



# Inverse modeling of soil water content to estimate the hydraulic properties of a shallow soil and the associated weathered bedrock

O. Le Bourgeois, Christophe Bouvier, P. Brunet, P. -A. Ayral

## ► To cite this version:

O. Le Bourgeois, Christophe Bouvier, P. Brunet, P. -A. Ayral. Inverse modeling of soil water content to estimate the hydraulic properties of a shallow soil and the associated weathered bedrock. *Journal of Hydrology*, 2016, Flash floods, hydro-geomorphic response and risk management, 541, pp.116–126. 10.1016/j.jhydrol.2016.01.067 . hal-02086691

**HAL Id: hal-02086691**

**<https://hal.science/hal-02086691>**

Submitted on 24 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Inverse modeling of soil water content to estimate the hydraulic properties of a shallow soil and the associated weathered bedrock

O. Le Bourgeois<sup>a,\*</sup>, C. Bouvier<sup>a</sup>, P. Brunet<sup>a</sup>, P.-A. Ayrat<sup>b</sup>

<sup>a</sup> Hydrosclences Montpellier, UMR 5569 CNRS-IRD-UM, 300, avenue du Professeur Emile Jeanbrau, 34000 Montpellier, France

<sup>b</sup> Ecole des Mines d'Alès, Centre de Recherche LGEI, 7, rue Jules Renard, 30319 Alès, France

## S U M M A R Y

Modeling soil water flow requires the knowledge of numerous parameters associated to the water content and the soil hydraulic properties. Direct estimations of those parameters in laboratory require expensive equipment and the obtained parameters are generally not representative at the field scale because of the limitation of core sample size. Indirect methods such as inverse modeling are known to get efficient estimations and are easier to set up and process for large-scale studies. In this study, we investigated the capacity of an inverse modeling procedure to estimate the soil and the bedrock hydrodynamic properties only from *in situ* soil water content measurements at multiple depths under natural conditions. Multi-objective parameter optimization was performed using the HYDRUS-1D software and an external optimization procedure based on the NSGA-II algorithm. In a midslope shallow soil, water content was monitored at 3 depths, 20, 40, and 60 cm during 12 intense rainfall events, whose amounts ranged between 50 and 250 mm and duration between 1 and 5 days. The vertical profile was considered as 2 layers of soils above a third layer representing the weathered schist rock. This deep layer acted as a deep boundary condition, which features the bedrock permeability and water storage. Each layer was described through the 6 parameters of the Mualem–van Genuchten formulation. The calibrated parameters appeared to have very low uncertainty while allowing a good modelisation of the observed water content variations. The calibrated saturated water content was close to the laboratory porosity measurements while the saturated hydraulic conductivity showed that the soil was highly permeable, as measured in the field. The inverse modeling approach allowed an estimation of the hydraulic properties of the bedrock layer where no measurement was available. The bedrock layer was found to have a low saturated hydraulic conductivity ( $<5 \text{ mm h}^{-1}$ ), which means that the schist bedrock is poorly weathered and that saturated area can be generated above this depth, as it was observed. The simulated water contents were generally close to the measured water contents, but the model failed sometimes to reproduce the saturation of the soil in the deeper layers, probably because of sub-surface flux at the soil/bedrock interface. In these cases, further investigation will have to be made by using a 2D-model.

## Keywords:

Inverse modeling

Soil moisture

Soil hydraulic properties

Bedrock hydraulic properties

## 0. Introduction

Soil water content and flow are crucial to understand small-scale processes involved in many hydrological applications, e.g. flash-flood genesis occurring on small watersheds (Braud et al., 2010). Hydrological models used to simulate soil water flows are generally based on Richards' equation (Richards, 1931) involving hydraulic conductivity, water pressure head and soil water content, as well as boundary conditions. The main difficulty in using such model lies in the access to the soil hydraulic properties, i.e.

the parameters that governs both the soil water retention and the soil water conductivity curves. Moreover, recent studies showed the importance of investigating the bedrock conductivity function since the water flow at the interface between soil and bedrock is one of the key processes involved in flood genesis (Hopp and McDonnell, 2009; James et al., 2010).

Soil hydraulic parameters can be estimated by many direct or indirect methods. Direct methods are based on small soil sample laboratory experiments or small scale field experiments. Laboratory methods allow a total control of boundary conditions and precise data acquisition (Durner and Lipsius, 2005) but according to Dirksen (1999), the presence of large unstable structural elements such as discontinuous granitic slabs or abundance of stones are overriding reasons to perform *in situ* experiments. Likewise, direct

\* Corresponding author. Tel.: +33 (0) 4 67 14 90 44.

E-mail address: [olivier.le-bourgeois@hotmail.fr](mailto:olivier.le-bourgeois@hotmail.fr) (O. Le Bourgeois).

field experiment methods involve high cost, significant time demand and are very sensitive to small scale heterogeneity (Tseng and Jury, 1993; Basile et al., 2003).

On the other hand, automatic-calibration (inverse modeling) aims to find the hydraulic parameters by minimizing the difference between simulated and observed data such as water content, head pressure or geophysical prospecting (for a review of inverse methods, see Vrugt et al., 2008). It is one of nowadays most used methods thanks to its ability to give good estimates of hydraulics properties from relatively simple data (Ritter et al., 2003; Loew and Mauser, 2008; Wollschläger et al., 2009; Scharnagl et al., 2011; Schelle et al., 2013; Werisch et al., 2014). However, much attention must be paid in the following aspects:

- (i) *Non-optimal solution*: inverse modeling is based on the minimization of an objective function, which can be performed by using several algorithms and methods. Many studies compared the search algorithms (Madsen, 2003; Wöhling et al., 2008; Efstratiadis and Koutsoyiannis, 2010) and concluded that a global optimization approach was more effective than local search optimization because it reduces the risk of being trapped in a local optimum during the optimization process. By definition, the multiple objective functions approach also allows a more complex constraint on the model which is preferred when the number of estimated parameters becomes important (Efstratiadis and Koutsoyiannis, 2010).
- (ii) *Non-uniqueness solution*: dependency between hydraulic parameters leads to the existence of multiple set of parameters, which give equally satisfying simulations. This problem is known as equifinality (Beven, 1993). It is one of the major difficulties for the interpretation of parameters values, their comparison in time or in space, as well as the extrapolation of the model beyond the observed conditions. Multi-objective methods were proved to be satisfactory to reduce the non-uniqueness problem (Madsen, 2003; Bekele and Nicklow, 2007; Vrugt and Robinson, 2007) because they provide the trade-off of all the objective function used leading to Pareto-optimal set of solutions (Horn et al., 1994; Wöhling et al., 2008; Efstratiadis and Koutsoyiannis, 2010). The compromise solution representing the best fit to multiple aspects of the problem leads to the most behavioral hydraulic properties.

In Mediterranean mountainous catchments, thin but pervious soils and high rainfall intensities cause the water storage at the hillslope scale to be largely influenced by the bedrock permeability. Thus, different responses can be induced by the geological features of the bedrock, for example shale or granites (Vannier et al., 2013).

In addition, temporal variations of water content are the result of combined complex processes such as infiltration and evapotranspiration, among others. An event-based inverse modeling approach can reduce the processes to be taken into account in the temporal variation of the water content, by, for example, neglecting the evapotranspiration processes during the short rain events. Such method was proved to be efficient for rainfall-runoff modeling (Berthet et al., 2009; Tramblay et al., 2010).

This study aims to understand the influence of the weathered bedrock on the vertical water flow during high intensity rainfall events which lead to flash floods in the Cévennes region (Southeast of France). To do so, we need to determine the hydraulic properties of a mountainous sandy-loam soil and the schist substratum under it. Beyond the study case presented here, it is expected to get a method in order to explore furthermore a wide area still poorly documented, with respect to the hydrodynamic properties of the

soils and their spatial variability. Inverse modeling is based on a multi-event based approach using the NSGA-II (Deb et al., 2002) multi-objective algorithm in order to avoid accounting for the evapotranspiration processes and to reduce the equifinality problem. In addition, special attention is paid to the estimation of the bedrock hydraulic properties and its influence on soil saturation and runoff.

## 1. Material and methods

### 1.1. Study area

The study area is a small plot about 420 m<sup>2</sup> located in Sumène (Fig. 1), 50 km north of Montpellier in the Cévennes mountains (43°58'55"N, 3°42'58"E). The plot presents a medium slope of 15% west oriented.

