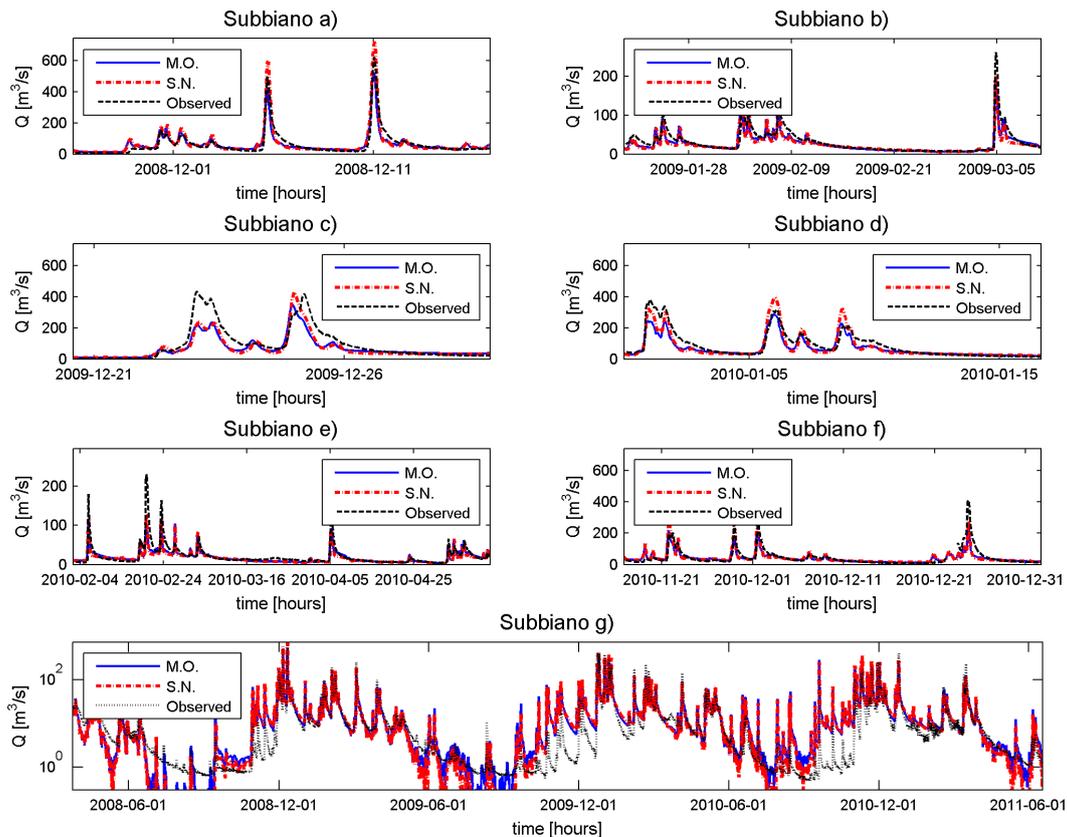


Figure 16. Observed streamflow for the Casentino basin compared with simulations obtained using M. O. and S. N. parameter sets for the period 2008 to 2011. The bottom subpanel shows the entire period with log scale on y axis. Time is reported on x axis and streamflow on y axis. The other subpanels show the main flood events with linear scale on y axis.

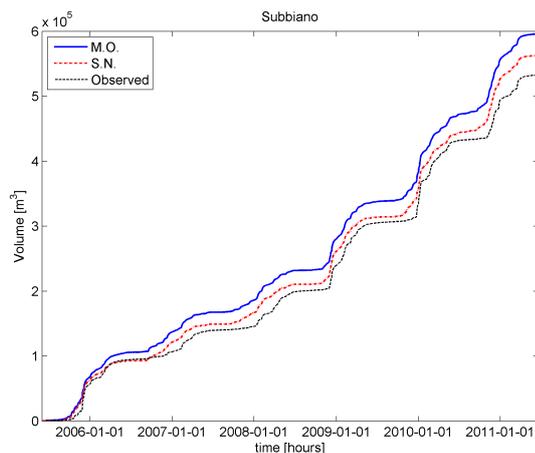


Figure 17. Accumulated runoff volumes of the Casentino basin obtained with the M. O. and S. N. parameter sets for the period 2005–2011. Time is reported on x axis and total accumulated volume on y axis.

output variables that can be derived from both gauge and remote-sensing data. This helps to reduce the possibility of obtaining similar results with a large number of parameter sets. We can thus state that the presented work explored the direction proposed by Seibert and McDonnell (2002) and Efstratiadis and Koutsoyiannis (2010), which consisted in obtaining a better overall performance of the model and ensure consistency across its various aspects.

