Simultaneous soil moisture and properties estimation for a drip irrigated field by assimilating cosmic-ray neutron intensity

Xujun Han, Harrie-Jan Hendricks Franssen, Rafael Rosolem, Heye Bogena, Fernando Martínez Alzamora, Andre Chanzy, Harry Vereecken

HAL Id: hal-02631847
https://hal.inrae.fr/hal-02631847v1

Submitted on 27 May 2020

Accepted Manuscript

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

Xujun Han, Harrie-Jan Hendricks Franssen, Miguel Ángel Jiménez Bello, Rafael Rosolem, Heye Bogena, Fernando Martínez Alzamora, André Chanzy, Harry Vereecken

Please cite this article as: Han, X., Franssen, H.H., Bello, M.A.J., Rosolem, R., Bogena, H., Alzamora, F.M., Chanzy, A., Vereecken, H., Simultaneous Soil Moisture and Properties Estimation for a Drip Irrigated Field by Assimilating Cosmic-ray Neutron Intensity, *Journal of Hydrology* (2016), doi: http://dx.doi.org/10.1016/j.jhydrol.2016.05.050

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

Xujun Han[1,2], Harrie-Jan Hendricks Franssen[1,2], Miguel Ángel Jiménez Bello[3], Rafael Rosolem[4], Heye Bogena[1], Fernando Martínez Alzamora[3], André Chanzy[5], Harry Vereecken[1,2]
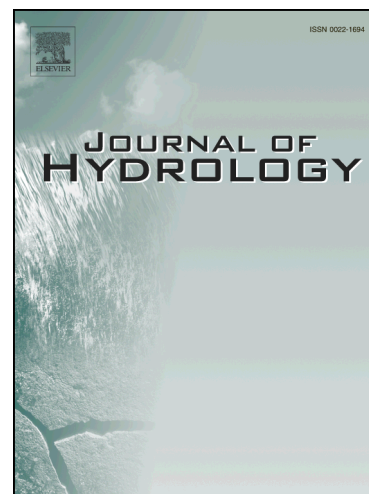
1. Forschungszentrum Jülich, Agrosphere (IBG 3), Leo-Brandt-Strasse, 52425 Jülich, Germany
2. Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Geoverbund ABC/J, Leo-Brandt-Strasse, 52425 Jülich, Germany
3. Institute of Water and Environmental Engineering (IIAMA). Universitat Politècnica de Valencia. Camino de Vera, s/n, 46022 Valencia, Spain
4. Department of Civil Engineering, University of Bristol, Bristol BS8 1TR, UK
5. INRA, UMR1114, Environnement Méditerranéen et Modélisation des AgroHydrosystèmes, Domaine Saint Paul, Site Agroparc, F-84914 Avignon, France

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

Abstract

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

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

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

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

1. Introduction

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

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

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

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

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

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

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

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

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

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

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

In Han et al., 2015a, we studied the joint updating of soil moisture, soil temperature and leaf area index by assimilating CRP neutron intensity and land surface temperature. In this study however, we considered in addition the joint updating of soil moisture and soil properties, or soil moisture and soil moisture bias, and the vertical and horizontal weighting for updating soil moisture in the footprint of a CRP. This implies that in this work states and parameters for many model grid cells in the CRP footprint are updated with a single CRP neutron intensity observation. This is therefore a small multiscale data assimilation experiment with the irrigation scheduling as one of the objectives.

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

2. Methodology

The main components of the methodology are: (i) measurement of above-ground neutron intensity, which is linked to field scale soil moisture content by a measurement operator (section 2.1) and horizontal weights (section 2.3); (ii) the land surface model CLM (version 4.5) which simulates the transport of water and energy in the soil-plant-atmosphere continuum (section 2.2); (iii) data assimilation according to the Local Ensemble Transform Kalman Filter (LETKF) methodology (Hunt et al., 2007) which optimally combines measurements and model predictions to update soil moisture (and possibly soil properties or soil moisture bias), taking into account uncertain atmospheric forcing and model parameters (or model bias) (section 2.3) and (iv) an optimization routine which calculates irrigation need for the ensemble of soil moisture forecasts (section 2.4).