#### 1.1.1. Soil properties

The soil is a thin Brunisol over schist bedrock. As represented in Fig. 2 (left), the five first centimeters consist in a C-org rich layer, followed by a thirty centimeters thick first mineral layer which contains lots of coarse fragments and roots. Finally, the lowest mineral layer is in contact with the bedrock. Soil thickness ranges from 60 to 90 cm down to the top of the schist bedrock across the plot. To estimate granulometry, 9 soil samples have been taken from 3 locations over the experimental plot. In every location, 3 samples have been taken at increasing depths: 5, 30, and 50 cm. The mean percentage of sand, silt, clay, and coarse fragments is shown in Fig. 2 (center). The soil is characterized by 30–37% of coarse fragments, and for the matrix component, 60–73% of sand, 20–27% of silt, and only 7–13% of clay.

Using 100 cm<sup>3</sup> steel cylinders, 7 soil samples have been taken from surface to 85 cm every 10 cm deep. The density ranges from 1.1 (in the first structural layer) to 1.5 g cm<sup>-3</sup> due to the increase of coarse fragments, while porosity decreases from 0.50 to 0.40 cm<sup>3</sup> cm<sup>-3</sup> (Fig. 2, right).

In order to get a direct estimation of the soil conductivity, we made 3 set of measurements using a tension disc infiltrometer (SDEC SW080B model). Two out of the 3 set of measurements were done at the soil surface, and the last one at 20 cm depth. The three near-saturated hydraulic conductivity estimations were 204, 208, and 219 mm h<sup>-1</sup> with respectively last measured potentials −5, −3, and −1 mm.

#### 1.1.2. Soil water content data

Soil water content was monitored by 3 Theta-probe ML2-X sensors installed at 3 depths: 20, 40, and 60 cm. In the neighboring of the probes, the soil depth is around 70 cm. Probes were spaced approximately 50 cm apart sideways to limit the vertical water flow disruption. Volumetric soil moisture ( $\theta$ ) values was derived from the soil dielectric capacity using the following equation given by the constructor (Delta-T devices Ltd, 1999):

$$\theta = \frac{[1.1 + 4.44V] - a_0}{a_1} \quad (1)$$

with  $\theta$  the volumetric water content (cm<sup>3</sup> cm<sup>-3</sup>),  $V$  the output voltage of the probe (V),  $a_0$  and  $a_1$  two calibrated constant (–).

The probes are calibrated by the constructor and generally need to be adjusted. But, thanks to the large majority of mineral particles in the studied soil, the calibrated parameters obtained oncore samples were very close to the original calibration so we choose to directly use the constructor calibration for mineral soil.

Water contents were monitored at a 15 min time step during three years from 2008 to 2010. The example of year 2008 exhibits a strong contrast between very dry soil during hot summer and

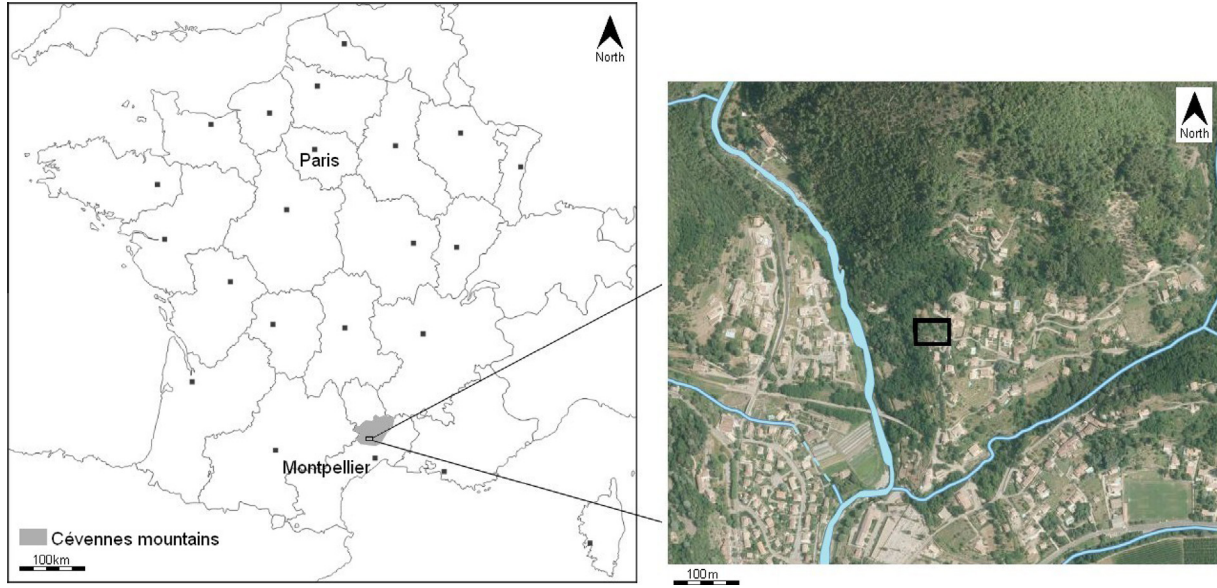


Fig. 1. Localization of the Sumène experimental land-plot in the Cévennes region, south of France.

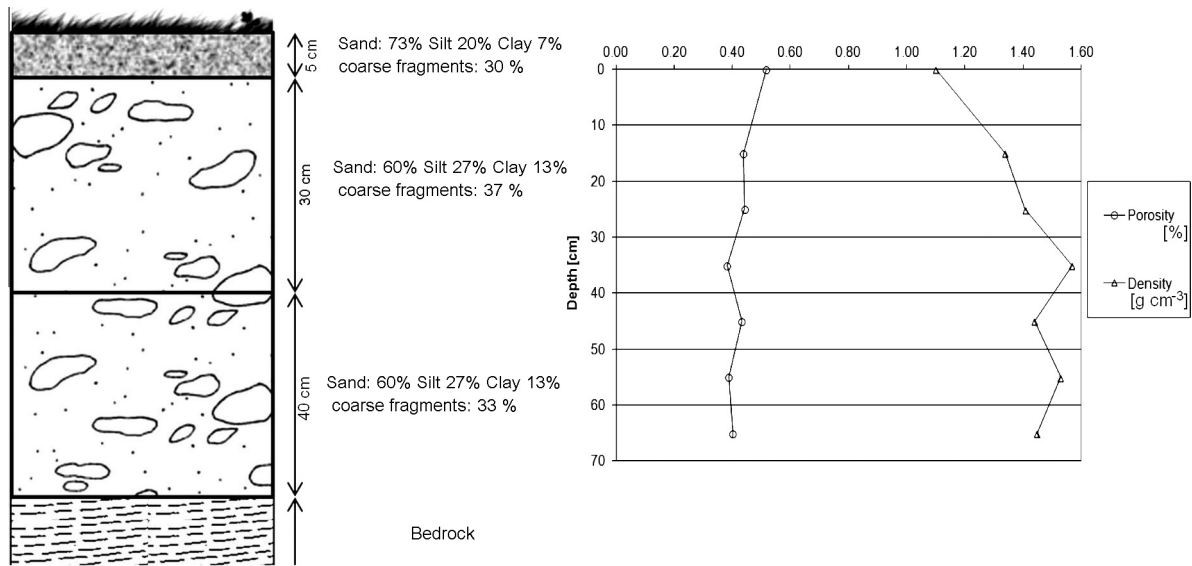


Fig. 2. Variation of density, porosity and granulometry along the soil profile in the Sumène plot.

high water content during autumn–spring period after important rainfall events (Fig. 3). This is representative of the Mediterranean climate, specifically in the Cévennes area. From the three years data set, we observed minimum water content of  $0.08 \text{ cm}^3 \text{ cm}^{-3}$  during summer and maxima of 0.43, 0.47, and  $0.48 \text{ cm}^3 \text{ cm}^{-3}$  respectively for the 20, 40, and 60 cm measurement depths during rainfalls.

#### 1.1.3. Rainfall data and events selection

A tipping-bucket pluviometer (*Précis Mécanique*,  $1000 \text{ cm}^2$ ) was located on the plot. Rain data were collected from 2008 to 2010 and processed at a 15 min time step.