In addition, remote-sensing data (in this specific case the LST and SSM) offer alternative ways to carry out parameter calibration in cases where no streamflow data might be available. Satellite-derived data such as DEM, SSM and LST are generally universally available.

The described methodologies can be adapted and applied to other hydrological models that have characteristics similar to Continuum and that can simulate LST and soil moisture as state variables in a prognostic way; moreover, these methods can be extended by referring to other remote-sensing data, and general observed data, that can be reproduced by the model. The more the model has the capability of reproducing observable quantities (e.g., evapotranspiration, soil humidity), the more the constraints imposed on the model can increase.

Finally, the results of the presented work can be read from two different points of view. On one hand, the results highlight the advantages of using distributed hydrological models that allow for reproduction, with some degree of detail, of the physical processes; as a matter of fact, such models simulate a larger number of variables which can also be observed. On the other hand, similar to what was demonstrated by other authors (Corbari and Mancini, 2014; Montzka et al., 2011; Sutanudjaja et al., 2014; Wanders et al., 2014), the results highlight the opportunities given by remote sensing and the necessity of augmenting the number (and the quality) of these data. Remotely sensed data can in fact be used to parameterize hydrological models and to set up constraints to the parameters in the calibration process, while offering an alternative way of calibrating these models where standard observation are lacking.

Appendix A

The snow accumulation–melting module was introduced in order to carry out multi-year simulations in alpine climates. It is a simple model that is derived from commonly used equations (Maidment, 1992) and it is forced by meteorological observations.

The equations that describe the snow mass conservation and its melting are the following:

$$\frac{\Delta \text{SWE}}{\Delta t} = S_f - \text{SM}, \quad (\text{A1})$$

where SWE is the snow water equivalent, S_f is the solid precipitation and SM is the snowmelt estimated as

$$\text{SM} = \frac{R_n}{\rho_w \lambda_f} + m_c \cdot (T_a - T_0), \quad (\text{A2})$$

where R_n is the net radiation, ρ_w the water density, λ_f the latent heat of melting and T_a the air temperature. T_0 and m_c are two parameters that represent the temperature at which the melting starts and the melting coefficient, respectively. These two parameters are estimated using values from the literature (Maidment, 1992), $T_0 = 0^\circ\text{C}$ and $m_c = 4 \text{ mm day}^{-1}$.

The mass balance is applied at cell scale for the entire domain of the model, so that a snow cover map can be generated with the same resolution of the DEM. The energy balance and, as a consequence, the evapotranspiration are inhibited for those cells where snow cover is present.

The applied approach is very simple and neglects the heat exchanges between the soil and the snow cover, but it is generally sufficient if the goal is the estimation of the snow contribution to the runoff, especially when the regime of the basin is not strongly influenced by snowmelt.

The precipitation is partitioned into solid or liquid if the air temperature is below or above a fixed threshold.

Acknowledgements. This work is supported by the Italian Civil Protection Department, and by the Italian regions of Valle d'Aosta and Liguria. We acknowledge the Italian Civil Protection Department for providing us with the data from the regional meteorological observation networks. We also thank the H-SAF project for the availability of surface soil moisture data.