2.1. Cosmic ray Soil Moisture Interaction Code (COSMIC)

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

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

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

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

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

where N (counts/h) is the number of high-energy neutrons at the soil surface, z is the soil layer depth (m), ρ_s is the dry soil bulk density (g cm^{-3}), ρ_w is the total soil water density, including the lattice water (g cm^{-3}); $m_s(z)$ and $m_w(z)$ are the integrated mass per unit area of dry soil and water (g cm^{-2}), φ is the angle between the vertical below the detector and the line between the detector and each point in the plane (Shuttleworth et al., 2013), L_1 is the high energy soil attenuation length with value of 162.0 gcm^{-2}, L_2 is the high energy water attenuation length of 129.1 gcm^{-2}, L_3 is the fast neutron soil attenuation length (gcm^{-2}) and L_4 is the fast neutron water attenuation length with value of 3.16 gcm^{-2} (Shuttleworth et al., 2013).

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

2.2. Community Land Model - CLM

The Community Land Model (CLM - version 4.5) developed by the National Center for Atmospheric Research (NCAR) was used to calculate soil moisture content and evapotranspiration (Oleson et al., 2013). CLM uses the simplified Richards equation to model water flow in the unsaturated zone and calculation of land surface energy fluxes is done by invoking the Monin-Obukhov similarity theory. In addition, the following processes can be simulated by CLM: transfer of solar radiation and longwave radiation, stomatal physiology and photosynthesis, crop dynamics and irrigation (Oleson et al., 2013). The land cover can be represented by 17 plant functional types (PFTs) and the calculation of energy fluxes is based on the PFTs. Hydraulic and thermal parameters in CLM are derived based on soil properties such as sand and clay fraction and organic matter density (Oleson et al., 2013).

2.3. Data Assimilation and Parameter Estimation

The Local Ensemble Transform Kalman Filter (LETKF) is a square root ensemble Kalman filter (Hunt et al., 2007; Miyoshi and Yamane, 2007), which is applied extensively in atmospheric data assimilation studies (Aravéquia et al., 2011; Baek et al., 2006; Lien et al., 2013; Miyoshi et al., 2014) and also in land data assimilation studies (Han et al., 2014a). In LETKF, the uncertainty of the model forecast is represented by ensemble members. In this study LETKF is used to estimate both soil moisture and soil properties with the state augmentation method (Bateni and

Entekhabi, 2012; Han et al., 2014a; Li and Ren, 2011) or to update soil moisture and soil moisture bias jointly.

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

$$\mathbf{X}^b = \left[\mathbf{x}_1^b - \overline{\mathbf{x}}^b, \ldots, \mathbf{x}_M^b - \overline{\mathbf{x}}^b \right] \tag{5}$$

$$\mathbf{y}_i^b = H\left(\mathbf{x}_i^b \right) \tag{6}$$

$$\mathbf{Y}^b = \left[\mathbf{y}_1^b - \overline{\mathbf{y}}^b, \ldots, \mathbf{y}_M^b - \overline{\mathbf{y}}^b \right] \tag{7}$$

where $\mathbf{x}_1^b, \ldots \mathbf{x}_M^b$ are vectors with the ensemble members, M is the ensemble size, $\overline{\mathbf{x}}^b$ is the vector with ensemble means calculated over $\mathbf{x}_1^b, \ldots \mathbf{x}_M^b$, H is the observation operator (i.e., COSMIC for soil moisture), $\mathbf{y}_i^b$ is the mapping of the ensemble members $\mathbf{x}_1^b, \ldots \mathbf{x}_M^b$ to the measurement space and $\overline{\mathbf{y}}^b$ is the vector of ensemble means of $\mathbf{y}_1^b, \ldots \mathbf{y}_M^b$. The vector $\overline{\mathbf{x}}^b$ contains i) the depth weighted average soil moisture θ_{cosmic}, which was derived from COSMIC and considers the contribution of different soil layers, ii) the soil moisture of 10 layers ($\theta_1 \ldots \theta_{10}$) and iii) soil properties (sand fraction, clay fraction and organic matter density) in case soil properties are estimated. In case of bias estimation, the vector $\overline{\mathbf{x}}^b$ contains soil moisture bias instead of soil properties. The dimensions of the augmented state vector $\overline{\mathbf{x}}^b$ were 11 for the state estimation only, 14 for joint state and parameter estimation, and 12 for joint state and bias estimation. Only soil properties and soil moisture bias for the upper soil layer were included in the state vector $\overline{\mathbf{x}}^b$. Soil properties for all 10 layers were updated based on the ratio of the soil properties between the upper soil layer and the lower soil layers (Han et al., 2014b). The soil moisture bias of deeper

layers was assumed to decrease exponentially according to equation (8).