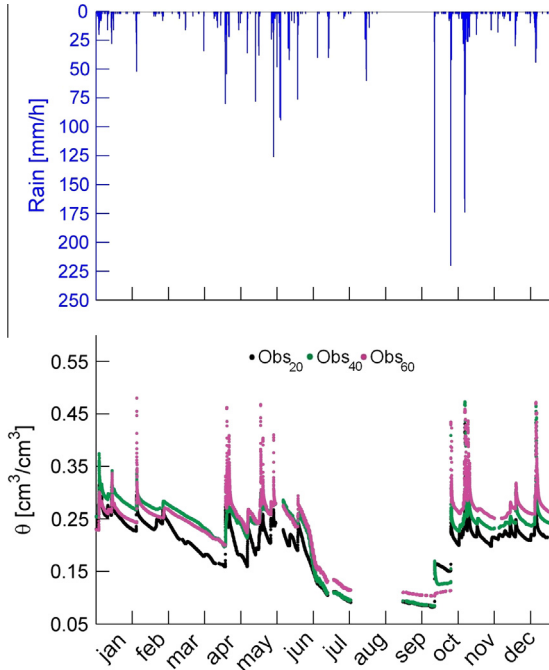
The selection of events is based on rainfall characteristics. They were automatically selected using two simple criteria:

- Accumulated rainfall  $\geq 50 \text{ mm}$ .
- Time separating two different episodes  $\geq 24 \text{ h}$  without rain.

A summary of the 12 selected episodes is given in Table 1 and the representation of the water content variations is given in Fig. 4. We also considered two additional episodes kept apart from the calibration process, which were used for the validation of the calibrated parameters.

Episodes were quite variable as their duration ranges from 1 to nearly 5 days and the cumulated rain varied from 50 to 230 mm. The initial water content varied from  $0.20$  to  $0.28 \text{ cm}^3 \text{ cm}^{-3}$  through all episodes and soil profile (except for episode 4 which is drier). Accordingly to the Mediterranean climate, the 3 episodes presenting the most intense rainfall 4, 5, and 10 – occurred during autumn while the others occurred during winter and spring.

Seven out of the twelve selected episodes and 1 out of the 2 validation episodes presented an observed water content plateau (Fig. 4) at least for the deepest probe. We assumed that those constant measurements during 3 or more time step (i.e. more than 45 min) were representative of the saturation of the soil profile.



**Fig. 3.** Rain and soil water content monitored at 20 (black), 40 (green), and 60 cm (purple) depth in 2008. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Those observations were important to estimate the water infiltration rate at the interface between the soil and the weathered bed-rock and consequently the creation of temporary water tables occurring during the rainfall events.

We noticed a measured saturated water content around  $0.45 \text{ cm}^3 \text{ cm}^{-3}$  with a standard deviation of  $\pm 0.03 \text{ cm}^3 \text{ cm}^{-3}$ . Moreover, saturation was also measured with similar values at 40 and/or 20 cm depth for episodes 4 and 5.

## 1.2. Numerical modeling

### 1.2.1. Model description

The numerical problem was set up using the HYDRUS-1D model (Šimůnek et al., 1998, 2008) to represent the variability of soil water content. Because we chose an event based approach and because most of the selected rainfall episodes occurred during autumn and winter, we assumed that the evapotranspiration

process can be neglected. Thus, soil water fluxes could be represented by the Richards equation (Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[ K(\theta) \left( \frac{\partial h(\theta)}{\partial z} - 1 \right) \right] \quad (2)$$

with  $h$  the soil pressure head (mm),  $\theta$  the soil water content ( $\text{cm}^3 \text{ cm}^{-3}$ ),  $t$  the time (h),  $z$  (positive downward) the vertical coordinate (mm),  $K(\theta)$  the soil hydraulic conductivity function ( $\text{mm h}^{-1}$ ).

We chose to use the Mualem–van Genuchten (MvG) model (Mualem, 1976; van Genuchten, 1980) to describe the water retention curve (WRC) and the water conductivity curve (WCC). The water retention curve was thus given by:

$$\begin{cases} \theta(h) = \theta_s & h \geq 0 \\ \theta(h) = \theta_r + (\theta_s - \theta_r) \left[ \frac{1}{1 + |zh|^\alpha} \right]^{1-\frac{1}{n}} & h < 0 \end{cases} \quad (3)$$

where  $\theta_r$  and  $\theta_s$  ( $\text{cm}^3 \text{ cm}^{-3}$ ) are the residual and saturated water content, and  $\alpha$  ( $\text{mm}^{-1}$ ) and  $n$  (–) are shape parameters.

Hydraulic conductivity was given by:

$$K(\theta) = K_s Se^l \left[ 1 - \left( 1 - Se^{n/(n-1)} \right)^{1-1/n} \right]^2 \quad (4)$$

with  $K_s$  the saturated hydraulic conductivity ( $\text{mm h}^{-1}$ ),  $Se = \frac{\theta - \theta_r}{\theta_s - \theta_r}$  the effective saturation (–), and  $l$  (–) the pore connectivity factor.

Six parameters were thus necessary to define the hydraulic properties of each soil material:  $\theta_r$ ,  $\theta_s$ ,  $\alpha$ ,  $n$ ,  $K_s$ , and  $l$ .

### 1.2.2. Spatial discretization

We chose to consider the  $\theta_{40}$  and  $\theta_{60}$  probes as being in the same soil horizon because the granulometry, porosity, and density measurements showed barely any variations below 30 cm deep. The soil profile was thus divided in three layers: (i) layer<sub>1</sub> from surface to 30 cm deep (including  $\theta_{20}$  probe), (ii) layer<sub>2</sub> from 30 to 70 cm deep (including  $\theta_{40}$  and  $\theta_{60}$  probes), and layer<sub>3</sub> from 70 to 170 cm. Layer<sub>1</sub> and layer<sub>2</sub> represented the soil understood with a pedological meaning, layer<sub>3</sub> represented the weathered schist bed-rock. We chose to regroup the first organic layer (from surface to a 2–5 cm depth) and the first organo-mineral layer to limit the number of estimated parameters since its thickness was negligible compared to the other layers.

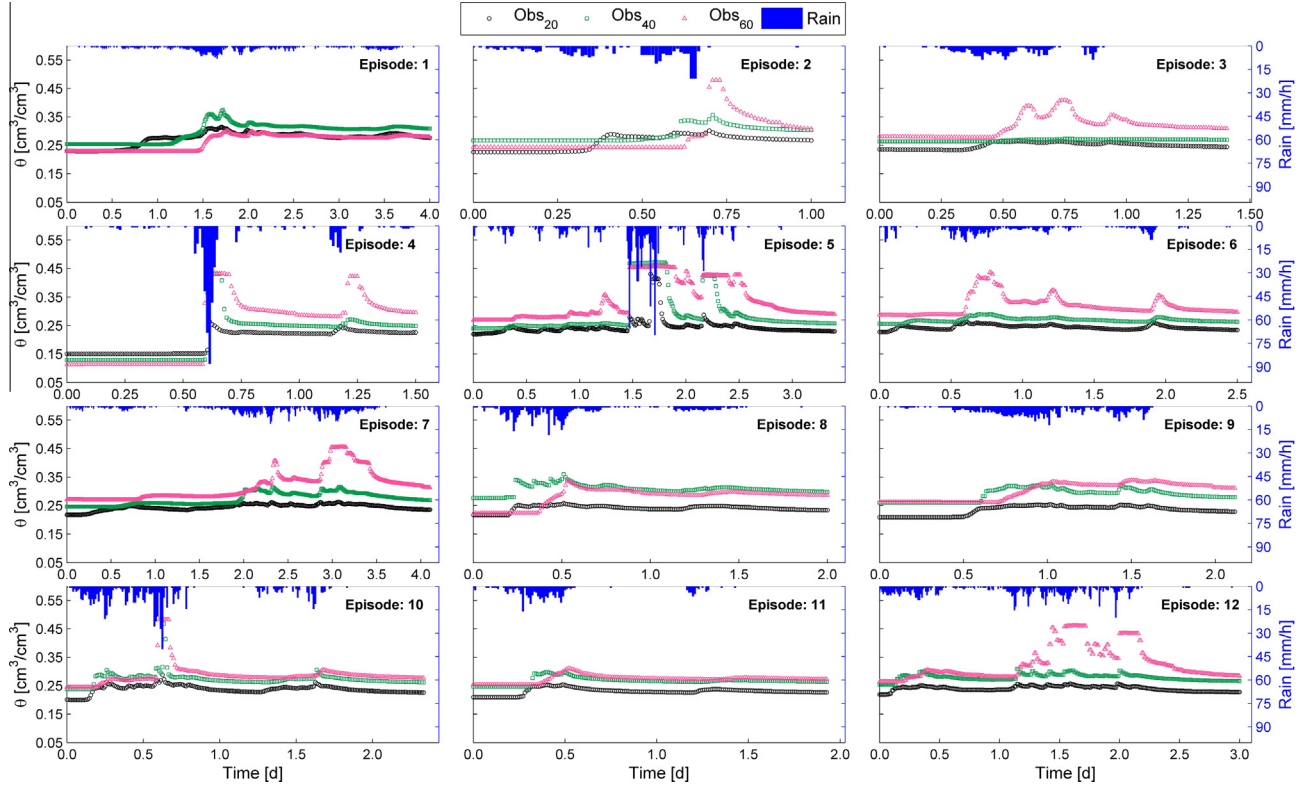
Because we worked with a sandy textured soil, we expected high  $n$  and  $\alpha$  values which required fine discretization especially near the surface and layers transition (Šimůnek et al., 2012). The nodal distance varied from 1 mm at the surface because it is the most weather influenced zone, to 2 mm near layer transitions, and up to 5 mm in the middle of soil layers.

