

Document downloaded from:

<http://hdl.handle.net/10251/81817>

This paper must be cited as:

Han, X.; Franssens, J.; Rosolem, R.; Jiménez Bello, MA.; Bogena, H.; Martínez Alzamora, F.; Chanzy, A.... (2016). Simultaneous soil moisture and properties estimation for a drip irrigated field by assimilating cosmic-ray neutron intensity. *Journal of Hydrology*. 539:611-624. doi:10.1016/j.jhydrol.2016.05.050.



The final publication is available at

<http://doi.org/10.1016/j.jhydrol.2016.05.050>

Copyright Elsevier

Additional Information

1 Simultaneous Soil Moisture and Properties Estimation  
2 for a Drip Irrigated Field by Assimilating Cosmic-ray  
3 Neutron Intensity

4  
5 Xujun Han<sup>1,2</sup>, Harrie-Jan Hendricks Franssen<sup>1,2</sup>, Miguel Ángel Jiménez Bello<sup>3</sup>, Rafael  
6 Rosolem<sup>4</sup>, Heye Bogena<sup>1</sup>, Fernando Martínez Alzamora<sup>3</sup>, André Chanzy<sup>5</sup>, Harry Vereecken<sup>1,2</sup>

- 7  
8 1. Forschungszentrum Jülich, Agrosphere (IBG 3), Leo-Brandt-Strasse, 52425 Jülich,  
9 Germany  
10 2. Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys,  
11 Geoverbund ABC/J, Leo-Brandt-Strasse, 52425 Jülich, Germany  
12 3. Institute of Water and Environmental Engineering (IiAMA). Universitat Politècnica de  
13 Valencia. Camino de Vera, s/n, 46022 Valencia, Spain  
14 4. Department of Civil Engineering, University of Bristol, Bristol BS8 1TR, UK  
15 5. INRA, UMR1114, Environnement Méditerranéen et Modélisation des AgroHydrosystèmes,  
16 Domaine Saint Paul, Site Agroparc, F-84914 Avignon, France

17  
18 Corresponding author: Xujun Han, Forschungszentrum Jülich, Agrosphere (IBG 3),  
19 Leo-Brandt-Strasse, 52425 Jülich, Germany. (x.han@fz-juelich.de)

20 **Abstract**

21 Neutron intensity measured by the aboveground cosmic-ray neutron intensity  
22 probe (CRP) allows estimating soil moisture content at the field scale. In this work,  
23 synthetic neutron intensities were used to remove the bias of simulated soil moisture  
24 content or update soil hydraulic properties (together with soil moisture) in the  
25 Community Land Model (CLM) using the Local Ensemble Transform Kalman Filter.  
26 The cosmic-ray forward model COSMIC was used as the non-linear measurement  
27 operator which maps between neutron intensity and soil moisture. The novel aspect of  
28 this work is that synthetically measured neutron intensity was used for real time  
29 updating of soil states and soil properties (or soil moisture bias) and posterior use for  
30 the real time scheduling of irrigation (data assimilation based real-time control  
31 approach). Uncertainty of model forcing and soil properties (sand fraction, clay  
32 fraction and organic matter density) were considered in the ensemble predictions of  
33 the soil moisture profiles. Horizontal and vertical weighting of soil moisture was  
34 introduced in the data assimilation in order to handle the scale mismatch between the  
35 cosmic-ray footprint and the CLM grid cell.

36 The approach was illustrated in a synthetic study with the real-time irrigation  
37 scheduling of fields of citrus trees. After adjusting soil moisture content by  
38 assimilating neutron intensity, the irrigation requirements were calculated based on  
39 the water deficit method. Model bias was introduced by using coarser soil texture in  
40 the data assimilation experiments than in reality. A series of experiments was done  
41 with different combinations of state, parameter and bias estimation in combination  
42 with irrigation scheduling.

43 Assimilation of CRP neutron intensity improved soil moisture characterization.  
44 Irrigation requirement was overestimated if biased soil properties were used. The soil  
45 moisture bias was reduced by 35% after data assimilation. The scenario of joint  
46 state-parameter estimation resulted in the best soil moisture characterization (50%  
47 decrease in root mean square error compared to open loop simulations), and the best  
48 estimate of needed irrigation amount (86% decrease in Hausdorff distance compared

49 to open loop). The coarse scale synthetic CRP observation was proven to be useful for  
50 the fine scale soil moisture and soil properties estimation for the objective of  
51 irrigation scheduling.

52

53 **Keywords:** Data Assimilation; Cosmic ray; Soil Moisture; Parameter Estimation; Bias  
54 Estimation; Irrigation Scheduling

55 **1. Introduction**

56 Globally, 70% of fresh water is used by agriculture (FAO - Food and Agriculture  
57 Organization of the United Nations). Therefore, it is necessary to increase the water  
58 use efficiency and reduce the water need for crop production, while maintaining crop  
59 yield. Enough water should be applied to meet the requirement of maximum crop  
60 evapotranspiration (ET). Farmers usually base irrigation scheduling on their own  
61 experience taking into account soil water status and crop growth. However, it is  
62 unlikely that the optimal scheduling of irrigation is acquired without the knowledge of  
63 crop water needs. Low cost sensors that measure soil moisture content can be of  
64 advantage. However, these sensors typically have a very small measurement volume  
65 which is much smaller than the scale of the fields where the crops are grown.  
66 Numerical models like crop growth models (Heng et al., 2009) and land surface  
67 models (Wood et al., 2011) can be used for the quantitative estimation of the irrigation  
68 requirement under specific soil water and crop growth conditions. The estimated  
69 irrigation amount can be applied accurately with new agricultural technology like drip  
70 irrigation (Sampathkumar et al., 2012). However, uncertain model input data and  
71 deficits in the model structure result in biased estimates of soil water status, crop  
72 transpiration and therefore irrigation requirement.

73 The optimal scheduling of irrigation is complicated given the high heterogeneity  
74 of soil moisture content in drip irrigated fields. An estimate of soil moisture content  
75 for the complete root zone is important in this context. It is difficult to achieve this  
76 with small-scale measurements (e.g., TDR-Time Domain Reflectometry,

77 FDR-Frequency Domain Reflectometry or TDT-Time Domain Transmission) as a  
78 prohibitively large number of sensors is needed to cover large irrigated areas. Soil  
79 moisture information from remote sensing on the other hand is limited to the upper  
80 few soil centimeters, and often has a very coarse horizontal resolution (>10 km)  
81 (Entekhabi et al., 2010; Kerr et al., 2010; Montzka et al., 2013). A further limitation  
82 of satellite-derived soil moisture content is that it is not reliable for highly vegetated  
83 areas (Njoku and Chan, 2006) and high uncertainties (Merlin et al., 2009; Montzka et  
84 al., 2013). The spatial variability of soil moisture is controlled by soil hydraulic  
85 properties, meteorological forcing, land cover patterns and topographic features at  
86 different measurement scales. Small scale variability is more driven by soil hydraulic  
87 properties while large scale variability is also more driven by the other factors. Hence,  
88 strengths and weaknesses of each measurement method rely on the additional  
89 uncertainty given by these additional controlling factors (Crow et al., 2012).

90 A new promising method which can determine integral root zone soil moisture  
91 from the measured above ground fast neutron intensity (defined as the number of  
92 counted neutrons per unit of time – e.g., counts per hour) has been proposed (Zreda et  
93 al., 2012). This synthetic study focuses on the assimilation of cosmic-ray probe (CRP)  
94 neutron intensity (Bogena et al., 2013; Desilets et al., 2010; Rosolem et al., 2014;  
95 Shuttleworth et al., 2013; Zreda et al., 2008; Zreda et al., 2012). Soil moisture  
96 measurements at the intermediate scale of the cosmic ray probe have the advantage  
97 that they are less affected by small scale variability of soil hydraulic properties. A  
98 further advantage is that soil moisture can be determined for a deeper layer (10-70 cm)

99 in higher temporal frequency than remote sensing (Rosolem et al., 2014).

100 Primary cosmic rays originate from our galaxy and eventually collide with  
101 atmospheric nuclei, generating secondary cosmic rays mainly consisting of neutrons  
102 (Lal and Peters, 1967). Primary cosmic rays create cascades of secondary high-energy  
103 neutrons through colliding with atmospheric nuclei and the high-energy neutrons can  
104 penetrate the atmosphere and collide with nuclei in soils. These collisions in the soil  
105 generate fast neutrons. Some of these fast neutrons are eventually scattered back to  
106 the atmosphere and the fast neutron intensity can be measured with the CRP. The  
107 measured intensity of fast neutrons above the ground depends strongly on soil  
108 moisture content (Hendrick and Edge, 1966; Zreda et al., 2012). CRPs make use of  
109 this principle to estimate soil moisture content for an area of about 600 m diameter  
110 and variable measurement depth (~10-70 cm) depending on the soil moisture  
111 conditions (Zreda et al., 2012).

112 Measured neutron intensities above ground need to be corrected for variations in  
113 incoming high-energetic neutrons and atmospheric pressure (Zreda et al., 2012).  
114 Moreover, as the measured neutron intensity depends on additional sources of  
115 hydrogen (besides of soil moisture), these need to be taken into account in order to  
116 isolate the soil moisture signal. Corrections have been proposed for other hydrogen  
117 sources like atmospheric vapor (Rosolem et al., 2013), lattice water and organic  
118 carbon in the soil (Franz et al., 2013), hydrogen atoms stored in the litter layer  
119 (Bogena et al., 2013) and above-ground biomass (Baatz et al., 2015). Data  
120 assimilation studies have shown the advantage of using measured multi-source soil

121 moisture observations for improving the soil moisture profile characterization of a  
122 land surface model (Crow et al., 2008; De Lannoy et al., 2007b; Han et al., 2012;  
123 Huang et al., 2008; Reichle et al., 2008; Walker et al., 2001). Measured neutron  
124 intensities have already been used for assimilation in a land surface model to improve  
125 estimates of soil moisture profiles, but the model parameters were calibrated a priori  
126 (Han et al., 2015a; Rosolem et al., 2014; Shuttleworth et al., 2013).

127 In this paper we will investigate the benefits of assimilating coarse scale (600 m)  
128 neutron intensity data into the Community Land Model (CLM) for the application of  
129 drip irrigation for citrus trees on a finer scale (100 m) than the CRP scale. The neutron  
130 intensity measured by a synthetic CRP affects a larger area than a typical irrigation  
131 management unit (1 ha in this work). In order to study the impact of soil moisture data  
132 assimilation on irrigation scheduling, the drip irrigation was therefore simulated at a  
133 finer spatial scale than the footprint of a CRP. The drip irrigation was applied at the  
134 vegetated area and resulted in a very heterogeneous soil moisture distribution with the  
135 alternation of patches of wet and dry soil. It is very CPU-intensive to explicitly model  
136 the irrigated patches and the non-irrigated parts, and a simplified implementation was  
137 adopted in this work, which will be further detailed in the methodology section. In the  
138 simulation experiments, CLM was driven by biased soil properties to mimic the  
139 intrinsic model uncertainties. The coarse scale CRP neutron intensity observations  
140 were used to update the field scale heterogeneous soil moisture field through data  
141 assimilation. The joint soil moisture and soil properties (or soil moisture bias)  
142 estimation scheme was evaluated. This is important because soil moisture content and



143 crop transpiration are sensitive to model parameters (Hou et al., 2012; Rosolem et al.,  
144 2012; Schwinger et al., 2010). Typically, field measurements of parameter values are  
145 scarce and very uncertain, especially because of the scale mismatch between a local  
146 measurement and the model scale (Waller et al., 2014). Model parameter estimation  
147 in the context of a data assimilation framework was proven to be successful, using  
148 either an augmented state vector approach (Chen and Zhang, 2006), dual state  
149 parameter estimation (Moradkhani et al., 2005b) or parameter estimation in a loop  
150 external to the data assimilation filter (Vrugt et al., 2005). Successful applications are  
151 reported for such diverse areas as groundwater hydrology (Franssen and Kinzelbach,  
152 2008; Kurtz et al., 2014; Schöniger et al., 2012), rainfall-runoff models (Moradkhani  
153 et al., 2005a; Vrugt et al., 2006), land surface models (Han et al., 2014a; Pauwels et  
154 al., 2009), vadose zone hydrology (Montzka et al., 2011; Wu and Margulis, 2013) and  
155 atmospheric models (Ruiz et al., 2013). A data assimilation framework can consider  
156 uncertain model forcing, model structure and initial conditions, as well as parameter  
157 uncertainties. Data assimilation has become a commonly used method for parameter  
158 estimation, especially for large scale applications (Wanders et al., 2014).

159 Joint soil moisture and soil moisture bias estimation has been proven to be helpful  
160 for improving data assimilation results (De Lannoy et al., 2007a; Kumar et al., 2012b)  
161 like soil temperature assimilation with bias correction (Bosilovich et al., 2007;  
162 Reichle et al., 2010). In this study, we also evaluated the impact of the soil moisture  
163 bias estimation method (Dee, 2005) on improving the soil moisture assimilation and  
164 irrigation scheduling and compared it with joint state-parameter estimation.

165 It is expected that a more accurate characterization of the heterogeneous soil  
166 moisture distribution can be obtained if the coarse scale CRP neutron intensity data  
167 are assimilated using a combination of data assimilation and parameter estimation (or  
168 bias estimation). Based on such results, it is then assumed that the estimated irrigation  
169 requirement could be improved. The objective of this study is to evaluate with help of  
170 a synthetic study: 1) the potential of measured neutron intensity data by the CRP for  
171 improving the characterization of soil moisture content and soil properties (or soil  
172 moisture bias), and 2) the impact of the assimilation of neutron intensity on better  
173 irrigation scheduling and the potential for real-time irrigation optimization. In this  
174 study, the spatial variability of soil properties and crop status will be considered in the  
175 data assimilation.