$$
\overline{\mathbf{x}}^b = \left\{
\begin{array}{ll}
\begin{bmatrix} \theta_{cosmic} \\ \theta_1 \\ \vdots \\ \theta_{10} \end{bmatrix} & state\ estimation \\[2em]
\begin{bmatrix} \theta_{cosmic} \\ \theta_1 \\ \vdots \\ \theta_{10} \\ Sand \\ Clay \\ Organic \end{bmatrix} & parameter\ estimation \\[2em]
\begin{bmatrix} \theta_{cosmic} \\ \theta_1 + \theta_{Bias}^{k-1} * exp\left(-1.0 * Z_1\right) \\ \vdots \\ \theta_{10} + \theta_{Bias}^{k-1} * exp\left(-1.0 * Z_{10}\right) \\ \theta_{Bias}^{k-1} \end{bmatrix} & bias\ estimation
\end{array}
\right\}
\tag{8}
$$

where Z_i (m) is the ith soil layer thickness of CLM and k is the time step.

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

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

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

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

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

$$
\overline{w}^a = \mathbf{P}^a \mathbf{Y}^{bT}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^b\right) \tag{11}
$$

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

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

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a + \bar{\mathbf{x}}^b \tag{12}$$

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

The forecast model of soil moisture bias $\boldsymbol{\theta}_{Bias}^k$ was defined as:

$$\boldsymbol{\theta}_{Bias}^k = \boldsymbol{\theta}_{Bias}^{k-1} \tag{13}$$

where k is the time step.

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

considered the contributions from all model grid cells contained within the CRP footprint. The size of the Gaussian window function was chosen as the diameter of the CRP footprint (600 m). The observation operator H (needed by equation (6)) was defined as the combination of a Gaussian window function and the COSMIC operator. Using equation (6) we get:

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

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

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

horizontally and vertically) will be updated according to the correlation between the simulated values by CLM and θ_{cosmic}:

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

2.4. Irrigation Requirement

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

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

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

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

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

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

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

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

$$
RF_i = \begin{cases} 0.5 \begin{bmatrix} \exp\left(-r_a Z_{h,i-1}\right) + \exp\left(-r_b Z_{h,i-1}\right) \\ -\exp\left(-r_a Z_{h,i}\right) + \exp\left(-r_b Z_{h,i}\right) \end{bmatrix} & 1 \le i < N_{levsoi} \\ 0.5 \begin{bmatrix} \exp\left(-r_a Z_{h,i-1}\right) + \exp\left(-r_b Z_{h,i-1}\right) \end{bmatrix} & i = N_{levsoi} \end{cases} \tag{18}
$$

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.

2.5. Performance Measures

In order to evaluate the results of data assimilation and irrigation, the Root Mean

Square Error (RMSE) was calculated for the hourly soil moisture results:

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

where θ_{est} is the hourly soil moisture ensemble mean for a given scenario, θ_{ref}

is the hourly soil moisture value of the reference scenario. K is equal to 8760. Soil

moisture contents at the time points of assimilation and prediction were included in

the calculations with Eq. 19. Lower RMSE values mean better performance.

The Hausdorff Distance (HD) is a quantitative measure of the similarity between

two spatial distributions (Kumar et al., 2012a). Lower HD values mean higher spatial

similarity. HD is defined as the maximum distance of a set to the nearest point in the

other set:

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

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

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

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

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{N_1} + \dfrac{s_2^2}{N_2}}} \tag{21}$$

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 optimized irrigation amounts, respectively. The subscripts 1 and 2 represent the reference scenario and estimation scenario, respectively.