**Table 1**

Characteristics of the 12 calibration episodes and the 2 validation episodes.  $\theta_{x,\text{ini}}$  the initial water content measured at  $x$  cm deep,  $\theta_{x,\text{sat}}$  the water content measured during the saturation plateau, CumRain the cumulated rain observed during the episode and MaxRain the maximum rain intensity observed over 15 min. Episodes presenting saturation plateau are in bold.

Episode	Starting	Ending	Duration (d)	$\theta_{20,\text{ini}}$	$\theta_{20,\text{sat}}$	$\theta_{40,\text{ini}}$	$\theta_{40,\text{sat}}$	$\theta_{60,\text{ini}}$	$\theta_{60,\text{sat}}$	CumRain (mm)	MaxRain ( $\text{mm h}^{-1}$ )
1	02/01/2008	07/01/2008	4.83	0.23	–	0.25	–	0.23	–	112	7
2	<b>03/02/2008</b>	<b>05/02/2008</b>	<b>1.64</b>	<b>0.23</b>	–	<b>0.27</b>	–	<b>0.24</b>	<b>0.48</b>	<b>55</b>	<b>13</b>
3	20/04/2008	21/04/2008	1.41	0.24	–	0.26	–	0.28	–	50	7
4	<b>19/10/2008</b>	<b>23/10/2008</b>	<b>3.43</b>	<b>0.15</b>	–	<b>0.13</b>	<b>0.43</b>	<b>0.11</b>	<b>0.47</b>	<b>120</b>	<b>62</b>
5	<b>31/10/2008</b>	<b>03/11/2008</b>	<b>3.40</b>	<b>0.22</b>	<b>0.43</b>	<b>0.24</b>	<b>0.47</b>	<b>0.27</b>	<b>0.46</b>	<b>232</b>	<b>33</b>
6	04/11/2008	06/11/2008	2.95	0.23	–	0.26	–	0.29	–	87	8
7	<b>31/01/2009</b>	<b>04/02/2009</b>	<b>4.10</b>	<b>0.22</b>	–	<b>0.25</b>	–	<b>0.27</b>	<b>0.46</b>	<b>171</b>	<b>9</b>
8	11/04/2009	13/04/2009	2.35	0.22	–	0.28	–	0.22	–	97	11
9	16/02/2010	18/02/2010	2.11	0.21	–	0.26	–	0.26	–	116	8
10	<b>30/10/2010</b>	<b>01/11/2010</b>	<b>2.33</b>	<b>0.20</b>	–	<b>0.24</b>	–	<b>0.25</b>	<b>0.48</b>	<b>154</b>	<b>23</b>
11	20/11/2010	22/11/2010	1.92	0.21	–	0.25	–	0.25	–	55	8
12	<b>21/12/2010</b>	<b>24/12/2010</b>	<b>3.43</b>	<b>0.22</b>	–	<b>0.25</b>	–	<b>0.26</b>	<b>0.46</b>	<b>158</b>	<b>10</b>
Val_1	09/12/2008	16/12/2008	7.3	0.22	–	0.24	–	0.26	–	104	7
Val_2	<b>26/12/2008</b>	<b>01/01/2009</b>	<b>6.7</b>	<b>0.21</b>	–	<b>0.24</b>	<b>0.47</b>	<b>0.26</b>	<b>0.47</b>	<b>211</b>	<b>15</b>





**Fig. 4.** Rain and soil water content monitored at 20 (black), 40 (green), and 60 cm (purple) depth during the 12 selected rainfall events. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Top boundary condition is only controlled by precipitation ( $\text{mm h}^{-1}$ ).

Soil initial water content has been fixed as follow:

- From top to 30 cm, set to the initial 20 cm measurement value.
- From 30.2 cm to 50 cm, set to the initial 40 cm measurement value.

$$\begin{cases} \text{of}_{20} = \{\text{RMSE}_{20} + \text{RMSE}_{20}(\theta_{\text{obs},20} \geq 0.9 * \theta_{\text{obs},20,\text{max}})\} \\ \text{of}_{40,60} = \frac{\{\text{RMSE}_{40} + \text{RMSE}_{40}(\theta_{\text{obs},40} \geq 0.9 * \theta_{\text{obs},40,\text{max}})\} + \{\text{RMSE}_{60} + \text{RMSE}_{60}(\theta_{\text{obs},60} \geq 0.9 * \theta_{\text{obs},60,\text{max}})\}}{2} \end{cases} \quad (6)$$

- From 50.2 cm to 70 cm, set to the initial 60 cm measurement value.
- For layer<sub>3</sub>, we assumed that layer<sub>3</sub> effective saturation was equal to layer<sub>2</sub> effective saturation so we could define layer<sub>3</sub> initial water content as:

$$\theta_{\text{ini},\text{layer}_3} = \theta_{r,\text{layer}_3} + (\theta_{\text{sat},\text{layer}_3} - \theta_{r,\text{layer}_3}) * \left( \frac{\theta_{\text{ini},\text{layer}_2} - \theta_{r,\text{layer}_2}}{\theta_{\text{sat},\text{layer}_2} - \theta_{r,\text{layer}_2}} \right) \quad (5)$$

Boundary condition at layer<sub>3</sub>'s bottom was considered as free drainage.

### 1.2.3. Inverse optimization algorithm

Inverse estimation of hydraulic properties was performed by minimization of two objective functions. The first one was defined as the sum of root mean square error (RMSE) between observed and simulated soil water. To encourage the model to reproduce the saturation periods and the most important variations, we chose to define the objective function at each monitored depth for a

given episode as the sum of the RMSE calculated from all the measurements and the RMSE calculated from the measurements included between the maximum water content and 90% of this value. For the second layer, which contains two probes, we calculated the mean objective function calculated for the  $\theta_{40}$  and  $\theta_{60}$  probes. Elementary objective functions were defined by:

Because the aim was to determine a representative set of MvG parameters, the inverse problem was solved using all rain episodes simultaneously. We therefore defined the objective functions used by the algorithm as follow:

$$\text{OF}_j(u) = \frac{1}{\text{nbEp}} \sum_{k=1}^{\text{nbEp}} \text{of}_j(u) \quad (7)$$

with  $\text{OF}_j$  the objective function for observed depth  $j \in \{20; 40 \cup 60\}$ ,  $\text{nbEp}$  the number of selected rain episodes,  $\text{of}_j(u)$  the elementary objective function at depth  $j$  for the  $u$  vector of MvG parameters.

To optimize the set up inverse problem, we used the NSGA-II algorithm proposed by Deb et al. (2002) which has been used with success in hydrology (Bekele and Nicklow, 2007; Dumedah et al., 2010) and in the multi-algorithm procedure AMALGAM (Vrugt and Robinson, 2007; Wöhling et al., 2008; Köhne et al., 2011; Wöhling and Vrugt, 2011; Werisch et al., 2014). The algorithm is based on a fast non-dominating sorting approach to find Pareto optimal solutions.