176

## 177 **2. Methodology**

178 The main components of the methodology are: (i) measurement of above-ground  
179 neutron intensity, which is linked to field scale soil moisture content by a  
180 measurement operator (section 2.1) and horizontal weights (section 2.3); (ii) the land  
181 surface model CLM (version 4.5) which simulates the transport of water and energy  
182 in the soil-plant-atmosphere continuum (section 2.2); (iii) data assimilation according  
183 to the Local Ensemble Transform Kalman Filter (LETKF) methodology (Hunt et al.,  
184 2007) which optimally combines measurements and model predictions to update soil  
185 moisture (and possibly soil properties or soil moisture bias), taking into account  
186 uncertain atmospheric forcing and model parameters (or model bias) (section 2.3) and

187 (iv) an optimization routine which calculates irrigation need for the ensemble of soil  
 188 moisture forecasts (section 2.4).

## 189 **2.1. Cosmic ray Soil Moisture Interaction Code (COSMIC)**

190 In order to assimilate neutron intensity, the relationship between neutron intensity  
 191 and depth-weighted soil moisture content should be reasonably represented. The  
 192 newly-developed COsmic ray Soil Moisture Interaction Code (COSMIC)  
 193 (Shuttleworth et al., 2013) was adopted as the forward observation operator to  
 194 simulate the equivalent neutron count rates from simulated soil moisture profiles (i.e.,  
 195 soil moisture contents for 10 vertical model layers of CLM from surface to 3 m depth,  
 196 in this study) and takes into account the weighted contribution of individual soil  
 197 layers with depth. The COSMIC operator calculates the number of fast neutrons  
 198 reaching the CRP  $N_{COSMOS}$  at a near-surface measurement point by:

$$199 \quad N_{COSMOS} = N \int_0^{\infty} \left\{ A(z) [\alpha \rho_s(z) + \rho_w(z)] \exp \left( - \left[ \frac{m_s(z)}{L_1} + \frac{m_w(z)}{L_2} \right] \right) \right\} dz \quad (1)$$

$$200 \quad A(z) = \left( \frac{2}{\pi} \right)^{\pi/2} \int_0^{\pi/2} \exp \left( \frac{-1}{\cos(\varphi)} \left[ \frac{m_s(z)}{L_3} + \frac{m_w(z)}{L_4} \right] \right) d\varphi \quad (2)$$

$$201 \quad \alpha = 0.405 - 0.102 \times \rho_s \quad (3)$$

$$202 \quad L_3 = -31.76 + 99.38 \times \rho_s \quad (4)$$

203 where  $N$  (counts/h) is the number of high-energy neutrons at the soil surface,  $z$   
 204 is the soil layer depth (m),  $\rho_s$  is the dry soil bulk density ( $\text{g cm}^{-3}$ ),  $\rho_w$  is the total  
 205 soil water density, including the lattice water ( $\text{g cm}^{-3}$ );  $m_s(z)$  and  $m_w(z)$  are the  
 206 integrated mass per unit area of dry soil and water ( $\text{g cm}^{-2}$ ),  $\varphi$  is the angle between  
 207 the vertical below the detector and the line between the detector and each point in the

208 plane (Shuttleworth et al., 2013),  $L_1$  is the high energy soil attenuation length with  
209 value of  $162.0 \text{ gcm}^{-2}$ ,  $L_2$  is the high energy water attenuation length of  $129.1 \text{ gcm}^{-2}$ ,  
210  $L_3$  is the fast neutron soil attenuation length ( $\text{gcm}^{-2}$ ) and  $L_4$  is the fast neutron water  
211 attenuation length with value of  $3.16 \text{ gcm}^{-2}$  (Shuttleworth et al., 2013).

212 In this study soil moisture contents for 10 vertical soil layers of CLM were used  
213 to drive the COSMIC operator. COSMIC interpolates the soil moisture to 300 layers  
214 with a soil profile depth of 3 meters and derives the fast neutron count rate from the  
215 depth-averaged soil moisture content based on the effective sensor depth, also  
216 calculated by COSMIC. The simulated fast neutron intensity was assumed to be  
217 observed and subsequently used for the data assimilation as will be explained below.  
218 Vertical variation of soil moisture content will be considered according to the  
219 contribution to the total neutron intensity of each soil layer.

220

## 221 **2.2. Community Land Model - CLM**

222 The Community Land Model (CLM - version 4.5) developed by the National  
223 Center for Atmospheric Research (NCAR) was used to calculate soil moisture content  
224 and evapotranspiration (Oleson et al., 2013). CLM uses the simplified Richards  
225 equation to model water flow in the unsaturated zone and calculation of land surface  
226 energy fluxes is done by invoking the Monin-Obukhov similarity theory. In addition,  
227 the following processes can be simulated by CLM: transfer of solar radiation and  
228 longwave radiation, stomatal physiology and photosynthesis, crop dynamics and  
229 irrigation (Oleson et al., 2013). The land cover can be represented by 17 plant

230 functional types (PFTs) and the calculation of energy fluxes is based on the PFTs.  
 231 Hydraulic and thermal parameters in CLM are derived based on soil properties such  
 232 as sand and clay fraction and organic matter density (Oleson et al., 2013).

233

### 234 **2.3. Data Assimilation and Parameter Estimation**

235 The Local Ensemble Transform Kalman Filter (LETKF) is a square root  
 236 ensemble Kalman filter (Hunt et al., 2007; Miyoshi and Yamane, 2007), which is  
 237 applied extensively in atmospheric data assimilation studies (Arav équia et al., 2011;  
 238 Baek et al., 2006; Lien et al., 2013; Miyoshi et al., 2014) and also in land data  
 239 assimilation studies (Han et al., 2014a). In LETKF, the uncertainty of the model  
 240 forecast is represented by ensemble members. In this study LETKF is used to estimate  
 241 both soil moisture and soil properties with the state augmentation method (Bateni and  
 242 Entekhabi, 2012; Han et al., 2014a; Li and Ren, 2011) or to update soil moisture and  
 243 soil moisture bias jointly.

244 First two matrices  $\mathbf{X}^b$  and  $\mathbf{Y}^b$  are constructed based on simulated soil moisture  
 245 and soil properties (or soil moisture bias) of the ensemble members:

$$246 \quad \mathbf{X}^b = [\mathbf{x}_1^b - \bar{\mathbf{x}}^b, \dots, \mathbf{x}_M^b - \bar{\mathbf{x}}^b] \quad (5)$$

$$247 \quad \mathbf{y}_i^b = H(\mathbf{x}_i^b) \quad (6)$$

$$248 \quad \mathbf{Y}^b = [\mathbf{y}_1^b - \bar{\mathbf{y}}^b, \dots, \mathbf{y}_M^b - \bar{\mathbf{y}}^b] \quad (7)$$

249 where  $\mathbf{x}_1^b, \dots, \mathbf{x}_M^b$  are vectors with the ensemble members,  $M$  is the ensemble  
 250 size,  $\bar{\mathbf{x}}^b$  is the vector with ensemble means calculated over  $\mathbf{x}_1^b, \dots, \mathbf{x}_M^b$ ,  $H$  is the  
 251 observation operator (i.e., COSMIC for soil moisture),  $\mathbf{y}_i^b$  is the mapping of the

252 ensemble members  $\mathbf{x}_1^b, \dots, \mathbf{x}_M^b$  to the measurement space and  $\bar{\mathbf{y}}^b$  is the vector of  
253 ensemble means of  $\mathbf{y}_1^b, \dots, \mathbf{y}_M^b$ . The vector  $\bar{\mathbf{x}}^b$  contains i) the depth weighted average  
254 soil moisture  $\theta_{cosmic}$ , which was derived from COSMIC and considers the  
255 contribution of different soil layers, ii) the soil moisture of 10 layers ( $\theta_1 \dots \theta_{10}$ ) and iii)  
256 soil properties (sand fraction, clay fraction and organic matter density) in case soil  
257 properties are estimated. In case of bias estimation, the vector  $\bar{\mathbf{x}}^b$  contains soil  
258 moisture bias instead of soil properties. The dimensions of the augmented state vector  
259  $\bar{\mathbf{x}}^b$  were 11 for the state estimation only, 14 for joint state and parameter estimation,  
260 and 12 for joint state and bias estimation. Only soil properties and soil moisture bias  
261 for the upper soil layer were included in the state vector  $\bar{\mathbf{x}}^b$ . Soil properties for all 10  
262 layers were updated based on the ratio of the soil properties between the upper soil  
263 layer and the lower soil layers (Han et al., 2014b). The soil moisture bias of deeper  
264 layers was assumed to decrease exponentially according to equation (8).

$$\bar{\mathbf{x}}^b = \left\{ \begin{array}{l} \left[ \begin{array}{c} \theta_{cosmic} \\ \theta_1 \\ \vdots \\ \theta_{10} \end{array} \right] \quad \textit{state estimation} \\ \left[ \begin{array}{c} \theta_{cosmic} \\ \theta_1 \\ \vdots \\ \theta_{10} \\ \textit{Sand} \\ \textit{Clay} \\ \textit{Organic} \end{array} \right] \quad \textit{parameter estimation} \\ \left[ \begin{array}{c} \theta_{cosmic} \\ \theta_1 + \theta_{Bias}^{k-1} * \exp(-1.0 * Z_1) \\ \vdots \\ \theta_{10} + \theta_{Bias}^{k-1} * \exp(-1.0 * Z_{10}) \\ \theta_{Bias}^{k-1} \end{array} \right] \quad \textit{bias estimation} \end{array} \right. \quad (8)$$

266 where  $Z_i$  (m) is the  $i^{\text{th}}$  soil layer thickness of CLM and  $k$  is the time step.

267 Next, the analysis error covariance matrix  $\mathbf{P}^a$  is calculated:

$$268 \quad \mathbf{P}^a = \left[ (M-1)\mathbf{I} + \mathbf{Y}^{bT} \mathbf{R}^{-1} \mathbf{Y}^b \right] \quad (9)$$

269 where  $\mathbf{R}$  is the observation error covariance matrix. The perturbations in the  
270 ensemble space  $\mathbf{W}^a$  are calculated according to:

$$271 \quad \mathbf{W}^a = \left[ (M-1)\mathbf{P}^a \right]^{1/2} \quad (10)$$

272 The analysis mean  $\bar{\mathbf{w}}^a$  is given by:

$$273 \quad \bar{\mathbf{w}}^a = \mathbf{P}^a \mathbf{Y}^{bT} \mathbf{R}^{-1} (\mathbf{y}^o - \bar{\mathbf{y}}^b) \quad (11)$$

274 where  $\mathbf{y}^o$  is the vector with the measured CRP neutron intensity. The analysis  
275 mean is added to each column of  $\mathbf{W}^a$  to get the analysis ensemble.

276 Finally, the new analysis  $\mathbf{X}^a$  is obtained according to:

$$277 \quad \mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a + \bar{\mathbf{x}}^b \quad (12)$$

278 where  $\mathbf{X}^a$  are the model ensemble members after analysis.  $\mathbf{X}^a$  includes the  
279 updated soil moisture and soil properties (or soil moisture bias), and will be used as  
280 initial condition for the next time step.

281 The forecast model of soil moisture bias  $\theta_{Bias}^k$  was defined as:

$$282 \quad \theta_{Bias}^k = \theta_{Bias}^{k-1} \quad (13)$$

283 where  $k$  is the time step.

284 In this study, the LETKF was applied as a 1D filter to update soil moisture in the  
285 cosmic ray footprint, which covered several fine scale model grid cells. Neutron  
286 intensity data were used to update the fine scale soil moisture within the expected  
287 CRP footprint and the contribution of different grid cells within the CRP footprint to  
288 the observed neutron intensity was taken into account by assigning different  
289 horizontal weights to the CLM grid cells. Following the horizontal weighting method  
290 proposed by (Bogena et al., 2013), we used a Gaussian window function (Harris, 1978)  
291 to define the horizontal weights of the different CLM grid cells in the CRP footprint.  
292 Because a single observation was used to update the soil moisture content of many  
293 CLM grid cells, a simple multi-scale data assimilation strategy was proposed. First,  
294 the COSMIC operator was run for each CLM grid cell to simulate the neutron  
295 intensity of each CLM grid cell. Next, a convolution with a Gaussian window  
296 function was applied to retrieve the integrated neutron intensity at the coarse scale.  
297 This convolution step was used in both the ensemble runs and the parallel runs  
298 (introduced in section 3.3). Therefore the CRP neutron intensity (measured/simulated)  
299 considered the contributions from all model grid cells contained within the CRP



300 footprint. The size of the Gaussian window function was chosen as the diameter of the  
301 CRP footprint (600 m). The observation operator  $H$  (needed by equation (6)) was  
302 defined as the combination of a Gaussian window function and the COSMIC operator.  
303 Using equation (6) we get:

$$304 \quad \mathbf{y}_i^b = f_{Gaussian}(f_{cosmic}(\mathbf{x}_i^b)) \quad (14)$$

305 where  $f_{cosmic}$  represents the COSMIC operator and  $f_{Gaussian}$  the Gaussian  
306 window function.