Edited by: A. Weerts

References

- Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Pignatelli, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrol. Earth Syst. Sci.*, 12, 1323–1337, doi:10.5194/hess-12-1323-2008, 2008.
- Albergel, C., de Rosnay, P., Gruhierand, C., Munoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, *Remote Sens. Environ.*, 118, 215–226, 2012.
- Bartholmes and Todini: Coupling meteorological and hydrological models for flood forecasting, *Hydrol. Earth Syst. Sci.*, 9, 333–346, doi:10.5194/hess-9-333-2005, 2005.
- Beven, K. and Binley, A. M.: The future of distributed models: model calibration and uncertainty prediction, *Hydrol. Process.*, 24, 43–69, 1992.
- Bitew, M. M. and Gebremichael, M.: Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model, *Water Resour. Res.*, 47, W06526, doi:10.1029/2010WR009917, 2011.
- Bjerklie, D. M., Dingman, S. L., Vorosmarty, C. J., Bolster, C. H., and Congalton, R. G.: Evaluating the potential for measuring river discharge from space, *J. Hydrol.*, 278, 17–38, 2003.
- Brakenridge, G. R., Nghiem, S. V., Anderson, E. A., and Chien, S.: Space-based Measurement of River Runoff, *EOS Trans. AGU*, 86, 185, 2005.
- Brakenridge, G. R., Cohen, S., Kettner, A. J., De Groeve, T., Nghiem, S. V., Syvitski, J. P., and Fekete, B. M.: Calibration of satellite measurements of river discharge using a global hydrology model, *J. Hydrol.*, 475, 123–136, 2012.
- Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Hasenauer, S.: ASCAT soil wetness index validation through in situ and modelled soil moisture data in central Italy, *Remote Sens. Environ.*, 114, 2745–2755, doi:10.1016/j.rse.2010.06.009, 2010.
- Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, *Remote Sens. Environ.*, 115, 3390–3408, 2011a.
- Brocca, L., Melone, F., and Moramarco, T.: Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting, *Hydrol. Process.*, 25, 2801–2813, 2011b.
- Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Albergel, C.: Scaling and filtering approaches for the use of satellite soil moisture observations “Remote Sensing of Land Surface Turbulent Fluxes and Soil Surface moisture Content: State of the Art”, Chapter 17, Taylor & Francis Group, London, UK, 415–430, 2013.
- Caparrini, F., Castelli, F., and Entekhabi, D.: Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences, *J. Hydrometeorol.*, 5, 145–159, 2004.
- Carpenter, T. M. and Georgakakos, K. P.: Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales, *J. Hydrol.*, 329, 174–185, 2006.
- Chiew, F. and McMahon, T.: Application of the daily rainfall–runoff model MODHYDROLOG to 28 Australian catchments, *J. Hydrol.*, 153, 383–416, 1994.
- Coccia, G., Mazzetti, C., Ortiz, E. A., and Todini, E.: Application of the TOPKAPI model within the DMIP 2 project, Proceedings of the International Conference, 23rd Conference on Hydrology, Phoenix, Arizona, 2009.
- Corbari, C. and Mancini, M.: Calibration and Validation of a Distributed Energy–Water Balance Model Using Satellite Data of Land Surface Temperature and Ground Discharge Measurements, *J. Hydrometeorol.*, 15, 376–392, doi:10.1175/JHM-D-12-0173.1, 2014.
- Crow, W. T., Wood, E. F., and Pan, M.: Multiobjective calibration of land surface model evapotranspiration predictions using streamflow observations and spaceborne surface radiometric temperature retrievals, *J. Geophys. Res.-Atmos.*, 108, 4725, doi:10.1029/2002JD003292, 2003.
- Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., de Rosnay, P., Ryu, D., and Walker, J. P.: Upscaling sparse ground based soil moisture observations for the validation of coarse resolution satellite soil moisture products, *Rev. Geophys.*, 50, RG2002, doi:10.1029/2011RG000372, 2012.
- Dickinson, R.: The force-restore method for surface temperature and its generalization, *J. Climate*, 1, 1086–1097, 1988.
- Efstratiadis, A. and Koustoyiannis, D.: One decade of Multi-Objective calibration approaches in hydrological modelling: a review, *Hydrolog. Sci. J.*, 55, 58–78, 2010.
- EUMETSAT: Product user manual, vegetation parameters, <http://landsaf.meteo.pt/> (last access: 8 January 2012), 2008.
- EUMETSAT: Algorithm Theoretical Basis Document for Land Surface Temperature (LST), <http://landsaf.meteo.pt/> (last access: 8 January 2012), 2009.
- Flores, A. N., Entekhabi, D., and Bras, R. L.: Reproducibility of soil moisture ensembles when representing soil parameter uncertainty using a Latin Hypercube–based approach with correlation control, *Water Resour. Res.*, 46, W04506, doi:10.1029/2009WR008155, 2010.
- Gabellani, S., Silvestro, F., Rudari, R., and Boni, G.: General calibration methodology for a combined Horton-SCS infiltration scheme in flash flood modeling, *Nat. Hazards Earth Syst. Sci.*, 8, 1317–1327, doi:10.5194/nhess-8-1317-2008, 2008.
- Giannoni, F., Roth, G., and Rudari, R.: A Semi-Distributed Rainfall–Runoff Model Based on a Geomorphologic Approach, *Phys. Chem. Earth*, 25, 665–671, 2000.
- Giannoni, F., Roth, G., and Rudari, R.: A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response, *Adv. Water Resour.*, 28, 567–581, 2005.