The offspring creation is ensured by crossing-over and mutation operators. The mutation operator has been proved to represent a possible bias when the parameters become close to the extreme boundaries (Hamdan, 2010, 2012). Indeed, the mutation operator initially set up by Srinivas and Deb (1995) creates the child by:

$$c_k = p_k + (p_k^u - p_k^l)\delta_k \quad (8)$$

with  $c_k$  the child chromosome  $k$ th parameter and  $p_k$  the parent chromosomes  $k$ th parameter,  $p_k^u$  and  $p_k^l$  are the upper and lower boundaries for the  $k$ th parameter and  $\delta_k$  the mutation factor. In the original operator,  $\delta_k$  was defined as  $\delta_k = \frac{\min\{p_k - p_k^l, p_k^u - p_k\}}{p_k^u - p_k^l}$ , the minimum distance between the  $k$ th parameter and its min/max boundaries divided by the boundaries difference. Hamdan (2010) exposed that if the parent parameter is close to one of its boundaries,  $\delta_k$  is very small and the mutation process may become useless because of the important likeness between parent and child chromosomes (Fig. 5a). This increases the risk of being trapped in local optima even in multi-objective algorithms. He proposed a version of the mutation process called highly-disruptive mutation.

Instead of having the same  $\delta_k$  distribution on both sides of  $p_k$  parameter, all the boundaries interval can be used as shown in Fig. 5b. It allows bigger jumps in parameters value and gives more chances to escape local optima.

However, those big jumps limited the possibility of getting a precise Pareto front. When the algorithm is close to the solution, too big jumps would add too much disturbance and then reduce its efficiency to find the global optima.

The solution proposed by Hamdan (2010) and used in this work was to use a hybridization of both mutation processes with a probability  $l-w$  of using the non-highly disruptive mutation and  $w$  of using the highly disruptive mutation. After experimental evaluation, we fixed  $w$  to 0.3 so that only 30% of mutations are highly disruptive.

To initialize the NSGA-II population, all rain episodes were treated separately using Hydrus-1D in forward mode and run from an external Matlab code (Matlab R2011b, The MathWorks, Inc.).

To reduce the number of calibrated parameters, we set  $\theta_r$  for each layer to the minimum value measured during summer period which was  $0.08 \text{ cm}^3 \text{ cm}^{-3}$  for the 3 observed depths. The remaining 15 parameters were estimated using the lower and upper boundaries listed in Table 2.

A total of 8000 randomly sampled 15-dimensional parameter sets were generated from parameter boundaries listed in Table 2. For each parameter set, the code attributed the elementary objective function values to each of the 12 episodes.

**Table 2**

Lower and upper boundaries for the calibrated Mualem-van Genuchten parameters.

		$\theta_s$ ( $\text{cm}^3 \text{ cm}^{-3}$ )	$\alpha$ ( $\text{mm}^{-1}$ )	$n$ (-)	$K_s$ ( $\text{mm h}^{-1}$ )	$l$ (-)
Layers 1 & 2	Min	0.30	0.0001	1.01	100	0.000
	Max	0.60	0.1000	5.00	3000	5.000
Layer 3	Min	0.15	0.0001	1.01	1	0.000
	Max	0.50	0.1000	5.00	200	5.000

From this pool of  $12 * 8000 = 96,000$  15-dimensional parameter sets, we defined  $n_i = \frac{N}{nbEp} * \frac{1}{nbOF}$  the number of parameter sets selected for each objective function to create the initial population with  $nbEp$  the number of studied rainfall episodes,  $nbOF$  the number of used objective functions, and  $N$  the size of NSGA-II initial population. To facilitate the procedure, we fixed the population size to 120, so we selected for every episode the 5 best parameter sets for each one of the 2 defined objective functions. This first step guaranteed the diversity in the initial population with multiple parameter sets basically optimized for every rain episodes.

Thanks to the chosen initialization procedure, we set the number of generation to 80.

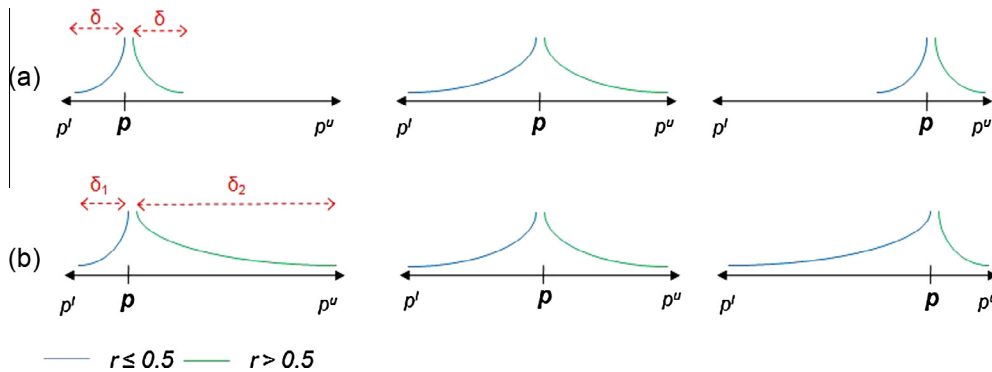
The inverse problem posed can thus be summarized by the following steps:

1. Production of 8000 random 15-dimensional parameter sets of hydraulic properties for each of the 12 selected rainfall episodes.
2. Select the  $n_i = 5$  best 15-dimensional parameter sets for each episode and each objective function to produce the initial population of size  $N = 120$ .
3. Run NSGA-II with crossover and mutation PDI = 5 for 70 generations.
4. Run NSGA-II for 10 more generations with crossover and mutation PDI = 15.

In this case the total number of generation was considered as the convergence criterion.

#### 1.2.4. Selection of the behavioral parameter sets

Besides the two Pareto extreme parameter sets which leads to the best performance for one objective function and the worst for the other, the most interesting solution is the one which could satisfy both objectives functions because it is the compromise-solution (Wöhling et al., 2008). This compromise solution was, for our case, the parameter set that reproduced the water content variation observed at the three monitored depths. This solution



**Fig. 5.** Difference between original mutation coded by Srinivas and Deb (1995) [a] and the highly disruptive mutation [b] introduced by Hamdan (2010) on the child parameter distribution.

was determined by minimizing the Euclidian-distance of the rank-1 Pareto individuals with the Euclidian-distance defined by:

$$ED(u) = \sqrt{(OF_{20}(u) + OF_{20,40}(u))^2} \quad (9)$$

with  $ED(u)$  the Euclidian-distance associated to the parameter set  $u$  depending on its objective functions  $OF_{20}(u)$  and  $OF_{20,40}(u)$ .

To evaluate the dispersion of the estimated parameters, we chose to consider as behavioral-solutions a third of the total population size. Consequently, the 40 best compromised solutions have been considered as behavioral.

## 2. Results and discussions

The aim of this study was to determine the soil hydraulic properties of a loamy sand using the NSGA-II algorithm from soil water content observation during 12 rainfall episodes.

To evaluate the convergence of the algorithm we represented the evolution of the two normalized objective functions best values and the best normalized compromised solution at every generation in Fig. 6.

We noticed a very sharp decrease of all objective functions during the first 10 generations, then they stabilized after 31 generations for the best compromise, after 66 generations for  $OF_{40,60}$ , and only in the last 10 generations for  $OF_{20}$ . Using a population size of 120 individuals, the achievement of the algorithm run represented 116,640 simulations.

After convergence of the algorithm, the solutions of the optimization procedure (i.e. the last generation) which were all part of a rank 1 Pareto Front are represented in Fig. 7.

The  $OF_{20}$  ranged from 0.022 to 0.031  $\text{cm}^3 \text{cm}^{-3}$  while the  $OF_{40,60}$  ranged from 0.057 to 0.065  $\text{cm}^3 \text{cm}^{-3}$ . The distribution was quite uniform and well distributed on all the objective function range which illustrated the capacity of NSGA-II to well explore the solution space in this complex scenario.

The mean optimal parameter values and their standard deviations calculated from the 40 best compromise solutions are given in Table 3.

Those results presented a very low deviation around the best compromised values for all the estimated parameters with the exception of  $l$  of the second layer. We thus considered that the uncertainty of the estimated parameters was very low. We then compared those estimations to the available data. The saturated water content  $\theta_s$  was close to the 0.4–0.44  $\text{cm}^3 \text{cm}^{-3}$  porosity measured on soil samples. Even if the total measured porosity

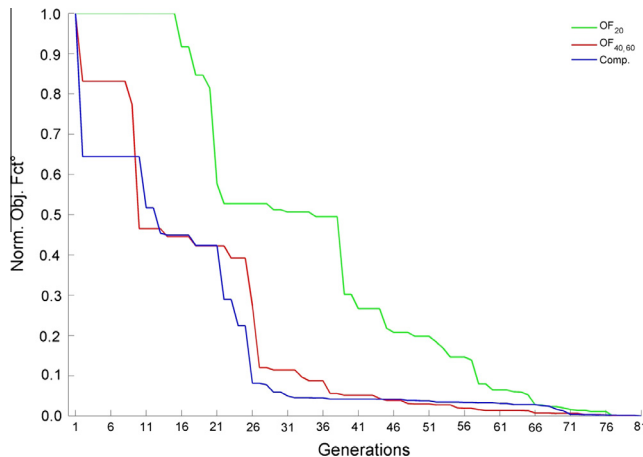


Fig. 6. Evolution of the best normalized objective functions and compromise from the initial algorithm population to the 81th generation.