307 The footprint of the CRP covered 43 CLM grid cells (i.e., ~600 m diameter). The  
308 data assimilation was done grid cell by grid cell in LETKF. Therefore the time series  
309 of a single neutron intensity observation was assigned uniformly to all grid cells of the  
310 CRP footprint and then the observations were assimilated separately for each CLM  
311 grid cell. The spatial localization was applied on the model states using equation (14),  
312 and therefore the observation localization was not used in this study (Greybush et al.,  
313 2011; Han et al., 2015b). The soil moisture measured by CRP is composed of the  
314 contribution of all horizontal and vertical CLM grid cells of the CRP footprint.  
315 Therefore, there is only one relevant soil moisture content value and neutron intensity  
316 value. Given equation (14), the upscaling of soil moisture from CLM will be done  
317 before assimilation, it means the upscaled soil moisture  $\theta_{cosmic}$  which incorporates the  
318 contribution from all surrounding grid cells will be used as the soil moisture content  
319 value to be updated with the CRP neutron intensity measurement. All the soil  
320 moisture content values for the individual grid cells within the CRP footprint (both  
321 horizontally and vertically) will be updated according to the correlation between the

322 simulated values by CLM and  $\theta_{cosmic}$  :

$$323 \quad \theta_{cosmic} = f_{Gaussian}(f_{cosmic}(\mathbf{x}_i^b)) \quad (15)$$

324

#### 325 **2.4. Irrigation Requirement**

326 CLM computes the water deficit between the current soil moisture content and a  
327 target soil moisture content. The target soil moisture in each soil layer is a weighted  
328 average of (1) the minimum needed soil moisture content to avoid water stress for that  
329 layer and (2) the saturated soil moisture content for that layer (Levis and Sacks,  
330 2011):

$$331 \quad \theta_{target,i} = (1-0.7) \times \theta_{o,i} + 0.7 \times \theta_{sat,i} \quad (16)$$

332 where  $i$  is the soil layer number,  $\theta_{o,i}$  is the minimum soil moisture content of  
333 each vegetation type so that stomata are completely opened and  $\theta_{sat,i}$  is the effective  
334 soil porosity.

335 The total water deficit  $W_{deficit}$  (mm) was defined as:

$$336 \quad W_{deficit} = \sum_i^N Root\_Fraction_i \times \max(\theta_{target,i} - \theta_{liq,i}, 0) \quad (17)$$

337 where  $\theta_{liq,i}$  is the current soil moisture content of layer  $i$ .

338 The estimated irrigation amount  $W_{deficit}$  was applied in CLM as an incoming  
339 water flux not subjected to interception by the canopy layer (precipitation on the  
340 contrary was subjected to interception).

341 The root fraction ( $RF_i$ ) of citrus trees for the soil layer  $i$  was parameterized as  
342 (Oleson et al., 2013):

$$RF_i = \begin{cases} 0.5 \left[ \begin{array}{l} \exp(-r_a Z_{h,i-1}) + \exp(-r_b Z_{h,i-1}) \\ -\exp(-r_a Z_{h,i}) + \exp(-r_b Z_{h,i}) \end{array} \right] & 1 \leq i < N_{levsoi} \\ 0.5 \left[ \exp(-r_a Z_{h,i-1}) + \exp(-r_b Z_{h,i-1}) \right] & i = N_{levsoi} \end{cases} \quad (18)$$

where  $Z_{h,i}$  (m) is the depth from the soil surface to the interface between soil layers  $i$  and  $i+1$  ( $Z_{h,0} = 0$  represents the soil surface),  $r_a$  and  $r_b$  are plant dependent root distribution parameters, for citrus trees:  $r_a = 8.992$ ,  $r_b = 8.992$  (Zeng, 2001),  $N_{levsoi} = 10$  is the total number of soil layers.

348

## 349 2.5. Performance Measures

350 In order to evaluate the results of data assimilation and irrigation, the Root Mean  
351 Square Error (RMSE) was calculated for the hourly soil moisture results:

$$RMSE = \sqrt{\frac{\sum_{k=i}^K (\theta_{est} - \theta_{ref})^2}{K}} \quad (19)$$

353 where  $\theta_{est}$  is the hourly soil moisture ensemble mean for a given scenario,  $\theta_{ref}$   
354 is the hourly soil moisture value of the reference scenario.  $K$  is equal to 8760. Soil  
355 moisture contents at the time points of assimilation and prediction were included in  
356 the calculations with Eq. 19. Lower RMSE values mean better performance.

357 The Hausdorff Distance (HD) is a quantitative measure of the similarity between  
358 two spatial distributions (Kumar et al., 2012a). Lower HD values mean higher spatial  
359 similarity. HD is defined as the maximum distance of a set to the nearest point in the  
360 other set:

$$h(Q, R) = \max_{q \in Q} \left\{ \min_{r \in R} \{q - r\} \right\} \quad (20)$$

362 where  $h(Q, R)$  is the HD value and  $q$  and  $r$  are the points of sets  $Q$  and

363  $R$ .  $q-r$  is the norm of the points in the space of  $Q$  and  $R$ , in terms of Euclidean  
364 distance.  $Q$  is the estimated annual irrigation amount of a CLM grid cell in the CRP  
365 footprint and  $R$  is the annual irrigation amount of the reference scenario of the CRP  
366 footprint.

367 The t-test can be used to determine whether two data sets are significantly  
368 different. The independent two-sample t-test was used to evaluate the statistical  
369 significance of the difference between the estimated irrigation amount and the  
370 reference irrigation amount (Welch, 1947). The definition of the t-test is as follows:

$$371 \quad t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \quad (21)$$

372 where  $\bar{X}_1 / \bar{X}_2$ ,  $s_1^2 / s_2^2$  and  $N_1 / N_2$  are the mean, variance and the number of  
373 optimized irrigation amounts, respectively. The subscripts 1 and 2 represent the  
374 reference scenario and estimation scenario, respectively.

375

### 376 **3. Synthetic Experiment**

377 A synthetic study was conducted to evaluate the methodology outlined in the  
378 previous sections. The synthetic study mimicked the Picassent site (close to Valencia,  
379 Spain) with citrus trees, which receives drip irrigation. The site is situated in a  
380 semi-arid region (39.38° N, 0.47° E) with yearly average precipitation of 454 mm (44  
381 precipitation days), average daily maximum temperature of 22.3°C and average daily  
382 minimum temperature of 13.4°C, and with a yearly irrigation period from April to  
383 October. In the synthetic experiments, a CRP (Zreda et al., 2012) measured the

384 neutron intensity which was assimilated in the land surface model CLM.

385

### 386 **3.1. Design of Synthetic Experiments**

387 The citrus tree was modeled as a broadleaf evergreen tropical PFT in CLM. In  
388 order to mimic the planting pattern of citrus trees at the Picassent site, odd CLM grid  
389 columns were modelled as bare ground without vegetation cover while even CLM  
390 columns were modelled as fully covered by broadleaf evergreen tropical trees. The  
391 mimicked ground cover of vegetation was similar to the real measured ground cover  
392 fraction of citrus trees. Soil properties (sand fraction and clay fraction) and organic  
393 matter density determine the soil thermal and hydraulic properties in CLM. Soil  
394 samples from the Picassent region were taken and a soil sand fraction of 32%, clay  
395 fraction of 33%, bulk density of  $1.5 \text{ gcm}^{-3}$ , and organic fraction of 1.2% (10 cm depth)  
396 and 0.7% (50 cm depth) were determined and applied uniformly in space in CLM.  
397 Spatially correlated noise was added to the uniform soil properties to represent the  
398 spatial heterogeneity. The noise was simulated by sequential Gaussian simulation with  
399 a correlation range of 100 m, variance of 100.0 (%) and mean value of 0.0 (%). The  
400 noise was constrained within the range of (-5.0%, 5.0%) after the sequential Gaussian  
401 simulation. The atmospheric forcing data measured by the Picassent weather station  
402 were used as CLM input. The maximum rate of carboxylation at  $25 \text{ }^{\circ}\text{C}$  ( $V_{cmax25}$ )  
403 controls the maximum rate of carboxylation and canopy transpiration in CLM. The  
404 default value of  $V_{cmax25}$  is  $55 \text{ } \mu\text{molm}^{-2}\text{s}^{-1}$  for broadleaf evergreen tropical tree in  
405 CLM. It was changed to  $100 \text{ } \mu\text{molm}^{-2}\text{s}^{-1}$  according to a reference value for citrus trees

406 (Velikova et al., 2012). The study area was discretized in  $40 \times 60$  grid cells at a spatial  
407 resolution of 100 m. However, our analysis will focus on the CRP footprint which  
408 contains only 43 CLM grid cells in total. Because of the spatial discretization, it is not  
409 easy to describe the exact CRP footprint. The diameter of the CRP footprint for the  
410 horizontal weighting calculation was chosen as the ratio between the CRP diameter  
411 and the spatial resolution of the CLM grid cell, which was 7 grid cells.

412 A model spin-up was made to obtain reasonable initial conditions for CLM. CLM  
413 was run for the spin-up period from 2010-01-01 to 2010-12-31 using an hourly time  
414 step. Next, a one-year (2011-01-01 to 2011-12-31) irrigation estimation using the true  
415 soil properties was performed in CLM and the irrigation requirement was calculated  
416 every three days during the period from 2011-01-01 to 2011-12-31 and subsequently  
417 applied. This reference model run also supplied the reference soil moisture  
418 distribution in space and time and the canopy transpiration, and a distribution in space  
419 and time (and total amounts) of irrigation.

420

### 421 **3.2. Ensemble Generation**

422 The simulation experiments evaluated how well the soil moisture,  
423 evapotranspiration, soil properties and irrigation requirement could be characterized if  
424 soil properties and model forcing were biased and/or uncertain, but the measurement  
425 data (albeit synthetically generated for this study) in the form of CRP neutron  
426 intensity were available. We assumed that the soil texture was systematically coarser  
427 than in reality. The soil properties used in the reference run were perturbed to

428 represent these conditions: the sand fraction was multiplied by 1.5 while the clay  
429 fraction and organic matter density were both multiplied by 0.75. In addition, for all  
430 scenarios the sand and clay fractions, and organic matter density were perturbed by a  
431 uniform distributed noise in the range of [-10.0%, 10.0%] (for soil texture) and [-10.0  
432  $\text{km m}^{-3}$ , 10.0  $\text{km m}^{-3}$ ] (for organic matter density) to generate 20 different soil  
433 hydraulic properties for the 20 different ensemble members. It is assumed that these  
434 perturbations represent a realistic representation of the uncertainty in practice. If the  
435 sum of the sand and clay fraction was larger than 98%, an amount equal to  
436  $((\text{Sand\%} + \text{Clay\%}) - 98\%) / 2.0$  was subtracted from both the sand and the clay fraction.  
437 For the ensembles for which initial soil moisture bias was updated instead of soil  
438 properties, soil moisture was perturbed with a spatially uniform value sampled from  
439 the uniform distribution with values between  $-0.04 \text{ m}^3\text{m}^{-3}$  and  $0.04 \text{ m}^3\text{m}^{-3}$ .

440 The atmospheric forcing data of precipitation, air temperature, shortwave incident  
441 radiation and longwave incident radiation were perturbed with a noise correlated in  
442 space and time (Han et al., 2013). The spatially correlated noise was generated using  
443 the Fast Fourier Transform approach and the temporal correlation was imposed by a  
444 first-order auto-regressive model approach (Kumar et al., 2009; Park and Xu, 2009).  
445 The perturbation parameters are summarized in Table 1.

446

### 447 **3.3. Synthetic Observation**

448 In this study, the “observed” synthetic CRP neutron intensity was generated using  
449 the COSMIC model, which used the soil moisture profile simulated by CLM as input.

450 A point which had to be considered in the simulation scenarios was that the soil  
451 moisture changed also as response to the irrigation amount applied in the CLM  
452 simulation. Therefore, the synthetic observations will be different for each simulation  
453 scenario due to the different irrigation amounts. In order to obtain the synthetic  
454 measurements of CRP neutron intensity for these scenario runs, a parallel run was  
455 made with the same soil properties and atmospheric forcing as the reference run for  
456 each of the simulation scenarios. However, the irrigation amount for this parallel run  
457 was the same as the optimized one for the specific simulation scenario. Therefore  
458 irrigation amounts differed among the parallel runs of the different scenarios. This in  
459 turn also affected the CRP intensity which was assimilated in the simulation scenarios.  
460 The soil moisture calculated in the parallel run was used now as the synthetic soil  
461 moisture measurement for this specific scenario and also used as input for COSMIC  
462 to estimate the synthetic CRP neutron intensity. This implies that the assimilated CRP  
463 neutron intensity differed among different scenarios, in correspondence with different  
464 irrigation amounts applied in the different scenarios.

465 In this synthetic study,  $N_{COSMOS}$  was set to 150 counts h<sup>-1</sup> and lattice water was 3%  
466 in COSMIC. The simulated CRP intensity measurements were perturbed in order to  
467 represent the observation error. This perturbation had a mean equal to zero and a  
468 variance equal to the square root of the neutron intensity to represent the observation  
469 error. The CLM grid cells located in the footprint of the CRP were updated by the data  
470 assimilation procedure. The neutron intensity observation was assimilated at 23:00  
471 each three days, prior to irrigation scheduling.



472 Fig. 1 shows a schematic overview of the complete calculation procedure for each  
473 time step.

474 [\[Insert Figure 1 here\]](#)

475

### 476 **3.4. Irrigation Scheduling**

477 In Valencia, citrus trees are typically irrigated 3~6 days per week from April to  
478 October. For simplicity, the irrigation duration was assumed to be two hours from  
479 06:00 AM onwards, and the irrigation was applied every three days for reasons of  
480 computational efficiency. CLM was run in prediction mode for three days to estimate  
481 the needed amount of irrigation (at 06:00) using the predefined atmospheric forcing  
482 data. The first assimilation of soil moisture data was done at 23:00, seven hours  
483 before the start of the second irrigation period.