3. Synthetic Experiment

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

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

3.1. Design of Synthetic Experiments

The citrus tree was modeled as a broadleaf evergreen tropical PFT in CLM. In order to mimic the planting pattern of citrus trees at the Picassent site, odd CLM grid columns were modelled as bare ground without vegetation cover while even CLM columns were modelled as fully covered by broadleaf evergreen tropical trees. The mimicked ground cover of vegetation was similar to the real measured ground cover fraction of citrus trees. Soil properties (sand fraction and clay fraction) and organic matter density determine the soil thermal and hydraulic properties in CLM. Soil samples from the Picassent region were taken and a soil sand fraction of 32%, clay fraction of 33%, bulk density of 1.5 gcm^{-3}, and organic fraction of 1.2% (10 cm depth) and 0.7% (50 cm depth) were determined using the Bouyoucos method, which is based on stokes law and involves use of a hydrometer (Bouyoucos, 1962) and applied uniformly in space in CLM. Spatially correlated noise was added to the uniform soil properties to represent the spatial heterogeneity. The noise was simulated by sequential Gaussian simulation with a correlation range of 100 m, variance of 100.0 (%) and mean value of 0.0 (%). The noise was constrained within the range of (-5.0%, 5.0%) after the sequential Gaussian simulation. The atmospheric forcing data measured by the Picassent weather station were used as CLM input. The maximum rate of carboxylation at 25 °C (V_{cmax25}) controls the maximum rate of carboxylation and canopy transpiration in CLM. The default value of V_{cmax25} is 55 $\mu molm^2s^{-1}$ for

broadleaf evergreen tropical tree in CLM. It was changed to 100 μmolm^{2}s^{-1} according

to a reference value for citrus trees (Velikova et al., 2012). The study area was

discretized in 40×60 grid cells at a spatial resolution of 100 m. However, our

analysis will focus on the CRP footprint which contains only 43 CLM grid cells in

total. Because of the spatial discretization, it is not easy to describe the exact CRP

footprint. The diameter of the CRP footprint for the horizontal weighting calculation

was chosen as the ratio between the CRP diameter and the spatial resolution of the

CLM grid cell, which was 7 grid cells.

A model spin-up was made to obtain reasonable initial conditions for CLM. CLM

was run for the spin-up period from 2010-01-01 to 2010-12-31 using an hourly time

step. Next, a one-year (2011-01-01 to 2011-12-31) irrigation estimation using the true

soil properties was performed in CLM and the irrigation requirement was calculated

every three days during the period from 2011-01-01 to 2011-12-31 and subsequently

applied. This reference model run also supplied the reference soil moisture

distribution in space and time and the canopy transpiration, and a distribution in space

and time (and total amounts) of irrigation.

3.2. Ensemble Generation

The simulation experiments evaluated how well the soil moisture,

evapotranspiration, soil properties and irrigation requirement could be characterized if

soil properties and model forcing were biased and/or uncertain, but the measurement

data (albeit synthetically generated for this study) in the form of CRP neutron

intensity were available. We assumed that the soil texture was systematically coarser than in reality. The soil properties used in the reference run were perturbed to represent these conditions: the sand fraction was multiplied by 1.5 while the clay fraction and organic matter density were both multiplied by 0.75. In addition, for all scenarios the sand and clay fractions, and organic matter density were perturbed by a uniform distributed noise in the range of [-10.0%, 10.0%] (for soil texture) and [-10.0 km m^{-3}, 10.0 km m^{-3}] (for organic matter density) to generate 20 different soil hydraulic properties for the 20 different ensemble members. It is assumed that these perturbations represent a realistic representation of the uncertainty in practice. If the sum of the sand and clay fraction was larger than 98%, an amount equal to ((Sand%+Clay%) − 98%)/2.0 was subtracted from both the sand and the clay fraction. For the ensembles for which initial soil moisture bias was updated instead of soil properties, soil moisture was perturbed with a spatially uniform value sampled from the uniform distribution with values between -0.04 m^{3}m^{-3} and 0.04 m^{3}m^{-3}.