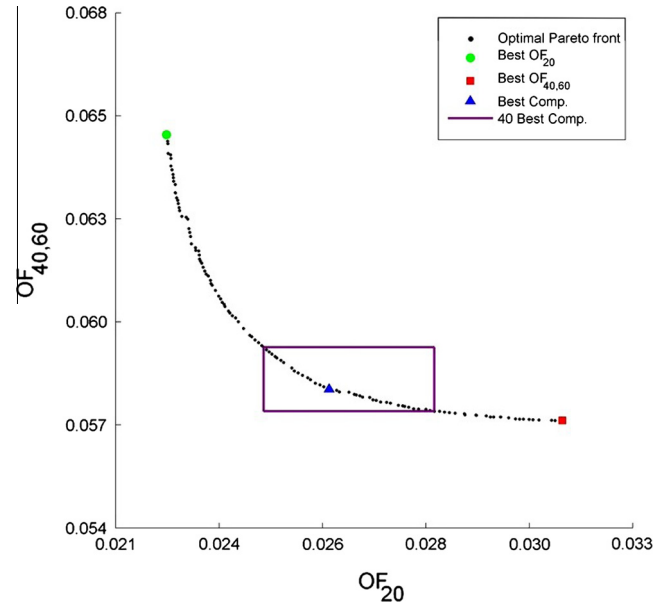


Fig. 7. Representation of the last generation individuals. The black dots represent the last generation rank-1 parameter sets, the best parameter set according to the  $OF_{20,40}$  is represented by the red square, the best parameter set according to the  $OF_{20}$  is represented by the green circle and the best compromise set is represented by the blue triangle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Estimated hydraulic parameters and the standard-deviation associated to each parameter calculated from the 40 best compromised parameter sets.

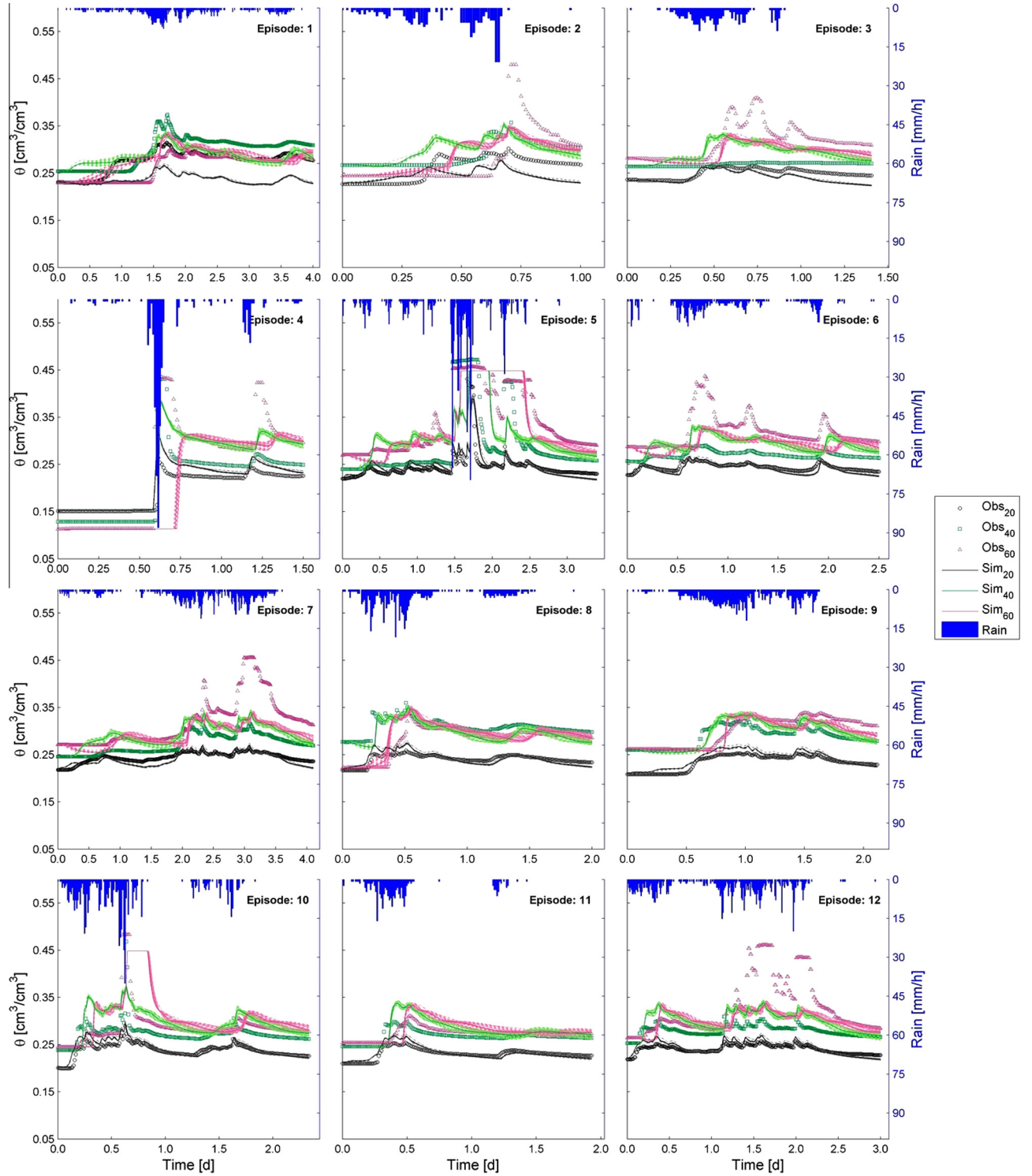
	$\theta_s$ ( $\text{cm}^3 \text{cm}^{-3}$ )	$\alpha$ ( $\text{mm}^{-1}$ )	$n$ (-)	$K_s$ ( $\text{mm h}^{-1}$ )	$l$ (-)
Layer <sub>1</sub>	$0.40 \pm 0.000$	$0.012 \pm 0.001$	$2.05 \pm 0.06$	$2283 \pm 70$	$3.21 \pm 0.06$
Layer <sub>2</sub>	$0.45 \pm 0.000$	$0.088 \pm 0.004$	$1.36 \pm 0.03$	$1923 \pm 5$	$1.44 \pm 0.37$
Layer <sub>3</sub>	$0.48 \pm 0.003$	$0.026 \pm 0.001$	$3.47 \pm 0.02$	$4 \pm 0$	$3.45 \pm 0.02$

was different from the saturated water content, the closeness of those results added confidence to the estimated parameters. The estimated saturated hydraulic conductivity was quite 10 times larger than the infiltrometry results. This important difference was assumed to be a consequence of both the infiltrometer setup which disturb soil surface and the scaling factor considering that the infiltrometer only gave information about a small depth while the modeling approach investigated the entire profile response. About the deep permeability, it is important to highlight the very low saturated hydraulic conductivity at 70 cm, which means that the deep infiltration at 70 mm does not exceed some  $\text{mm h}^{-1}$ . The third layer can thus be considered as schist rocks with a low level of weathering.

The simulations performed with the 40 best compromised solutions are then compared to observations (Fig. 8).

It can be seen that the water contents simulated from the 40 best compromised solutions were always in a narrow range. In addition, the water contents simulated from the best compromise 15-dimensional parameter set were very close to the measurements at 20 cm deep except for episodes 1, 2 and 4. They also fitted together at 40 cm except for episodes 2, 3, 4, 5, and 10, where the model over-estimated the measured water content. At 60 cm deep, the measured water content variation were well simulated only for episodes 1, 5, 8, 9, and 11 so only 5 episode out of 12. It is important to notice that for the most intense episode, the 5th, the measured water contents were well represented at all depths, even if the simulated variation at 40 cm deep was over-estimated during





**Fig. 8.** For the 12 calibration episodes: representation of the simulated (lines) and observed (markers) water content variations at the three monitored depth: 20, 40, and 60 cm. For each monitored depth, the simulation corresponding to the best compromised solution is represented in darker and full line whereas the 39 next best compromised solutions are represented in lighter dashed lines.

the first 36 h. The plateau observed at 40 and 60 cm was especially well represented in time and in amplitude. However, the important variations observed at 60 cm during episodes 2, 3, 4, 6, 7, and 12 were not reproduced by the model, while variations at 20 and 40 cm were well represented.