484 Five irrigation estimation scenarios were designed to assimilate the CRP neutron  
485 intensity in the land surface model and to evaluate the impact of data assimilation,  
486 parameter estimation and bias estimation on the characterization of the soil moisture  
487 profile, irrigation amount and evapotranspiration:

488 (1) Irrigation estimation with the true soil properties and atmospheric forcing  
489 (Reference).

490 (2) Irrigation estimation with biased soil properties, without CRP neutron  
491 intensity assimilation (No\_DA, i.e. Open loop).

492 (3) Irrigation estimation with biased soil properties, with the assimilation of CRP  
493 neutron intensity (every three days) (Only\_DA\_SM).

494 (4) Irrigation estimation with biased soil properties, with both CRP neutron  
495 intensity assimilation (every three days) and soil moisture bias estimation  
496 (DA\_SM\_Bias).

497 (5) Irrigation estimation with biased soil properties, with neutron intensity  
498 assimilation (every three days) including soil properties estimation (DA\_SM\_Par).  
499 For the scenario of DA\_SM\_Par, the study period included both 2010 and 2011. In  
500 2010, joint soil moisture and soil properties estimation were carried out in order to  
501 update the biased soil properties (sand fraction, clay fraction and organic matter  
502 density). Next, in 2011 the updated soil properties were used for soil moisture  
503 assimilation without soil properties updating. For the scenarios without soil properties  
504 estimation, the experiments were only carried out for 2011.

505 The estimation of soil properties during the year 2010 deteriorated in case of  
506 intensive irrigation. In case of intensive irrigation, the soil column is close to  
507 saturation, and the ensemble members show a limited spread. In addition, the  
508 sensitivity of soil moisture with respect to soil properties becomes small. Parameter  
509 estimation (i.e., updating of soil properties) is not very efficient during these periods.  
510 Therefore, soil properties were not updated if the accumulated irrigation amount  
511 between two data assimilation time steps was larger than 10 mm.

512

#### 513 **4. Results**

514 In this section we evaluate time series for the different simulation scenarios at the  
515 CRP location. Spatial patterns of (estimated) soil properties and optimized irrigation

516 amounts for different simulation scenarios are also compared. This comparison is  
517 made at the scale of the complete CRP footprint. The temporal evolution of soil  
518 moisture content at 30 cm and 50 cm depth for different simulation scenarios is shown  
519 in Fig. 2. The scenario No\_DA underestimated soil moisture content even although  
520 (too) high irrigation amounts were scheduled, which is related to the erroneous soil  
521 texture in this simulation scenario and the associated bias in soil properties like  
522 saturated hydraulic conductivity. In the data assimilation scenario Only\_DA\_SM soil  
523 moisture characterization was improved by assimilating the CRP neutron intensity,  
524 but the bias was also very high due to the large bias of soil properties. In order to  
525 reduce the soil moisture bias, the joint soil moisture state and bias estimation was  
526 evaluated in the scenario DA\_SM\_Bias by use of the state augmentation method. The  
527 soil moisture estimation became better than Only\_DA\_SM after the soil moisture bias  
528 was reduced. The best soil moisture results were obtained in the scenario DA\_SM\_Par,  
529 in which sand fraction, clay fraction and organic matter density were updated  
530 sequentially using the joint state and parameter estimation method.

531 [\[Insert Figure 2 here\]](#)

532 The RMSE values for soil moisture characterization are summarized in Fig. 3.  
533 Compared with the Reference scenario, the scenarios No\_DA and Only\_DA\_SM  
534 underestimated soil moisture content, which resulted in higher irrigation requirement  
535 than for the reference case. Compared with the scenario No\_DA, the RMSE values  
536 for soil moisture content at 30 cm depth decreased by 33%, 40% and 52% for the  
537 scenarios Only\_DA\_SM, DA\_SM\_Bias and DA\_SM\_Par, respectively. At 50cm

538 depth these RMSE-decreases were 39%, 35% and 51% for the scenarios  
539 Only\_DA\_SM, DA\_SM\_Bias and DA\_SM\_Par, respectively. These results illustrate  
540 the benefit of joint state-parameter estimation. Model results were strongly influenced  
541 by the biased soil properties for the scenario Only\_DA\_SM, where measurement data  
542 were not used to estimate model bias or soil properties. As model bias was related to  
543 biased soil properties in these simulations, joint state and parameter estimation  
544 performed better than the scenario with joint state and bias estimation.

545 [\[Insert Figure 3 here\]](#)

546 Fig. 4 shows the soil moisture bias for the different ensemble members (scenario  
547 DA\_SM\_Bias). The temporal evolution of the bias is shown for the first soil layer and  
548 the CRP location. The time series of true soil moisture bias (the scenario No\_DA) is  
549 shown for comparison. The mean bias value for No\_DA is  $0.051 \text{ m}^3\text{m}^{-3}$  and  $0.033$   
550  $\text{m}^3\text{m}^{-3}$  for DA\_SM\_Bias. The joint state-bias estimation could reduce 35% of the soil  
551 moisture bias introduced by the biased soil properties.

552 [\[Insert Figure 4 here\]](#)

553 The CLM model derives the soil hydraulic parameters using a predefined  
554 pedotransfer function (Oleson et al., 2013). The updated soil texture resulted therefore  
555 in updated soil hydraulic parameters in CLM. In order to show the influence of  
556 calibrated soil texture on the soil hydraulic parameters, the temporal evolution of the  
557 calibrated saturated hydraulic conductivity  $K$  and the empirical parameter  $B$  which  
558 represent the slope of water retention curve in the Clapp–Hornberger parameterization  
559 (Clapp and Hornberger, 1978; Oleson et al., 2013) are shown in Fig. 5 (scenario

560 DA\_SM\_Par). The Clapp–Hornberger parameterization can be used to calculate the  
561 hydraulic conductivity when the soil water retention data are not available. All soil  
562 hydraulic parameters could be improved in this scenario by assimilating the CRP  
563 neutron intensity. The mean ensemble values for the soil hydraulic parameters  
564 approached the reference values over the assimilation period. As the soil properties  
565 were not updated during the intensive irrigation period, the convergence was slow.

566 [\[Insert Figure 5 here\]](#)

567 Fig. 6 shows the time series of irrigation amount at the location of the CRP for  
568 the reference case and for the different estimation scenarios. The reference irrigation  
569 amount (707.2 mm) was calculated in CLM on the basis of the water deficit method.  
570 Fig. 6 illustrates that the intensive irrigation period from June to October coincided  
571 with limited precipitation. The reference irrigation amount was around twice of the  
572 annual precipitation amount (356.1 mm). The sum of the reference irrigation amount  
573 (707.2 mm) and annual precipitation (356.1 mm) was about 6% higher than the  
574 documented potential annual evapotranspiration of citrus trees (Ballester et al., 2011;  
575 Jimenez-Bello et al., 2015). We used a t-test to compare the estimated annual  
576 irrigation requirements by the different scenarios with the reference case. Large p  
577 values (Fig. 6) indicate that the scenarios with data assimilation did not have a  
578 significantly different irrigation scheduling as compared with the reference case. The  
579 scenario No\_DA has a p value <0.05, which indicates that the estimated irrigation  
580 amounts were significantly different from the reference case. Data assimilation  
581 (Only\_DA\_SM) improved results with higher p values (p=0.205). Scenarios with bias

582 correction ( $p=0.819$ ) and parameter estimation ( $p=0.755$ ) gave better results than the  
583 scenario of state estimation only (Only\_DA\_SM). The annual irrigation amounts for  
584 the grid cell at the location of the CRP are summarized in Fig. 7. Obviously, joint soil  
585 moisture and bias (or parameter estimation) improved the characterization of  
586 irrigation requirement.

587 [\[Insert Figure 6 here\]](#)

588 [\[Insert Figure 7 here\]](#)

589 Now we analyze the irrigation results for the scenarios in detail. The better  
590 characterization of the irrigation demand was found by combining CRP neutron  
591 intensity assimilation and soil properties estimation (or soil moisture bias estimation).  
592 Scenarios DA\_SM\_Bias and DA\_SM\_Par estimated an irrigation requirement of  
593 691.0 mm and 685.6 mm, respectively. The scenario Only\_DA\_SM (i.e., data  
594 assimilation with state estimation only) gave an irrigation estimation of 797.7 mm.  
595 CRP neutron intensity assimilation without parameter estimation provided much  
596 better results than scenarios without data assimilation (No\_DA). The estimated  
597 irrigation amount for the scenario No\_DA was 1107.3 mm. These results illustrate that  
598 irrigation estimation for the case of biased soil properties (i.e., sandier soil in model  
599 than in reference run) can be improved significantly by the assimilation of CRP  
600 neutron intensity and even better by including soil properties estimation or bias  
601 estimation. Data assimilation improved the estimation of irrigation demand and even  
602 resulted in a slightly lower irrigation than for the reference case.

603 The CRP footprint was composed of 43 CLM grid cells which were irrigated

604 separately. The irrigation requirements estimated for the different scenarios are  
605 displayed in Fig. 8 and compared with the reference scenario. The spatial irrigation  
606 patterns of scenarios DA\_SM\_Bias and DA\_SM\_Par are closer to the reference case  
607 than the scenario without data assimilation (No\_DA).

608 [\[Insert Figure 8 here\]](#)

609 The single CRP neutron intensity measurement for the coarse scale (600 m) was  
610 used to update the 43 CLM grid cells at the fine scale (100 m). The spatial distribution  
611 of soil properties and annual irrigation amount were compared with the reference  
612 spatial distributions. A higher spatial similarity is associated with lower HD values.  
613 The HD values were evaluated according to the distance between a CLM grid cell and  
614 the CRP location. Three classes were defined: (i) distance CRP- grid cell  $\leq 100$  m; (ii)  
615 distance CRP- grid cell  $> 100$  m and  $\leq 200$  m; (iii) distance CRP- grid cell  $> 200$  m  
616 and  $\leq 300$  m. The HD values for the comparison of the spatial patterns of soil  
617 properties for the scenarios No\_DA and DA\_SM\_Par are shown in Fig. 9. The  
618 similarity of soil properties between background and reference is small due to the  
619 imposed bias, and the HD values were 580.83, 1186.53, 1203.93 for sand fraction,  
620 clay fraction and organic matter density, respectively. The scenario DA\_SM\_Par  
621 resulted in a spatial distribution of soil properties closer to the reference case, with  
622 HD values of 69.83 for sand fraction (580.83 for No\_DA), 149.96 for clay fraction  
623 (1186.53 for No\_DA) and 185.21 for organic matter density (1203.93 for No\_DA)  
624 HD values for sand fraction, clay fraction and organic matter density for the region  
625 with distance CRP- grid cell  $\leq 100$  m decreased by 80%, 82% and 64%, compared to

626 No\_DA, respectively. For the region with distance CRP- grid cell  $> 100$  m and  $\leq 200$   
627 m the decreases were 82%, 66% and 67%, respectively. Finally, for the region with a  
628 distance CRP- grid cell  $> 200$  m and  $\leq 300$  m the decreases were 78%, 82% and 40%,  
629 respectively.

630 Figure 10 shows the HD values for the comparison of the spatial distribution of  
631 irrigation amounts with the reference. The figure includes comparisons for all  
632 different data assimilation scenarios. It is clear that the assimilation of CRP neutron  
633 intensity (Only\_DA\_SM) increased the similarity between the spatial distribution of  
634 estimated annual irrigation amount and the reference irrigation distribution. Soil  
635 properties estimation and soil moisture bias estimation increased the similarity in  
636 spatial irrigation pattern further. The HD values for the scenario DA\_SM\_Bias  
637 decreased 89% (distance CRP- grid cell  $\leq 100$  m), 81% (distance CRP- grid cell  $> 100$   
638 m and  $\leq 200$  m) and 82% (distance CRP- grid cell  $> 200$  m and  $\leq 300$  m) compared to  
639 No\_DA. The HD values for the scenario DA\_SM\_Par, for the same three distance  
640 classes and in the same order, decreased by 88%, 87% and 85% compared to No\_DA.

641 [\[Insert Figure 9 here\]](#)

642 [\[Insert Figure 10 here\]](#)

643 The total annual ET for the different scenarios was also calculated, and also the  
644 contributions from ground evaporation, evaporation of intercepted water by the  
645 canopy and canopy transpiration. The ET for the reference scenario was 756.6 mm.  
646 The ET for all other scenarios was very close to the reference. In case of the scenario  
647 No\_DA, too much water was irrigated so that drought stress did not occur. The other



648 scenarios with data assimilation resulted in less irrigation, but ET was also close to  
649 the reference value, indicating that less irrigation was not associated with plant stress.  
650 For all the scenarios, the ET values did not deviate much from the reference value.  
651 This is because in all cases an overestimation of the percentage of sand led to  
652 excessive irrigation and sustainment of potential ET. The largest contribution to the  
653 ET was the canopy transpiration. The irrigated grid cells were assumed to be fully  
654 covered by the vegetation, and therefore the ground evaporation was low. The low  
655 evaporation from the canopy intercepted water maybe related to the fact that the rain  
656 events occurred mainly in the spring and winter seasons.