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

3.3. Synthetic Observation

In this study, the "observed" synthetic CRP neutron intensity was generated using the COSMIC model, which used the soil moisture profile simulated by CLM as input. A point which had to be considered in the simulation scenarios was that the soil moisture changed also as response to the irrigation amount applied in the CLM simulation. Therefore, the synthetic observations will be different for each simulation scenario due to the different irrigation amounts. In order to obtain the synthetic measurements of CRP neutron intensity for these scenario runs, a parallel run was made with the same soil properties and atmospheric forcing as the reference run for each of the simulation scenarios. However, the irrigation amount for this parallel run was the same as the optimized one for the specific simulation scenario. Therefore irrigation amounts differed among the parallel runs of the different scenarios. This in turn also affected the CRP intensity which was assimilated in the simulation scenarios. The soil moisture calculated in the parallel run was used now as the synthetic soil moisture measurement for this specific scenario and also used as input for COSMIC to estimate the synthetic CRP neutron intensity. This implies that the assimilated CRP neutron intensity differed among different scenarios, in correspondence with different irrigation amounts applied in the different scenarios.

In this synthetic study, N_{COSMOS} was set to 150 counts h^{-1} and lattice water was 3% in COSMIC. The simulated CRP intensity measurements were perturbed in order to represent the observation error. This perturbation had a mean equal to zero and a variance equal to the square root of the neutron intensity to represent the observation error. The CLM grid cells located in the footprint of the CRP were updated by the data

assimilation procedure. The neutron intensity observation was assimilated at 23:00 each three days, prior to irrigation scheduling.

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

[Insert Figure 1 here]

3.4. Irrigation Scheduling

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

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

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

(2) Irrigation estimation with biased soil properties, without CRP neutron intensity assimilation (No_DA, i.e. Open loop).

(3) Irrigation estimation with biased soil properties, with the assimilation of CRP neutron intensity (every three days) (Only_DA_SM).

(4) Irrigation estimation with biased soil properties, with both CRP neutron intensity assimilation (every three days) and soil moisture bias estimation (DA_SM_Bias).

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

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

4. Results

In this section we evaluate time series for the different simulation scenarios at the CRP location. Spatial patterns of (estimated) soil properties and optimized irrigation amounts for different simulation scenarios are also compared. This comparison is made at the scale of the complete CRP footprint. The temporal evolution of soil moisture content at 30 cm and 50 cm depth for different simulation scenarios is shown in Fig. 2. The scenario No_DA underestimated soil moisture content even although (too) high irrigation amounts were scheduled, which is related to the erroneous soil texture in this simulation scenario and the associated bias in soil properties like saturated hydraulic conductivity. In the data assimilation scenario Only_DA_SM soil moisture characterization was improved by assimilating the CRP neutron intensity, but the bias was also very high due to the large bias of soil properties. In order to reduce the soil moisture bias, the joint soil moisture state and bias estimation was evaluated in the scenario DA_SM_Bias by use of the state augmentation method. The soil moisture estimation became better than Only_DA_SM after the soil moisture bias was reduced. The best soil moisture results were obtained in the scenario DA_SM_Par, in which sand fraction, clay fraction and organic matter density were updated sequentially using the joint state and parameter estimation method.

[Insert Figure 2 here]

The RMSE values for soil moisture characterization are summarized in Fig. 3. Compared with the Reference scenario, the scenarios No_DA and Only_DA_SM underestimated soil moisture content, which resulted in higher irrigation requirement than for the reference case. Compared with the scenario No_DA, the RMSE values

for soil moisture content at 30 cm depth decreased by 33%, 40% and 52% for the scenarios Only_DA_SM, DA_SM_Bias and DA_SM_Par, respectively. At 50cm depth these RMSE-decreases were 39%, 35% and 51% for the scenarios Only_DA_SM, DA_SM_Bias and DA_SM_Par, respectively. These results illustrate the benefit of joint state-parameter estimation. Model results were strongly influenced by the biased soil properties for the scenario Only_DA_SM, where measurement data were not used to estimate model bias or soil properties. As model bias was related to biased soil properties in these simulations, joint state and parameter estimation performed better than the scenario with joint state and bias estimation.

[Insert Figure 3 here]

Fig. 4 shows the soil moisture bias for the different ensemble members (scenario DA_SM_Bias). The temporal evolution of the bias is shown for the first soil layer and the CRP location. The time series of true soil moisture bias (the scenario No_DA) is shown for comparison. The mean bias value for No_DA is 0.051 m^3m^{-3} and 0.033 m^3m^{-3} for DA_SM_Bias. The joint state-bias estimation could reduce 35% of the soil moisture bias introduced by the biased soil properties.