- Göttsche, F. M. and Hulley, G. C.: Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region, *Remote Sens. Environ.*, 124, 149–158, 2012.
- Khan, S. I., Hong, Y., Wang, J., Yilmaz, K. K., Gourley, J. J., Adler, R. F., Brakenridge, G. R., Policelli, F., Habib, S., and Irwin, D.: Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria Basin: Implications for hydrologic prediction in ungauged basins, *IEEE T. Geoscience Remote*, 49, 85–95, 2011.
- Kim, S. M., Benham, B. L., Brannan, K. M., Zeckoski, R. W., and Doherty, J.: Comparison of hydrologic calibration of HSPF using automatic and manual methods, *Water Resour. Res.*, 43, W01402, doi:10.1029/2006WR004883, 2007.
- Koren, V., Moreda, F., and Smith, M.: Use of soil moisture observations to improve parameter consistency, *J. Phys. Chem. Earth*, 33, 1068–1080, doi:10.1016/j.pce.2008.01.003, 2008.
- Kozak, J. A., Ahuja, L. R., Green, T. R., and Ma, L.: Modelling crop canopy and residue rainfall interception effects on soil hydrological components for semi-arid agriculture, *Hydrol. Process.*, 21, 229–241, 2007.
- Lin, J. D.: On the Force-Restore method for prediction of ground surface temperature, *J. Geophys. Res.*, 85, 3251–3254, 1980.
- Liu, Z., Martina, M. L. V., and Todini, E.: Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment, *Hydrol. Earth Syst. Sci.*, 9, 347–364, doi:10.5194/hess-9-347-2005, 2005.
- Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, *J. Hydrol.*, 235, 276–288, 2000.
- Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Adv. Water Resour.*, 26, 205–216, 2003.
- Maidment, D.: *Handbook of Hydrology*, McGraw-Hill, Inc., New York, San Francis among others, USA, 1992.
- Montaldo, N., Rondena, R., Albertson, J. D., and Mancini, M.: Parsimonious modelling of vegetation dynamics for ecohydrological studies of waterlimited ecosystems, *Water Resour. Res.*, 41, W10416, doi:10.1029/2005WR004094, 2005.
- Montzka, C., Moradkhani, H., Weihermuller, L., Franssen, H.-J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, 399, 410–421, doi:10.1016/j.jhydrol.2011.01.020, 2011.
- Murray, S. J., Watson, I. M., and Prentice, I. C.: The use of dynamic global vegetation models for simulating hydrology and the potential integration of satellite observations, *Prog. Phys. Geogr.*, 37, 63–97, doi:10.1177/0309133312460072, 2013.
- Nash, J. E. and Sutcliffe, J. V.: River flood forecasting through conceptual models I: a discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., and Scipal, K.: Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale, *Hydrol. Earth Syst. Sci.*, 10, 353–368, doi:10.5194/hess-10-353-2006, 2006.
- Pavelsky, T. M.: Using width based rating curves from spatially discontinuous satellite imagery to monitor river discharge, *Hydrol. Process.*, 28, 3035–3040, 2014.
- Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242, 275–301, 2001.
- Rakovec, O., Hill, M. C., Clark, M. P., Weerts, A. H., Teuling, A. J., and Uijlenhoet, R.: Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models, *Water Resour. Res.*, 50, 409–426, doi:10.1002/2013WR014063, 2014.
- Ridler, M. E., Sandholt, I., Butts, M., Lerer, S., Mouglin, E., Timouk, F., Kergoat, L., and Madsen, E.: Calibrating a soil–vegetation–atmosphere transfer model with remote sensing estimates of surface temperature and soil surface moisture in a semi arid environment, *J. Hydrol.*, 436–437, 1–12, 2012.
- Rhoads, J., Dubayah, R., Lettenmaier, D., O'Donnell, G., and Lakshmi, V.: Validation of land surface models using satellite derived surface temperature, *J. Geophys. Res.*, 106, 20085–20099, 2001.
- Santanello, J., Peters-Lidard, C. D., Garcia, M. E., Mocko, D. M., Tischler, M. A., Moran, M. S., and Thoma, D.: Using remotely-sensed estimates of soil moisture to infer soil texture and hydraulic properties across a semi-arid watershed, *Remote Sens. Environ.*, 110, 79–97, doi:10.1016/j.rse.2007.02.007, 2007.
- Saulnier, G. M., Beven, K., and Obled, C.: Including spatially variable effective soil depths in TOPMODEL, *J. Hydrol.*, 202, 158–172, 1997.
- Savenije, H. H. G.: Equifinality, a blessing in disguise?, *Hydrol. Process.*, 15, 2835–2838, 2001.
- Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, 38, 23-1–23-14, doi:10.1029/2001WR000978, 2002.
- Shen, Z. Y., Chen, L., and Chen, T.: Analysis of parameter uncertainty in hydrological and sediment modeling using GLUE method: a case study of SWAT model applied to Three Gorges Reservoir Region, China, *Hydrol. Earth Syst. Sci.*, 16, 121–132, doi:10.5194/hess-16-121-2012, 2012.
- Silvestro, F., Gabellani, S., Delogu, F., Rudari, R., and Boni, G.: Exploiting remote sensing land surface temperature in distributed hydrological modelling: the example of the Continuum model, *Hydrol. Earth Syst. Sci.*, 17, 39–62, doi:10.5194/hess-17-39-2013, 2013.
- Singh, S. K. and Bárdossy, A.: Calibration of hydrological models on hydrologically unusual events, *Adv. Water Resour.*, 38, 81–91, doi:10.1016/j.advwatres.2011.12.006, 2012.
- Sini, F., Boni, G., Caparrini, F., and Entekhabi, D.: Estimation of large scale evaporation fields based on assimilation of remotely sensed land temperature, *Water Resour. Res.*, 44, W06410, doi:10.1029/2006WR005574, 2008.
- Stroud, P. D.: A recursive exponential filter for time-sensitive data, Los Alamos national Laboratory, LAUR-99-5573, available at: <http://public.lanl.gov/stroud/ExpFilter/ExpFilter995573.pdf> (last access: July 2008), 1999.
- Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, *Water Resour. Res.*, 50, 687–705, doi:10.1002/2013WR013807, 2014.
- Todini, E.: History and perspective of hydrological catchment modelling, in: *Water, Environment, Energy and Society*, edited by: Jain, S. K., Singh, V. P., Kumar, V., Kumar, R., Singh, R. D., and