657

## 658 **5. Discussion**

659 The proposed data assimilation and parameter estimation (or bias estimation) can  
660 improve the soil moisture and irrigation estimation. The joint state-parameter  
661 estimation is the best scenario, and reduced the RMSE values of soil moisture content  
662 more than 50%, the spatial similarity of irrigation amount was increased and the HD  
663 values were decreased by 86% on average. The novelty of this work was the  
664 assimilation of the new CRP data in combination with irrigation scheduling. In  
665 general, classical parameter estimation tends to focus on uncertainty in the parameter  
666 estimates only, while neglecting partial or all of the other uncertainty sources (Liu and  
667 Gupta, 2007). We did not aim to compare the parameter estimation methodology with  
668 other methodologies in this study. The synthetic CRP neutron intensity observations  
669 were assimilated in CLM and the synthetic study potentially overestimated the

670 performance of the proposed method.

671 In a real-world application, the model will show systematical biases and also the  
672 implementation of the project area in the model is a strong simplification which might  
673 generate additional bias. A complication for the application of the data assimilation  
674 system in a real-world application is therefore the presence of model structural bias,  
675 and parameter estimation could compensate for this bias so that the estimated  
676 parameter values are not necessarily closer to the true parameter values. The approach  
677 will try to identify the effective parameter values that maximize model performance at  
678 that scale (Wagener et al., 2007). Therefore, as an alternative, instead of updating  
679 states and parameters jointly, also states and bias could be estimated jointly. It was  
680 shown in this paper that both approaches gave improvements. Although uncertainty of  
681 soil hydraulic parameters is important in the context of irrigation scheduling, it might  
682 be difficult to infer better estimates of soil hydraulic parameters due to other sources  
683 of uncertainty like model structural bias. On the other hand, this does not need to  
684 hamper successful operational implementation of the proposed method. Hendricks  
685 Franssen et al (2011) demonstrated the feasibility of operational prediction of  
686 groundwater levels (Franssen et al., 2011), coupled to operational optimization of  
687 groundwater management at the same site (Bauser et al., 2012). The water works  
688 Zurich applied this methodology (Franssen et al., 2011) now for the period 2009-2015,  
689 with consistent better predictions than for the open loop run. However, in the  
690 operational implementation on-line parameter estimation was avoided and only states  
691 were updated. It is therefore possible that for on-line irrigation scheduling a

692 conservative, potentially less successful strategy should be followed where only states  
693 are updated. This synthetic study showed that state updating only also would improve  
694 irrigation scheduling considerably. We believe therefore that although in a real-world  
695 case study results will be less favorable than in the synthetic study, data assimilation  
696 with updating states only, or joint updating of states and bias (or parameters) in case  
697 of a systematic bias, will improve irrigation scheduling compared to a scenario  
698 without data assimilation.

699 An additional challenge for the real world application is the forward modeling of  
700 the CRP neutron intensity. The measured CRP neutron intensity needs to be corrected  
701 for variations in the incoming high-energetic neutrons, the atmospheric pressure and  
702 humidity, lattice water and organic carbon content of the soil, and aboveground  
703 biomass. The aboveground biomass of citrus trees is temporally variable related to the  
704 growth of the oranges (or lemons) over the year. The impact of vegetation water  
705 content on the CRP neutron intensity is still under active study. In principle, an  
706 empirical methodology is suited to correct for the influence of aboveground biomass  
707 on measured neutron count intensity (Baatz et al., 2015). In this study, the  
708 synthetically measured CRP neutron intensity was applied uniformly at the CRP  
709 footprint and a simple multiscale data assimilation scheme was proposed to update the  
710 field scale CLM simulation using coarse scale CRP neutron intensity. This may not be  
711 optimal as all the grid cells within the CRP footprint contribute differently to the  
712 measured CRP neutron intensity. The soil spatial heterogeneity in the CRP footprint  
713 was introduced by adding a random spatially correlated noise. Heterogeneous land

714 cover was not considered in this study. However, the spatial variability of ecosystem  
715 parameters could be a further confounding parameter influencing the results.  
716 Altogether, accounting for temporally variable biomass in the COSMIC operator does  
717 not seem a large limitation, but spatially variable soil moisture conditions within the  
718 cosmic ray probe footprint, are a serious challenge.

719 Furthermore, the weather forecast is essential to the irrigation scheduling. This  
720 aspect was not considered in this work. If the precise precipitation forecast cannot be  
721 obtained, the irrigation requirement cannot be estimated accurately.

722 A further important complication for real-world applications is that farmers want  
723 to irrigate the citrus based on their own experience, and in combination with the low  
724 water prices they might not want to follow the suggested irrigation scheduling.  
725 Altogether, we feel that the methodology is suited for real-world applications and can  
726 improve irrigation scheduling compared to more traditional scheduling, but that the  
727 farmer participation is the most critical factor, besides model structural bias and soil  
728 moisture heterogeneity within the cosmic ray probe footprint.

729 Therefore, the successful real application of the proposed method needs: a  
730 calibrated land surface model, an improved COSMIC operator in which the measured  
731 cosmic-ray neutron intensity is corrected for above and below ground biomass, not  
732 too large spatial variability of soil moisture content within the cosmic ray probe  
733 footprint and a precise weather forecast including uncertainty characterization and  
734 participation of farmers.

735 A further possible improvement is the consideration of irrigation below ET

736 requirement, known as deficit irrigation (DI), which can reduce water demand to meet  
737 the maximum ET (Fereres and Soriano, 2007).

738

## 739 **6. Conclusions**

740 This study investigated the assimilation of synthetic measurements of coarse  
741 scale CRP neutron intensity in CLM for updating field scale root zone soil moisture  
742 content. The synthetic study mimicked a drip irrigated citrus farmland near Valencia,  
743 Spain. CLM was driven by biased soil properties and the joint estimation of soil  
744 moisture and soil properties (or soil moisture bias) was evaluated in a data  
745 assimilation framework using the state augmentation method. The non-linear  
746 measurement operator COSMIC was used to simulate the CRP neutron intensity on  
747 the basis of the soil moisture profile estimated by the CLM model. Fast neutron  
748 intensity was assimilated directly, and both soil moisture and soil properties (soil  
749 moisture bias) were updated using the LETKF in combination with the CLM model.  
750 The horizontal and vertical weights for the different CLM grid cells in the CRP  
751 footprint were also considered using a Gaussian window function.

752 The results show that assimilating CRP neutron intensity can improve joint soil  
753 moisture and soil properties estimation, and irrigation scheduling. Data assimilation  
754 schemes that remove soil moisture bias or update soil properties on the basis of CRP  
755 neutron intensity outperform data assimilation without bias or parameter estimation.  
756 The joint soil moisture and soil parameter estimation with simple multiscale  
757 assimilation strategy of CRP neutron intensity can potentially be used for irrigation

758 scheduling in the future. The main challenges for the real world application are:  
759 model calibration to remove the bias, forward modeling of cosmic-ray neutron  
760 intensity under high vegetation coverage, precise weather forecasts and cooperation of  
761 farmers.

762

### 763 **Acknowledgments**

764 This work was supported by AGADAPT (adapting water use by the agriculture sector)  
765 financed by Climate Knowledge and Innovation Community (Climate-KIC) of the  
766 European Union. AGADAPT focuses on the development and deployment of novel  
767 methods to reduce and optimize the water usage of rain-fed and irrigated agriculture  
768 by combining knowledge-based innovative technologies, modelling and transfer of  
769 technologies and innovative practices. The work was also supported by Transregional  
770 Collaborative Research Centre 32, and the NSFC project (grant number: 41271357,  
771 91125001). The support of the supercomputing facilities of Forschungszentrum Jülich  
772 (JUROPA) is gratefully acknowledged.  
773

774 **References:**

- 775 Aravéquia, J.A. et al., 2011. Evaluation of a Strategy for the Assimilation of Satellite Radiance  
776 Observations with the Local Ensemble Transform Kalman Filter. *Mon Weather Rev*, 139(6):  
777 1932-1951.
- 778 Baatz, R. et al., 2015. An empirical vegetation correction for soil water content quantification using  
779 cosmic ray probes. *Water Resour Res*, 51(4): 2030-2046.
- 780 Baek, S.J., Hunt, B.R., Kalnay, E., Ott, E., Szunyogh, I., 2006. Local ensemble Kalman filtering in the  
781 presence of model bias. *Tellus A*, 58(3): 293-306.
- 782 Ballester, C., Castel, J., Intrigliolo, D.S., Castel, J.R., 2011. Response of Clementina de Nules citrus trees  
783 to summer deficit irrigation. Yield components and fruit composition. *Agr Water Manage*,  
784 98(6): 1027-1032.
- 785 Bateni, S.M., Entekhabi, D., 2012. Surface heat flux estimation with the ensemble Kalman smoother:  
786 Joint estimation of state and parameters. *Water Resour Res*, 48.
- 787 Bauser, G. et al., 2012. A comparison study of two different control criteria for the real-time  
788 management of urban groundwater works. *J Environ Manage*, 105: 21-9.
- 789 Bogena, H.R., Huisman, J.A., Baatz, R., Hendricks Franssen, H.J., Vereecken, H., 2013. Accuracy of the  
790 cosmic-ray soil water content probe in humid forest ecosystems: The worst case scenario.  
791 *Water Resour Res*, 49(9): 5778-5791.
- 792 Bosilovich, M.G., Radakovich, J.D., da Silva, A., Todling, R., Verter, F., 2007. Skin temperature analysis  
793 and bias correction in a coupled land-atmosphere data assimilation system. *J Meteorol Soc*  
794 *Jpn*, 85A: 205-228.
- 795 Chen, Y., Zhang, D.X., 2006. Data assimilation for transient flow in geologic formations via ensemble  
796 Kalman filter. *Adv Water Resour*, 29(8): 1107-1122.
- 797 Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Water*  
798 *Resour Res*, 14(4): 601-604.
- 799 Crow, W.T. et al., 2012. Upscaling Sparse Ground-Based Soil Moisture Observations for the Validation  
800 of Coarse-Resolution Satellite Soil Moisture Products. *Rev Geophys*, 50(2): n/a-n/a.
- 801 Crow, W.T., Kustas, W.P., Prueger, J.H., 2008. Monitoring root-zone soil moisture through the  
802 assimilation of a thermal remote sensing-based soil moisture proxy into a water balance  
803 model. *Remote Sens Environ*, 112(4): 1268-1281.
- 804 De Lannoy, G.J.M., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007a. State and bias estimation for  
805 soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and  
806 frequency. *Water Resour Res*, 43(6).
- 807 De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007b. Correcting for  
808 forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resour Res*,  
809 43(9): n/a-n/a.
- 810 Dee, D.P., 2005. Bias and data assimilation. *Q J Roy Meteor Soc*, 131(613): 3323-3343.
- 811 Desilets, D., Zreda, M., Ferre, T.P.A., 2010. Nature's neutron probe: Land surface hydrology at an  
812 elusive scale with cosmic rays. *Water Resour Res*, 46(11): W11505.
- 813 Entekhabi, D. et al., 2010. The Soil Moisture Active Passive (SMAP) Mission. *P IEEE*, 98(5): 704-716.
- 814 Fereres, E., Soriano, M.A., 2007. Deficit irrigation for reducing agricultural water use. *J Exp Bot*, 58(2):  
815 147-159.
- 816 Franssen, H.J.H. et al., 2011. Operational real-time modeling with ensemble Kalman filter of variably

817 saturated subsurface flow including stream-aquifer interaction and parameter updating.  
818 Water Resour Res, 47(2): n/a-n/a.

819 Franssen, H.J.H., Kinzelbach, W., 2008. Real-time groundwater flow modeling with the Ensemble  
820 Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem.  
821 Water Resour Res, 44(9): W09408.

822 Franz, T.E., Zreda, M., Rosolem, R., Ferre, T.P.A., 2013. A universal calibration function for  
823 determination of soil moisture with cosmic-ray neutrons. Hydrology and Earth System  
824 Sciences, 17(2): 453-460.

825 Greybush, S.J., Kalnay, E., Miyoshi, T., Ide, K., Hunt, B.R., 2011. Balance and Ensemble Kalman Filter  
826 Localization Techniques. Mon Weather Rev, 139(2): 511-522.

827 Han, X. et al., 2015a. Correction of systematic model forcing bias of CLM using assimilation of  
828 cosmic-ray Neutrons and land surface temperature: a study in the Heihe Catchment, China.  
829 Hydrology and Earth System Sciences, 19(1): 615-629.

830 Han, X., Li, X., Franssen, H.J.H., Vereecken, H., Montzka, C., 2012. Spatial horizontal correlation  
831 characteristics in the land data assimilation of soil moisture. Hydrology and Earth System  
832 Sciences, 16(5): 1349-1363.

833 Han, X., Li, X., Rigon, R., Jin, R., Endrizzi, S., 2015b. Soil moisture estimation by assimilating L-band  
834 microwave brightness temperature with geostatistics and observation localization. Plos One,  
835 10(1): e0116435.

836 Han, X.J. et al., 2013. Joint Assimilation of Surface Temperature and L-Band Microwave Brightness  
837 Temperature in Land Data Assimilation. Vadose Zone J, 12(3): 0.

838 Han, X.J., Franssen, H.J.H., Montzka, C., Vereecken, H., 2014a. Soil moisture and soil properties  
839 estimation in the Community Land Model with synthetic brightness temperature  
840 observations. Water Resour Res, 50(7): 6081-6105.

841 Han, X.J., Jin, R., Li, X., Wang, S.G., 2014b. Soil Moisture Estimation Using Cosmic-Ray Soil Moisture  
842 Sensing at Heterogeneous Farmland. Ieee Geoscience and Remote Sensing Letters, 11(9):  
843 1659-1663.