Episodes 3, 6, 7 and 12, were specials because they presented important water content variations at 60 cm, which did not seem

to be in relation to intense rainfall as those during episode 4 or 5. Episodes 2 and 4 did present intense rainfall but the models did not reproduce correctly the deepest water content variation either.

At this step, we presumed that the primary hypothesis of a vertical one-dimensional flow was not sufficient to represent all possible scenarios and that for some events sub-surface lateral flows

generated an increase of water content from the bottom of the soil profile at the limit with the weathered bedrock. Further simulations would have to be done with a 2D-subsurface software (e.g. Hydrus-2D) to evaluate this new hypothesis.

For the evaluation of the global performance of the model, we computed the errors between the measured and the simulated water content for 3 significant ranges of the observed water contents. Those ranges were:

- $\theta$  varying from 0.15 to 0.25  $\text{cm}^3 \text{cm}^{-3}$  corresponds to the initial water content of the majority of the episodes. It represents the starting period of the episode or the period between two rainfalls.
- $\theta$  varying from 0.25 to 0.35  $\text{cm}^3 \text{cm}^{-3}$  corresponds to small and medium water content variations.
- $\theta \in$  varying from 0.35 to  $\infty$   $\text{cm}^3 \text{cm}^{-3}$  corresponds to the most important water content variations and saturation periods.

The errors were calculated at each time step of all the selected events, and the distribution of these errors was reported in Fig. 9.

Thus, the model was able to reproduce 80–95% of the  $[0.15\text{--}0.25]$   $\text{cm}^3 \text{cm}^{-3}$  water content range with less than 0.05  $\text{cm}^3 \text{cm}^{-3}$  errors. The model reproduced the  $[0.25\text{--}0.35]$   $\text{cm}^3 \text{cm}^{-3}$  water content range with less than 0.05  $\text{cm}^3 \text{cm}^{-3}$  errors in more than 90% of the cases. But the model was able to reproduce only 40% of the water content variations at 20 and 40 cm depth, and 50% at 60 cm depth for the highest water content range:  $[0.35\text{--}0.50]$   $\text{cm}^3 \text{cm}^{-3}$ .

The representation of the simulated water content using the best compromise parameter set on the two validation episodes is

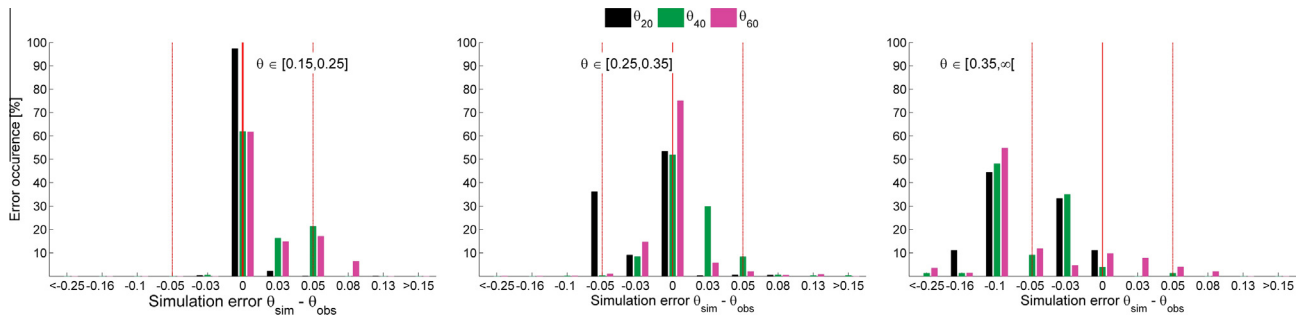
given in Fig. 10. We only used two validation episodes because there were not enough intense rainfall events during the monitored period and we chose not to use too small event because they were not relevant to this flash flood oriented study.

On both episodes, the water content variations at 20 cm deep were represented with less than 0.03  $\text{cm}^3 \text{cm}^{-3}$  error. At 40 cm deep, the simulated water content was globally over-estimated from 0.02 to 0.08  $\text{cm}^3 \text{cm}^{-3}$ . At 60 cm deep, the peaks were over-estimated for episode 1, especially the first one, but the inter-rainfall period was well represented, while for the second episode both inter-rainfall and peaks were well represented with a less than 0.03  $\text{cm}^3 \text{cm}^{-3}$  error. The model was able to make the separation between the two saturation periods during the second episode along with a good representation of the recession at the end of the episode.

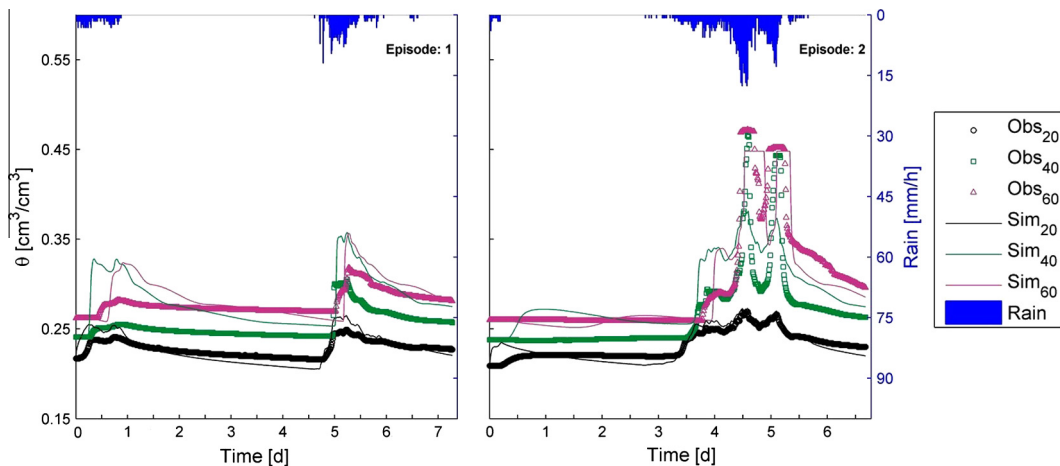
Those results confirmed the performance of the compromise parameter set and the efficiency of the inverse modeling approach used.

### 3. Conclusion

In this study, we investigated an inverse modeling method based on the NSGA-II algorithm and *in situ* soil water content measurements. The study area was a mid-slope land-plot in the Cevennes area (south of France) with a shallow sandy-loam soil. It has been discretized into a 170 cm deep profile with two soil layers of respectively 30 and 40 cm deep above a 100 cm thick layer representing the underlying weathered schist bedrock. The objectives were to assess both the hydrodynamical properties of the soils and



**Fig. 9.** Errors associated for 3 water content range at the three monitored depths. The center thick red vertical line represents an error inferior to 1% between the observed and simulated water content value, and thin vertical lines on the sides represent an error of 5%. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 10.** For the 2 validation episodes: representation of the simulated (lines) and observed (markers) water content variations at the three monitored depth: 20, 40, and 60 cm. For each monitored depth, only the simulation corresponding to the best compromised solution is represented in full line.

the permeability of the underlying bedrock. The NSGA-II algorithm was used to calibrate the parameters of the 2-layers soil as well as the parameters of the third layer considered as the weathered bedrock, where no measurement was available.

First, the inverse modeling set-up converged to a unique and well distributed Pareto front. The 40 behavioral parameter sets led to a very low uncertainty estimation of the Mualem–van Genuchten parameters with a standard-deviation between 3‰ and 4‰ for 14 of the 15 calibrated parameters and 25% for the pore connectivity factor of the second soil layer. The ability of the model to reproduce the measured water content in the first soil layer has been proved for 10 out of 12 calibration episodes and for the 2 validation episodes. Model performance for the second soil layer was more contrasted. In one hand, 8 episodes presented an underestimation of the water-content variation at 60 cm deep and an overestimation of the water content at 40 cm deep. In the other hand, 6 episodes presented a simulation of the water content variation with less than 5% errors, including the saturation plateau on the most intense rainfall event. The two episodes devoted to the validation of the calibrated parameters confirmed those results with a low intensity episode presenting an over-estimation of the water content in the second soil layer and a high-intensity episode with a good representation of the saturation plateau and the recession dynamic at the end of the rainfall event. The fact that the model under-estimated the important water content variations at 60 cm depth for 6 out of the 9 episodes brought into question the hypothesis of one-dimensional transfer. It appeared that for those episodes, sub-surface flux may have influenced the deep water content variations, and would have to be taken into account by 2D-modeling.