- Sharma, K. D., Proc. of WEES-2009, Allied Publishers Pvt. Ltd, New Delhi, India, 512–523, 2009.
- Todini, E. and Ciarapica, L.: The TOPKAPI Model. Mathematical Models of Large Watershed Hydrology, in: Chapter 12, edited by: Singh, V. P. and Frevert, D. K., Water Resources Publications, Littleton, Colorado, 2001.
- Vörösmarty, C., Askew, A., Grabs, W., Barry, R. G., Birkett, C., Döll, P., Goodison, B., Hall, A., Jenne, R., Kitaev, L., Landwehr, J., Keeler, M., Leavesley, G., Schaake, J., Strzepek, K., Sundarvel, S. S., Takeuch, K., and Webster, F.: Global water data: A newly endangered species, *Eos Trans. Am. Geophys. Un.*, 82, 54–58, 2001.
- Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39, 1201, doi:10.1029/2002WR001642, 2003.
- Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Environ.*, 70, 191–207, 1999.
- Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Zuger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Bloeschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., and Rubel, F.: The ASCAT soil moisture product: specifications, validation results, and emerging applications, *Meteorol. Z.*, 22, 5–33, doi:10.1127/0941-2948/2013/0399, 2013.
- Wanders, N., Bierkens, M. F. P., Jong, S. M., Roo, A., and Karssen-berg, D.: The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models, *Water Resour. Res.*, 50, 6874–6891, doi:10.1002/2013WR014639, 2014.
- White, C. D. and Lewis, M.: A new approach to monitoring spatial distribution and dynamics of wetlands and associated flows of Australian Great Artesian Basin springs using QuickBird satellite imagery, *J. Hydrol.*, 408, 140–152, 2011.
- Wooding, R. A.: A hydraulic modeling of the catchment-stream problem, 1. Kinematic wave theory, *J. Hydrol.*, 3, 254–267, 1965.
- Xu, D., Wang, W., Chau, K., Cheng, C., and Chen, S.: Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters, *J. Hydroinform.*, 15.1, 174–193, 2013.
- Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-Objective global optimization for hydrological models, *J. Hydrol.*, 204, 83–97, 1998.
- Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M. Y., Vinnikov, K. Y., Sun, D., and Tian, Y.: Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties, *IEEE T. Geosci. Remote*, 50, 704–713, 2012.
- Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F.: Superposition of three sources of uncertainties in operational flood forecasting chains, *Atmos. Res.*, 100, 246–262, doi:10.1016/j.atmosres.2010.12.005, 2011.
- Zhang, T., Stackhouse Jr., P. W., Gupta, S. K., Cox, S. J., Colleen Mikovitz, J., and Hinkelman, L. M.: The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach, *J. Quant. Spectrosc. Ra.*, 122, 127–140, 2013.