844 Harris, F.J., 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. P  
845 IEEE, 66(1): 51-83.

846 Hendrick, L.D., Edge, R.D., 1966. Cosmic-Ray Neutrons near the Earth. Physical Review, 145(4):  
847 1023-1025.

848 Heng, L.K., Hsiao, T., Evett, S., Howell, T., Steduto, P., 2009. Validating the FAO AquaCrop Model for  
849 Irrigated and Water Deficient Field Maize. Agron J, 101(3): 488-498.

850 Hou, Z.S., Huang, M.Y., Leung, L.R., Lin, G., Ricciuto, D.M., 2012. Sensitivity of surface flux simulations  
851 to hydrologic parameters based on an uncertainty quantification framework applied to the  
852 Community Land Model. J Geophys Res-Atmos, 117.

853 Huang, C.L., Li, X., Lu, L., Gu, J., 2008. Experiments of one-dimensional soil moisture assimilation  
854 system based on ensemble Kalman filter. Remote Sens Environ, 112(3): 888-900.

855 Hunt, B.R., Kostelich, E.J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A  
856 local ensemble transform Kalman filter. Physica D: Nonlinear Phenomena, 230(1-2): 112-126.

857 Jimenez-Bello, M.A., Castel, J.R., Testi, L., Intrigliolo, D.S., 2015. Assessment of a Remote Sensing  
858 Energy Balance Methodology (SEBAL) Using Different Interpolation Methods to Determine  
859 Evapotranspiration in a Citrus Orchard. Ieee J-Stars, 8(4): 1465-1477.

860 Kerr, Y.H. et al., 2010. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water



861 Cycle. *IEEE*, 98(5): 666-687.

862 Kumar, S.V. et al., 2012a. Land surface Verification Toolkit (LVT) – a generalized framework for land  
863 surface model evaluation. *Geosci Model Dev*, 5(3): 869-886.

864 Kumar, S.V. et al., 2012b. A comparison of methods for a priori bias correction in soil moisture data  
865 assimilation. *Water Resour Res*, 48(3): n/a-n/a.

866 Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., Peters-Lidard, C.D., 2009. Role of Subsurface Physics  
867 in the Assimilation of Surface Soil Moisture Observations. *J Hydrometeorol*, 10(6): 1534-1547.

868 Kurtz, W., Hendricks Franssen, H.J., Kaiser, H.P., Vereecken, H., 2014. Joint assimilation of piezometric  
869 heads and groundwater temperatures for improved modeling of river-aquifer interactions.  
870 *Water Resour Res*, 50(2): 1665-1688.

871 Lal, D., Peters, B., 1967. Cosmic Ray Produced Radioactivity on the Earth. In: Sitte, K. (Ed.), *Kosmische  
872 Strahlung II / Cosmic Rays II. Handbuch der Physik / Encyclopedia of Physics*. Springer Berlin  
873 Heidelberg, pp. 551-612.

874 Levis, S., Sacks, W.J., 2011. Technical descriptions of the interactive crop management (CLM4CNcrop)  
875 and interactive irrigation models in version 4 of the Community Land Model. [Available online  
876 at <http://www.cesm.ucar.edu/models/cesm1.0/clm/index.shtml>].

877 Li, C., Ren, L., 2011. Estimation of Unsaturated Soil Hydraulic Parameters Using the Ensemble Kalman  
878 Filter. *Vadose Zone J*, 10(4): 1205.

879 Lien, G.Y., Kalnay, E., Miyoshi, T., 2013. Effective assimilation of global precipitation: simulation  
880 experiments. *Tellus A*, 65(0).

881 Liu, Y.Q., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation  
882 framework. *Water Resour Res*, 43(7): n/a-n/a.

883 Merlin, O., Walker, J.P., Panciera, R., Escorihuela, M.J., Jackson, T.J., 2009. Assessing the SMOS Soil  
884 Moisture Retrieval Parameters With High-Resolution NAFE'06 Data. *IEEE Geoscience and  
885 Remote Sensing Letters*, 6(4): 635-639.

886 Miyoshi, T., Kondo, K., Imamura, T., 2014. The 10,240-member ensemble Kalman filtering with an  
887 intermediate AGCM. *Geophys Res Lett*, 41(14): 5264-5271.

888 Miyoshi, T., Yamane, S., 2007. Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48  
889 Resolution. *Mon Weather Rev*, 135(11): 3841-3861.

890 Montzka, C. et al., 2013. Brightness Temperature and Soil Moisture Validation at Different Scales  
891 During the SMOS Validation Campaign in the Rur and Erft Catchments, Germany. *IEEE T  
892 Geosci Remote*, 51(3): 1728-1743.

893 Montzka, C. et al., 2011. Hydraulic parameter estimation by remotely-sensed top soil moisture  
894 observations with the particle filter. *Journal of Hydrology*, 399(3-4): 410-421.

895 Moradkhani, H., Hsu, K.-L., Gupta, H., Sorooshian, S., 2005a. Uncertainty assessment of hydrologic  
896 model states and parameters: Sequential data assimilation using the particle filter. *Water  
897 Resour Res*, 41(5): n/a-n/a.

898 Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005b. Dual state – parameter estimation of  
899 hydrological models using ensemble Kalman filter. *Adv Water Resour*, 28(2): 135-147.

900 Njoku, E.G., Chan, S.K., 2006. Vegetation and surface roughness effects on AMSR-E land observations.  
901 *Remote Sens Environ*, 100(2): 190-199.

902 Oleson, K. et al., 2013. Technical Description of version 4.5 of the Community Land Model (CLM). Ncar  
903 Technical Note NCAR/TN-503+STR, National Center for Atmospheric Research, Boulder, CO,  
904 422 pp.

905 Park, S.K., Xu, L., 2009. Data assimilation for atmospheric, oceanic and hydrologic applications.  
906 Springer, Berlin/Heidelberg, 730 pp.

907 Pauwels, V.R.N. et al., 2009. Optimization of Soil Hydraulic Model Parameters Using Synthetic Aperture  
908 Radar Data: An Integrated Multidisciplinary Approach. *Ieee T Geosci Remote*, 47(2): 455-467.

909 Reichle, R.H., Crow, W.T., Koster, R.D., Sharif, H.O., Mahanama, S.P.P., 2008. Contribution of soil  
910 moisture retrievals to land data assimilation products. *Geophys Res Lett*, 35(1).

911 Reichle, R.H., Kumar, S.V., Mahanama, S.P.P., Koster, R.D., Liu, Q., 2010. Assimilation of  
912 Satellite-Derived Skin Temperature Observations into Land Surface Models. *J Hydrometeorol*,  
913 11(5): 1103-1122.

914 Rosolem, R., Gupta, H.V., Shuttleworth, W.J., Zeng, X., de Gonçaves, L.G.G., 2012. A fully  
915 multiple-criteria implementation of the Sobol' method for parameter sensitivity analysis.  
916 *Journal of Geophysical Research: Atmospheres*, 117(D7): D07103.

917 Rosolem, R. et al., 2014. Translating aboveground cosmic-ray neutron intensity to high-frequency soil  
918 moisture profiles at sub-kilometer scale. *Hydrology and Earth System Sciences*, 18(11):  
919 4363-4379.

920 Rosolem, R. et al., 2013. The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray  
921 Soil Moisture Observing System. *J Hydrometeorol*, 14(5): 1659-1671.

922 Ruiz, J.J., Pulido, M., Miyoshi, T., 2013. Estimating Model Parameters with Ensemble-Based Data  
923 Assimilation: A Review. *J Meteorol Soc Jpn*, 91(2): 79-99.

924 Sampathkumar, T., Pandian, B.J., Mahimairaja, S., 2012. Soil moisture distribution and root characters  
925 as influenced by deficit irrigation through drip system in cotton-maize cropping sequence. *Agr*  
926 *Water Manage*, 103: 43-53.

927 Schöniger, A., Nowak, W., Hendricks Franssen, H.J., 2012. Parameter estimation by ensemble Kalman  
928 filters with transformed data: Approach and application to hydraulic tomography. *Water*  
929 *Resour Res*, 48(4): n/a-n/a.

930 Schwinger, J., Kollet, S.J., Hoppe, C.M., Elbern, H., 2010. Sensitivity of Latent Heat Fluxes to Initial  
931 Values and Parameters of a Land-Surface Model. *Vadose Zone J*, 9(4): 984-1001.

932 Shuttleworth, J., Rosolem, R., Zreda, M., Franz, T.E., 2013. The COsmic-ray Soil Moisture Interaction  
933 Code (COSMIC) for use in data assimilation. *Hydrology and Earth System Sciences*, 17(8):  
934 3205-3217.

935 Velikova, V. et al., 2012. The impact of winter flooding with saline water on foliar carbon uptake and  
936 the volatile fraction of leaves and fruits of lemon (*Citrus limon*) trees. *Functional Plant*  
937 *Biology*, 39(3): 199.

938 Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., Verstraten, J.M., 2005. Improved treatment of  
939 uncertainty in hydrologic modeling: Combining the strengths of global optimization and data  
940 assimilation. *Water Resour Res*, 41(1).

941 Vrugt, J.A., Gupta, H.V., Nuallain, B.O., 2006. Real-time data assimilation for operational ensemble  
942 streamflow forecasting. *J Hydrometeorol*, 7(3): 548-565.

943 Wagener, T., Sivapalan, M., Troch, P., Woods, R., 2007. Catchment Classification and Hydrologic  
944 Similarity. *Geography Compass*, 1(4): 901-931.

945 Walker, J.P., Willgoose, G.R., Kalma, J.D., 2001. One-dimensional soil moisture profile retrieval by  
946 assimilation of near-surface observations: a comparison of retrieval algorithms. *Adv Water*  
947 *Resour*, 24(6): 631-650.

948 Waller, J.A., Dance, S.L., Lawless, A.S., Nichols, N.K., Eyre, J.R., 2014. Representativity error for

949 temperature and humidity using the Met Office high-resolution model†. Q J Roy Meteor Soc,  
950 140(681): 1189-1197.

951 Wanders, N., Bierkens, M.F.P., de Jong, S.M., de Roo, A., Karssenber, D., 2014. The benefits of using  
952 remotely sensed soil moisture in parameter identification of large-scale hydrological models.  
953 Water Resour Res, 50(8): 6874-6891.

954 Welch, B.L., 1947. The Generalization of Students Problem When Several Different Population  
955 Variances Are Involved. Biometrika, 34(1-2): 28-35.

956 Wood, E.F. et al., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for  
957 monitoring Earth's terrestrial water. Water Resour Res, 47(5): W05301.

958 Wu, C.-C., Margulis, S.A., 2013. Real-Time Soil Moisture and Salinity Profile Estimation Using  
959 Assimilation of Embedded Sensor Datastreams. Vadose Zone J, 12(1): 0.

960 Zeng, X.B., 2001. Global vegetation root distribution for land modeling. J Hydrometeorol, 2(5):  
961 525-530.

962 Zreda, M., Desilets, D., Ferre, T.P.A., Scott, R.L., 2008. Measuring soil moisture content non-invasively  
963 at intermediate spatial scale using cosmic-ray neutrons. Geophys Res Lett, 35(21): L21402.

964 Zreda, M. et al., 2012. COSMOS: the COsmic-ray Soil Moisture Observing System. Hydrology and Earth  
965 System Sciences, 16(11): 4079-4099.

966

967

968 **Table List**

969 Table 1 Summary of perturbation parameters for atmospheric forcing data

970

971

Table 1 Summary of perturbation parameters for atmospheric forcing data

<b>Variables</b>	<b>Noise</b>	<b>Standard deviation</b>	<b>Time Correlation scale</b>	<b>Forcing Cross Correlation</b>
Precipitation	Multiplicative	0.5	24 h	[ 1.0,-0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.3	24 h	-0.8, 1.0,-0.5, 0.4,
Longwave radiation	Additive	20 W/m <sup>2</sup>	24 h	0.5, -0.5, 1.0, 0.4,
Air temperature	Additive	1 K	24 h	0.0, 0.4, 0.4, 1.0]

972

973

974 **Figure List**

975 Figure 1. Schematic overview of the different steps of the irrigation optimization  
976 procedure

977 Figure 2. Soil moisture content at 30 cm (upper graph) and 50 cm (lower graph) depth  
978 at the cosmic-ray location for the different simulation scenarios

979 Figure 3. RMSE values for soil moisture content at 30 cm depth (left graph) and 50  
980 cm depth (right graph) for the different scenarios at the CRP location

981 Figure 4. Temporal evolution (collected every three days) of soil moisture bias for the  
982 first soil layer at the CRP location (scenario DA\_SM\_Bias), the true soil moisture bias  
983 was calculated from scenario No\_DA is showed in blue