In parallel with the performance of the inverse modeling set-up results, we showed that this method allowed us to access to the hydrodynamic properties of the weathered bedrock although no measurements were available considering the high density of coarse rock fragments in this layer. We estimated the saturated hydraulic conductivity of this bedrock was less than  $5 \text{ mm h}^{-1}$ . This was able to generate saturated areas at the bottom of the soil column, as it was observed. Here, saturation reached at least 50 cm above the bottom of the soil, and probably more in case of an extreme event. Thus, the selected inverse modeling method appeared to be pertinent to assess the hydrodynamic properties of the soils and the underlying bedrock. In both cases, this information is valuable for hydrological rainfall–runoff modeling. This method is planned to be extended to a larger area with a multiplication of the monitoring station in order to assess the spatial variability of the soil hydrodynamic properties in the Cévennes area.

## Acknowledgements

This study was performed within the frame of the FloodScale project, which is funded by the French National Research Agency (ANR) under Contract no ANR 2011 BS56 027, which contributes to the HyMeX program. This study was also supported by the French Observatoire Hydro-Météorologique Cévennes-Vivarais (OHM-CV).

## References

Basile, A., Ciollaro, G., Coppola, A., 2003. Hysteresis in soil water characteristics as a key to interpreting comparisons of laboratory and field measured hydraulic properties. *Water Resour. Res.* 39 (12), 1355.  
 Bekele, E.G., Nicklow, J.W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* 341, 165–176.  
 Berthet, L., Andréassian, V., Perrin, C., Javelle, P., 2009. How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event based and continuous approaches on 178 catchments. *Hydrol. Earth Syst. Sci.* 13, 819–831.

Beven, K.J., 1993. Prophecy, reality and uncertainty in disturbed hydrological modeling. *Adv. Water Resour.* 16, 41–51.  
 Braud, I. et al., 2010. The use of distributed hydrological models for the Gard 2002 flash flood event: analysis of associated hydrological processes. *J. Hydrol.* 394 (1–2), 162–181.  
 Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-2. *IEEE Trans. Evolutionary Comput.* 6 (2), 182–197.  
 Delta-T Devices Ltd, 1999. ThetaProbe Soil Moisture Sensor, Type ML2-X User Manual. Cambridge, England, p. 22.  
 Dirksen, C., 1999. Direct conductivity measurements for evaluating approximate and indirect determinations. In: van Genuchten, M.T. et al. (Eds.), *International Workshop on Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*. University of California, Riverside, pp. 271–278.  
 Dumedah, G., Berg, A.A., Wineberg, M., Collier, R., 2010. Selecting model parameter sets from a trade-off surface generated from the non-dominated sorting genetic algorithm-II. *Water Resour. Manage* 24, 4469–4489.  
 Durner, W., Lipsius, K., 2005. *Encyclopedia of Hydrological Sciences*, vol. 6. John Wiley & Sons, Ltd.  
 Efstratiadis, A., Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrol. Sci. J.* 55 (1), 58–78.  
 Hamdan, M., 2010. On the disruption-level of polynomial mutation for evolutionary multi-objective optimisation algorithms. *Comput. Inform.* 29, 783–800.  
 Hamdan, M., 2012. The distribution index in polynomial mutation for evolutionary multiobjective optimisation algorithms: an experimental Study. In: *International Conference on Electronics Computer Technology*, Kanyakumari, India.  
 Hopp, L., McDonnell, J.J., 2009. Connectivity at the hillslope scale: identifying interactions between storm size, bedrock permeability, slope angle and soil depth. *J. Hydrol.* 376 (3–4), 378–391.  
 Horn, J., Nafpliotis, N., Goldberg, D.E., 1994. A niched Pareto genetic algorithm for multiobjective optimization. In: *Piscataway (Ed.), Forst IEEE Conference on Evolutionary Computation*. Z. Michalewicz, pp. 82–87.  
 James, A.L., McDonnell, J.J., Tromp-van Meerveld, I., Peters, N.E., 2010. Gypsies in the palace: experimentalist's view on the use of 3-D physics-based simulation of hillslope hydrological response. *Hydrol. Process.* 24 (6), 3878–3893.  
 Köhne, J.M. et al., 2011. Coupled simulation of surface runoff and soil water flow using multi-objective parameter estimation. *J. Hydrol.*, 403.  
 Loew, A., Mauser, W., 2008. Inverse modeling of soil characteristics from surface soil moisture observations: potential and limitations. *Hydrol. Earth Syst. Sci. Discuss.* 5, 95–145.  
 Madsen, H., 2003. Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives. *Adv. Water Resour.* 26, 205–216.  
 Mualem, Y., 1976. A new model predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12, 513–522.  
 Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *J. Appl. Phys.* 1, 318.  
 Ritter, A., Hupet, F., Munoz-Carpena, R., Lambot, S., Vanlooster, M., 2003. Using inverse methods for estimating soil hydraulic properties from field data as an alternative to direct methods. *Agric. Water Manag.* 59, 77–96.  
 Schelle, H., Durner, W., Iden, S.C., Fank, J., 2013. Simultaneous estimation of soil hydraulic and root distribution parameters from lysimeter data by inverse modeling. *Proc. Environ. Sci.* 19, 564–573.  
 Scharnagl, B., Vrugt, J.A., Vereecken, H., Herbst, M., 2011. Inverse modeling of in situ soil water dynamics: investigation the effect of different prior distributions of the soil hydraulic parameters. *Hydrol. Earth Syst. Sci.* 15, 3043–3059.  
 Šimůnek, J., Šejna, J.M., van Genuchten, M.T., 1998. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. IGWMC – TPS – 70, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, p. 186.  
 Šimůnek, J., van Genuchten, M.T., Šejna, M., 2008. Development and applications of the HYDRUS and STANMOD software packages and related codes. *Vadose Zone J.* 7 (2), 587–600.  
 Šimůnek, J., van Genuchten, M.T., Šejna, M., 2012. HYDRUS: model use, calibration and validation. *Am. Soc. Agric. Biol. Eng.* 55 (4), 1261–1274.  
 Srinivas, N., Deb, K., 1995. Multiobjective function optimization using nondominated sorting genetic algorithms. *Evol. Comput.* 2 (3), 221–248.  
 Trambly, Y. et al., 2010. Assessment of initial soil moisture conditions for event-based rainfall–runoff modelling. *J. Hydrol.* 387 (3–4), 176–187.  
 Tseng, P.-H., Jury, W.A., 1993. Simulation of field measurement of hydraulic conductivity in unsaturated heterogeneous soil. *Water Resour. Res.* 29, 2087–2099.  
 van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892–898.  
 Vannier, O., Braud, I., Anquetin, S., 2013. Regional estimation of catchment-scale soil properties by means of stream flow recession analysis for use in distributed hydrological models. *Hydrol. Process.* 28 (26), 6276–6291.  
 Vrugt, J.A., Robinson, B.A., 2007. Improved evolutionary optimization from genetically adaptive multimethod search. *Proc. Natl. Acad. Sci. USA* 104, 432–442.  
 Vrugt, J.A., Stauffer, P.H., Wöhling, T., Robinson, B.A., Vesselinov, V.V., 2008. Inverse modeling of subsurface flow and transport properties: a review with new developments. *Vadose Zone J.* 7 (2), 843–864.

Werisch, S., Grundmann, J., Al-Dhuhli, H., Algharibi, E., Lennartz, F., 2014. Multiobjective parameter estimation of hydraulic properties for a sandy soil in Oman. *Environ. Earth Sci.* 72 (12), 4935–4956.

Wöhling, T., Vrugt, J.A., 2011. Multiresponse multilayer vadose zone model calibration using Markov chain Monte Carlo simulation and field water retention data. *Water Resour. Res.* 47 (W04510), 1–19.

Wöhling, T., Vrugt, J.A., Barkle, G.F., 2008. Comparison of three multiobjective optimization algorithms for inverse modeling of vadose zone hydraulic properties. *Soil Sci. Soc. Am. J.* 72 (2), 305–319.

Wollschläger, U., Pfaff, T., Roth, K., 2009. Field scale effective hydraulic parameterization obtained from TDR time series and inverse modeling. *Hydrol. Earth Syst. Sci.* 6, 1489–1522.