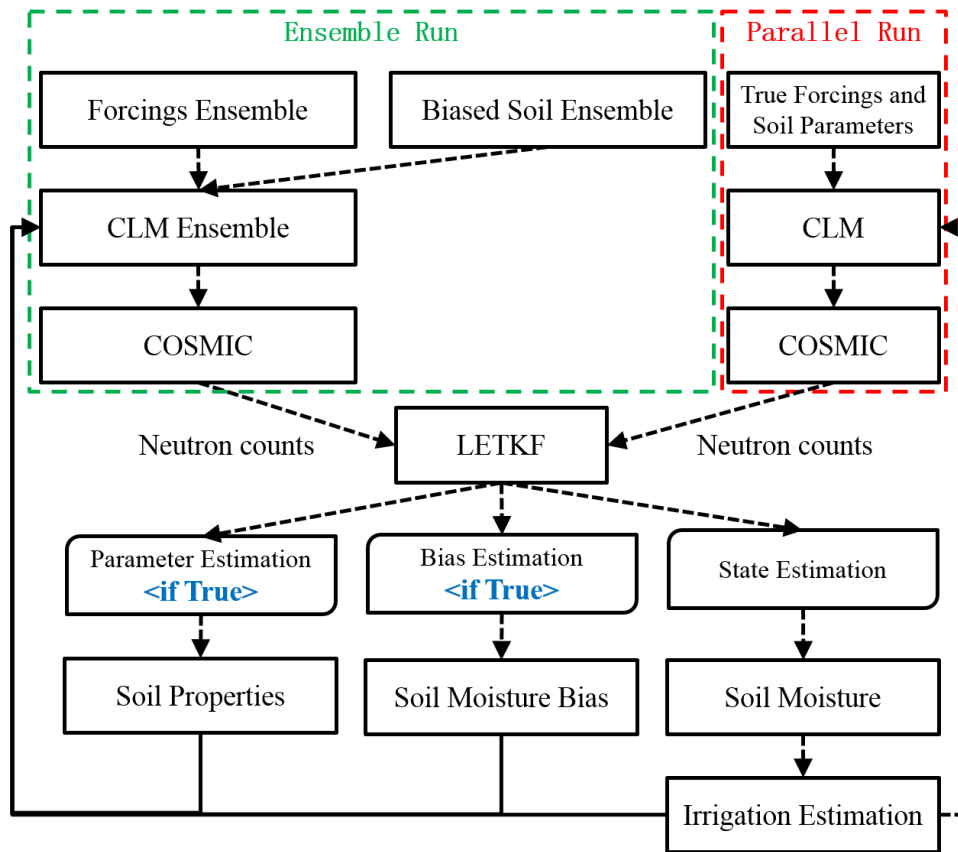
984 Figure 5. Temporal evolution (collected every three days) of saturated hydraulic  
985 conductivity  $K$  of soil ( $K_{10cm}$ ) and the empirical parameter  $B$  of the Clapp–  
986 Hornberger parameterization ( $B_{10cm}$ ) at the CRP location for the scenario  
987 DA\_SM\_Par

988 Figure 6. Irrigation requirement as function of time at the CRP location for the  
989 different scenarios; t-test statistics (p-value) with significance level 0.05 for  
990 comparing the calculated irrigation distribution with reference irrigation are also  
991 shown for the different scenarios (large p-values indicate high similarity)

992 Figure 7. Annual irrigation requirement according to the different scenarios at the  
993 CRP location

994 Figure 8. Annual irrigation calculated for different simulations scenarios and  
995 compared to the reference scenario

996 Figure 9. Hausdorff distance values of calculated annual irrigation requirement,  
997 compared to reference irrigation, for different scenarios. Results are plotted as  
998 function of distance between model grid cells and CRP location  
999



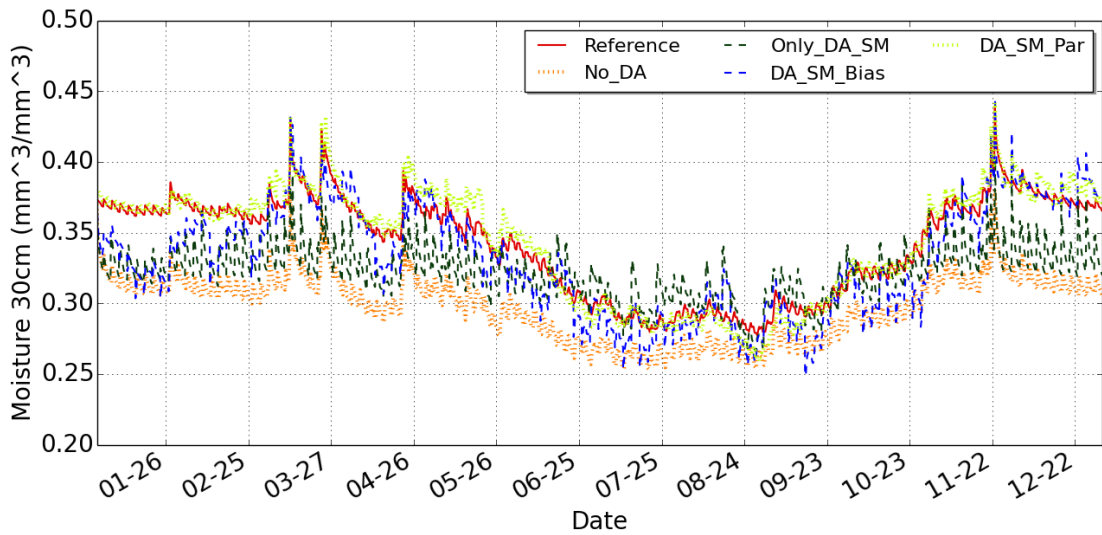
1000

1001

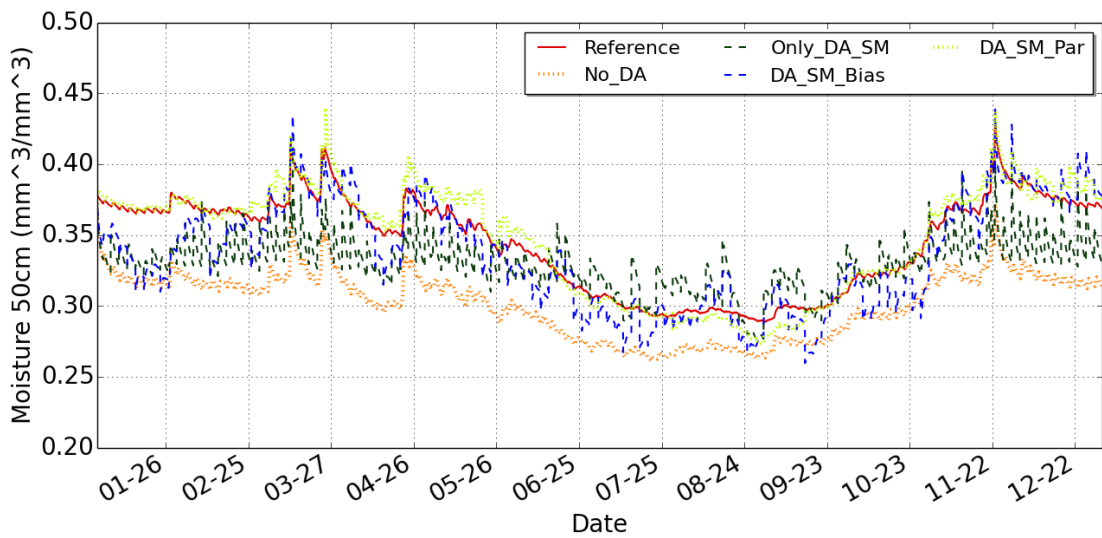
1002

1003

Figure 1. Schematic overview of the different steps of the irrigation optimization procedure



1004



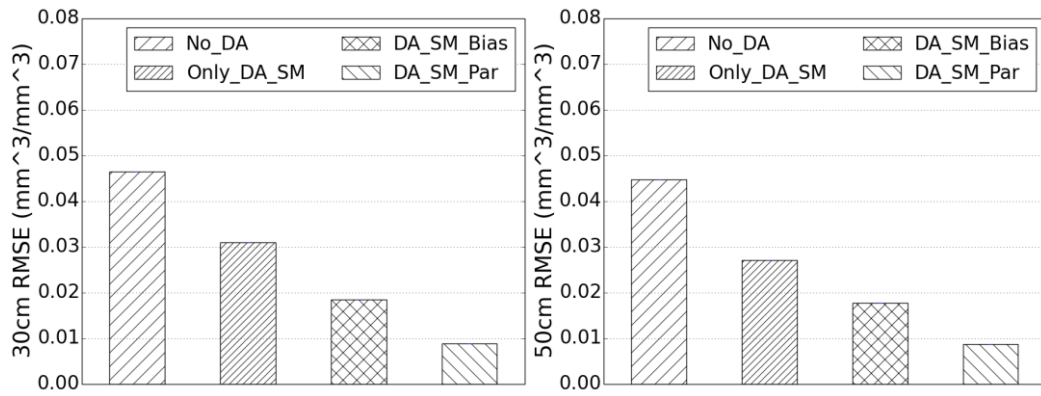
1005

1006 Figure 2. Soil moisture content at 30 cm (upper graph) and 50 cm (lower graph)

1007 depth at the CRP location for the different simulation scenarios

1008





1009

1010

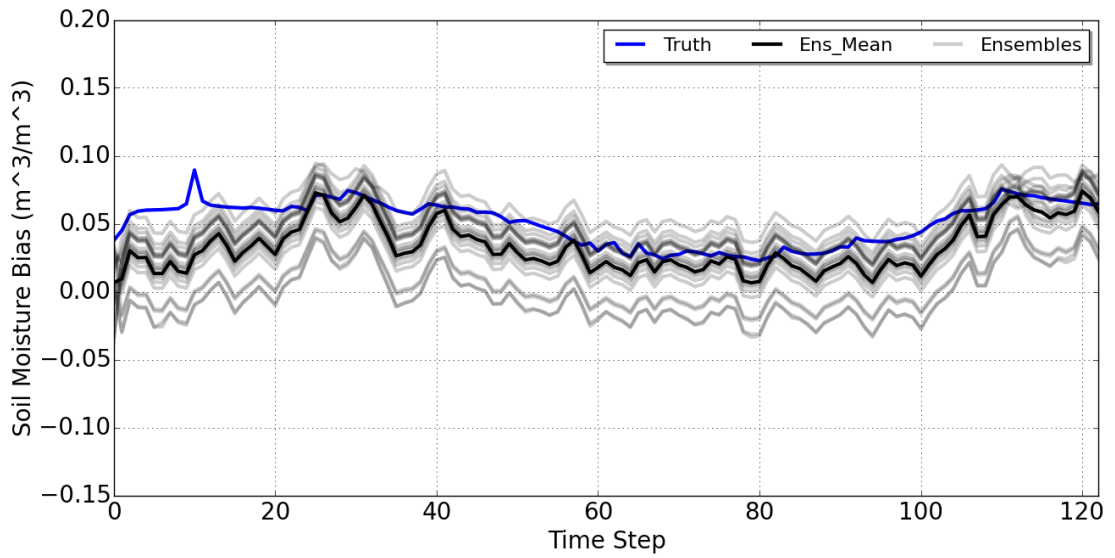
Figure 3. RMSE values for soil moisture content at 30 cm depth (left graph) and 50 cm depth (right graph) for the different scenarios at the CRP location

1011

1012

1013

1014

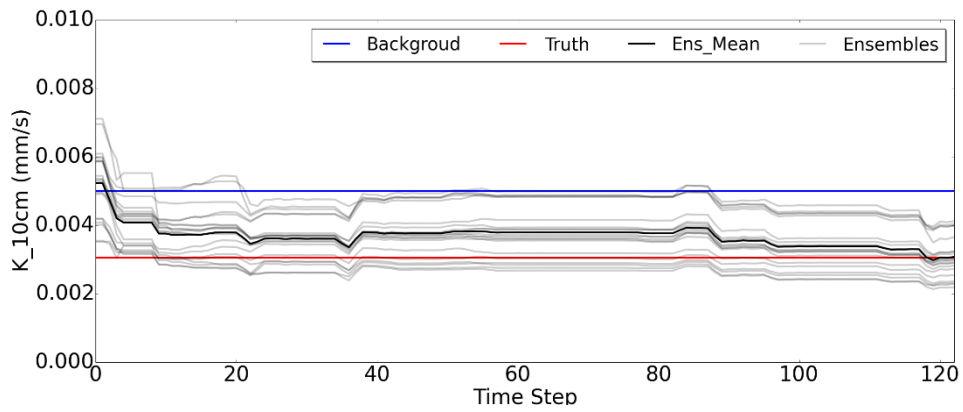


1015

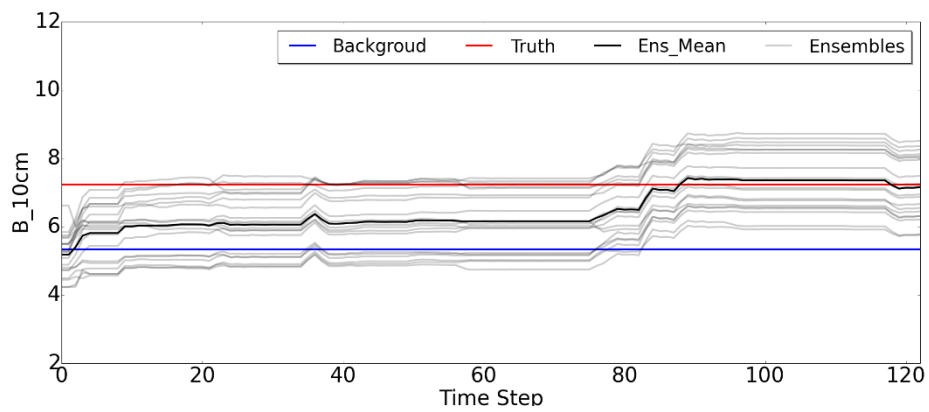
1016 Figure 4. Temporal evolution (collected every three days) of soil moisture bias for the  
1017 first soil layer at the CRP location (scenario DA\_SM\_Bias). The true soil moisture  
1018 bias was calculated from the scenario No\_DA and is shown in blue. The unit of x-axis  
1019 is for time steps of 3 days.

1020

1021



1022



1023

1024

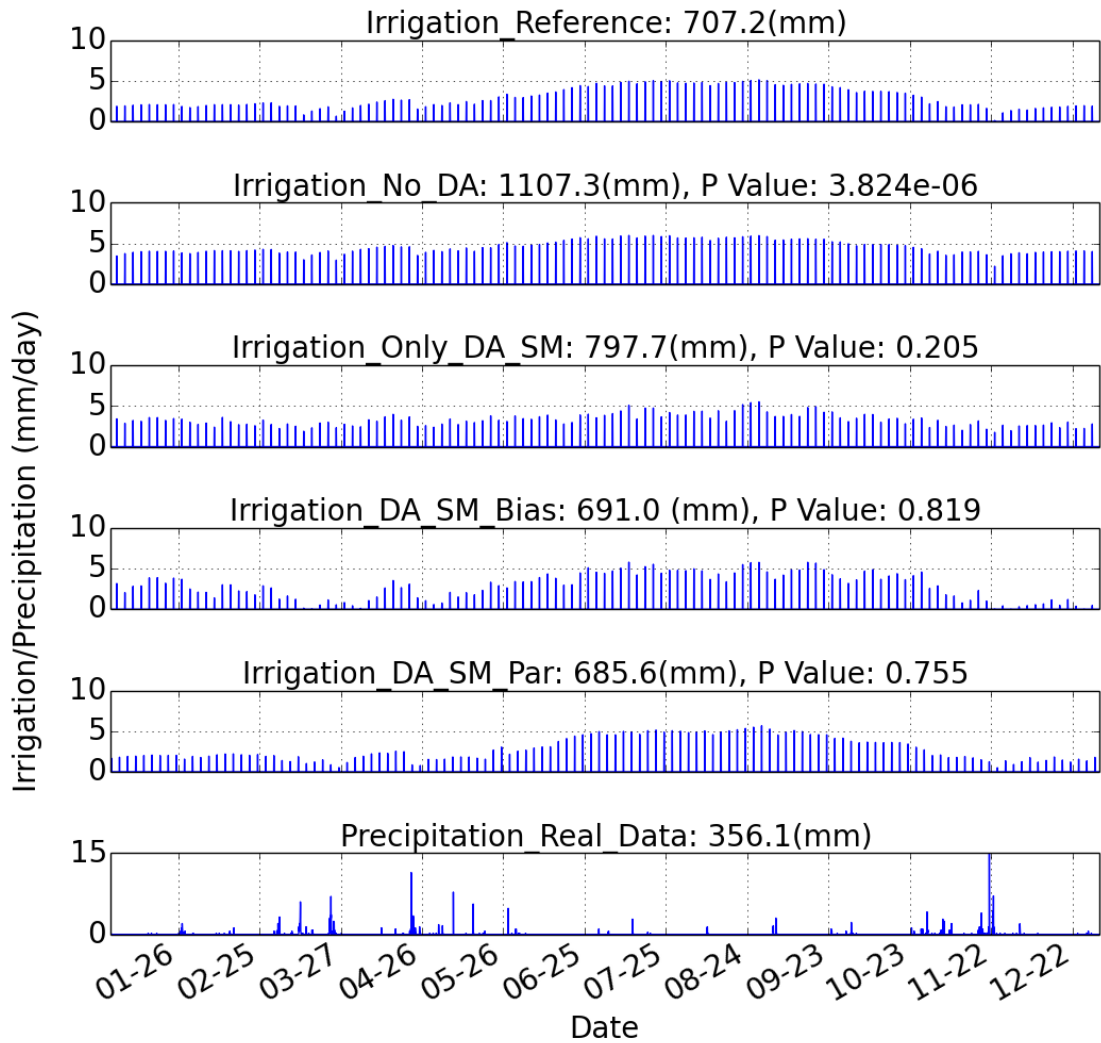
1025

1026

1027

1028

Figure 5. Temporal evolution (collected every three days) of saturated hydraulic conductivity  $K$  of soil ( $K_{10\text{cm}}$ ) and the empirical parameter  $B$  of the Clapp–Hornberger parameterization ( $B_{10\text{cm}}$ ) at the CRP location for the scenario DA\_SM\_Par. The unit of x-axis is for time steps of 3 days.



1029

1030

1031

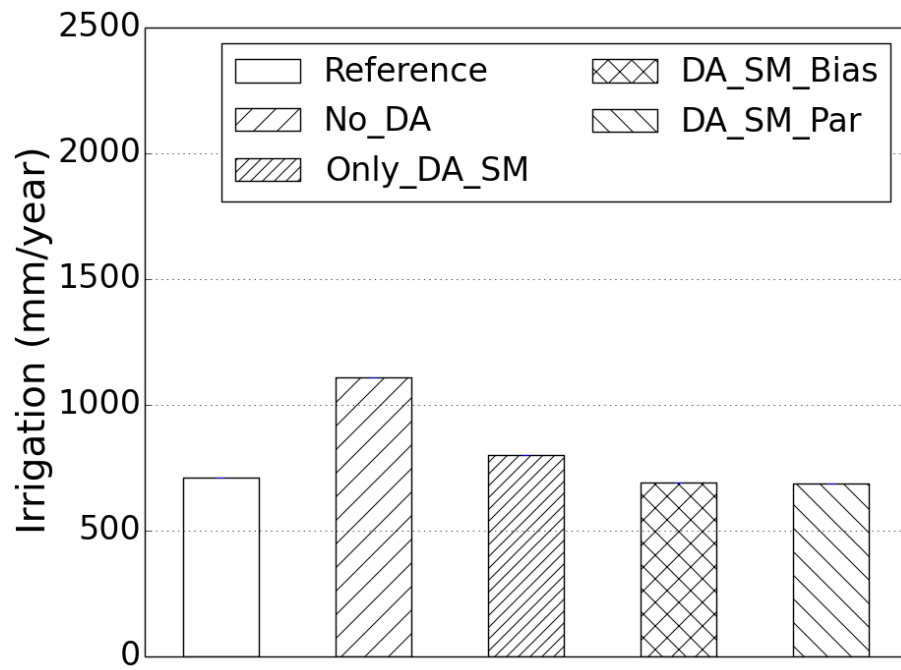
1032

1033

1034

Figure 6. Irrigation requirement as function of time at the CRP location for the different scenarios; t-test statistics (p-value) with significance level 0.05 for comparing the calculated irrigation distribution with reference irrigation are also shown for the different scenarios (large p-values indicate high similarity)

1035



1036

1037 Figure 7. Annual irrigation requirement according to the different scenarios at the

1038

CRP location

1039

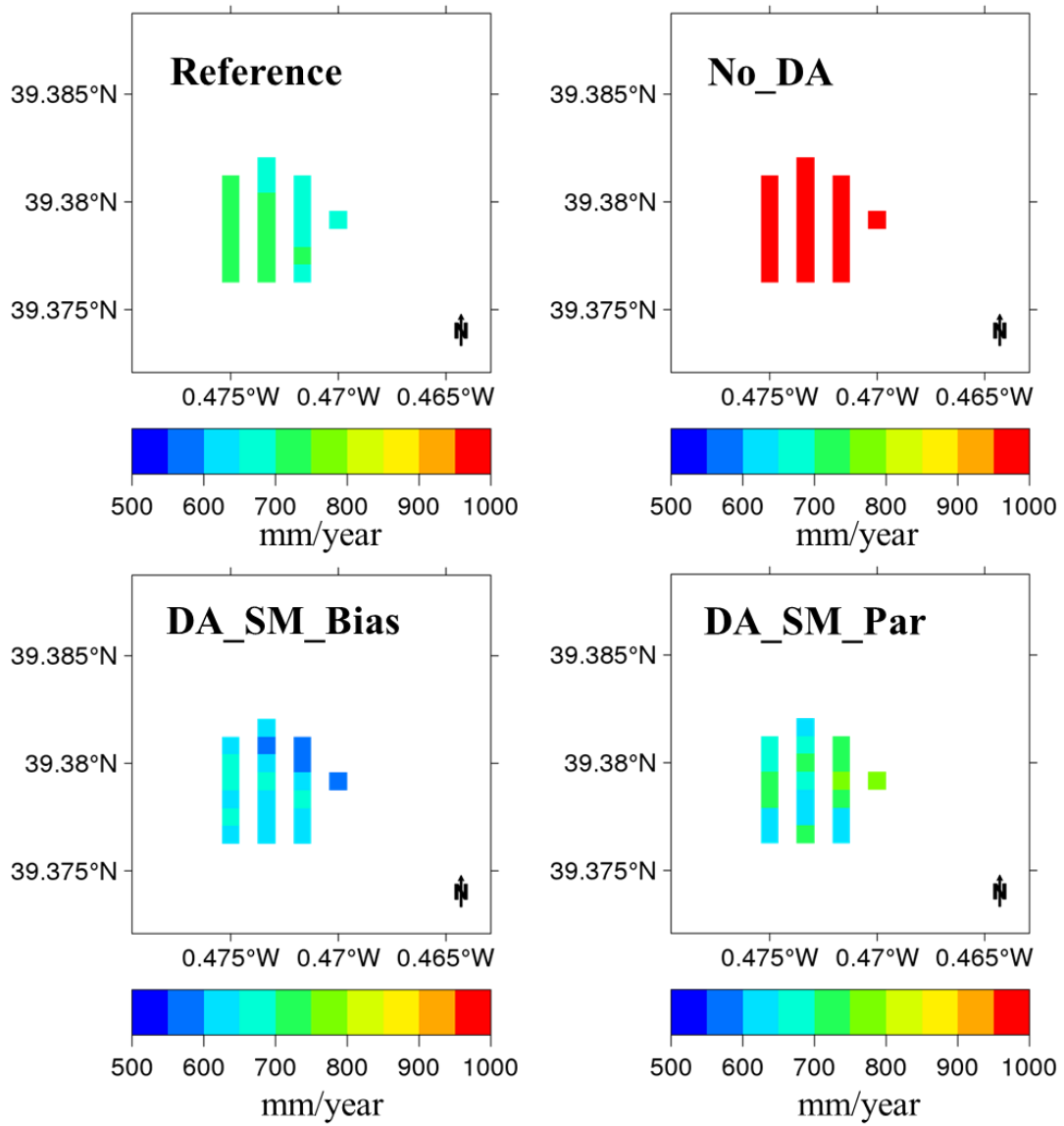
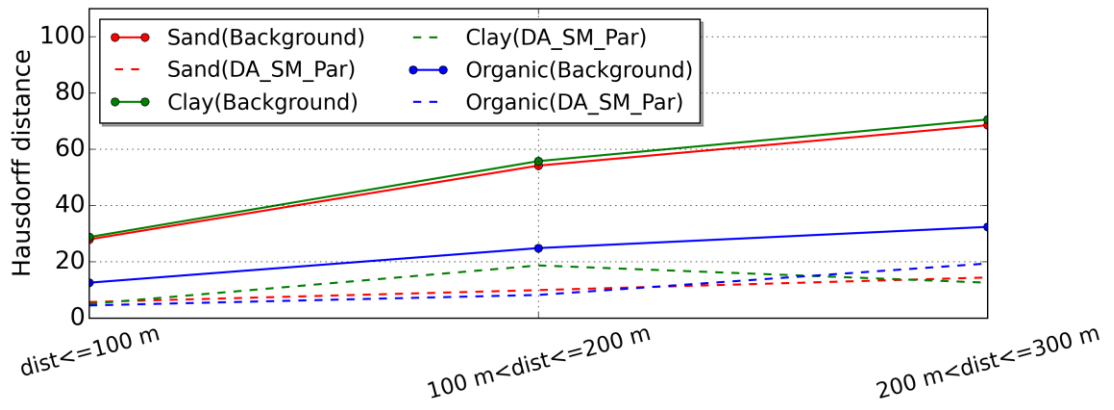


Figure 8. Annual irrigation calculated for different simulations scenarios and

compared to the reference scenario



1045

1046 Figure 9. Hausdorff distance values for background soil properties (sand fraction, clay  
 1047 fraction and organic matter density) and estimated soil properties (scenario

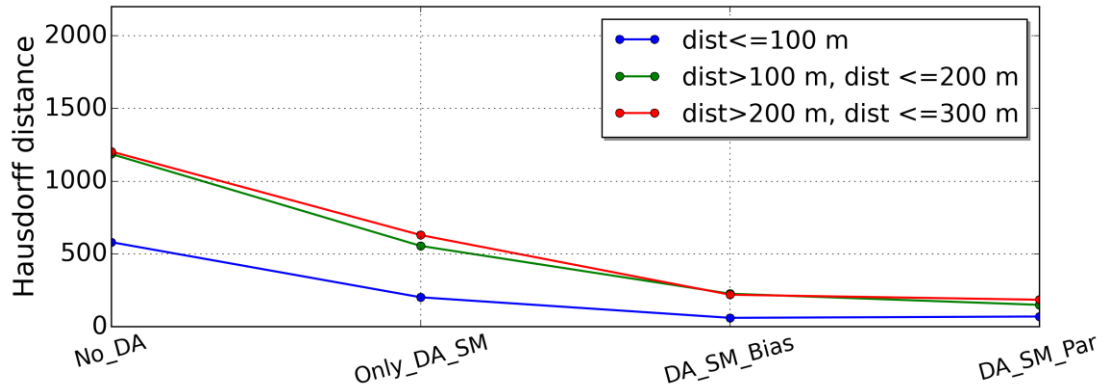
1048 DA\_SM\_Par). Results are plotted as function of the distance between model grid cells

1049 and the CRP location

1050

1051

1052



1053

1054

Figure 10. Hausdorff distance values of calculated annual irrigation requirement,

1055

compared to reference irrigation, for different scenarios. Results are plotted as

1056

function of distance between model grid cells and CRP location

1057

1058