[Insert Figure 4 here]

The CLM model derives the soil hydraulic parameters using a predefined pedotransfer function (Oleson et al., 2013). The updated soil texture resulted therefore in updated soil hydraulic parameters in CLM. In order to show the influence of calibrated soil texture on the soil hydraulic parameters, the temporal evolution of the calibrated saturated hydraulic conductivity K and the empirical parameter B which

represent the slope of water retention curve in the Clapp–Hornberger parameterization (Clapp and Hornberger, 1978; Oleson et al., 2013) are shown in Fig. 5 (scenario DA_SM_Par). The Clapp–Hornberger parameterization can be used to calculate the hydraulic conductivity when the soil water retention data are not available. All soil hydraulic parameters could be improved in this scenario by assimilating the CRP neutron intensity. The mean ensemble values for the soil hydraulic parameters approached the reference values over the assimilation period. As the soil properties were not updated during the intensive irrigation period, the convergence was slow.

[Insert Figure 5 here]

Fig. 6 shows the time series of irrigation amount at the location of the CRP for the reference case and for the different estimation scenarios. The reference irrigation amount (707.2 mm) was calculated in CLM on the basis of the water deficit method. Fig. 6 illustrates that the intensive irrigation period from June to October coincided with limited precipitation. The reference irrigation amount was around twice of the annual precipitation amount (356.1 mm). The sum of the reference irrigation amount (707.2 mm) and annual precipitation (356.1 mm) was about 6% higher than the documented potential annual evapotranspiration of citrus trees (Ballester et al., 2011; Jimenez-Bello et al., 2015). We used a t-test to compare the estimated annual irrigation requirements by the different scenarios with the reference case. Large p values (Fig. 6) indicate that the scenarios with data assimilation did not have a significantly different irrigation scheduling as compared with the reference case. The scenario No_DA has a p value <0.05, which indicates that the estimated irrigation

amounts were significantly different from the reference case. Data assimilation (Only_DA_SM) improved results with higher p values (p=0.205). Scenarios with bias correction (p=0.819) and parameter estimation (p=0.755) gave better results than the scenario of state estimation only (Only_DA_SM). The annual irrigation amounts for the grid cell at the location of the CRP are summarized in Fig. 7. Obviously, joint soil moisture and bias (or parameter estimation) improved the characterization of irrigation requirement.

[Insert Figure 6 here]

[Insert Figure 7 here]

Now we analyze the irrigation results for the scenarios in detail. The better characterization of the irrigation demand was found by combining CRP neutron intensity assimilation and soil properties estimation (or soil moisture bias estimation). Scenarios DA_SM_Bias and DA_SM_Par estimated an irrigation requirement of 691.0 mm and 685.6 mm, respectively. The scenario Only_DA_SM (i.e., data assimilation with state estimation only) gave an irrigation estimation of 797.7 mm. CRP neutron intensity assimilation without parameter estimation provided much better results than scenarios without data assimilation (No_DA). The estimated irrigation amount for the scenario No_DA was 1107.3 mm. These results illustrate that irrigation estimation for the case of biased soil properties (i.e., sandier soil in model than in reference run) can be improved significantly by the assimilation of CRP neutron intensity and even better by including soil properties estimation or bias estimation. Data assimilation improved the estimation of irrigation demand and even

resulted in a slightly lower irrigation than for the reference case.

The CRP footprint was composed of 43 CLM grid cells which were irrigated separately. The irrigation requirements estimated for the different scenarios are displayed in Fig. 8 and compared with the reference scenario. The spatial irrigation patterns of scenarios DA_SM_Bias and DA_SM_Par are closer to the reference case than the scenario without data assimilation (No_DA).

[Insert Figure 8 here]

The single CRP neutron intensity measurement for the coarse scale (600 m) was used to update the 43 CLM grid cells at the fine scale (100 m). The spatial distribution of soil properties and annual irrigation amount were compared with the reference spatial distributions. A higher spatial similarity is associated with lower HD values. The HD values were evaluated according to the distance between a CLM grid cell and the CRP location. Three classes were defined: (i) distance CRP- grid cell ≤ 100 m; (ii) distance CRP- grid cell > 100 m and ≤ 200 m; (iii) distance CRP- grid cell > 200 m and ≤ 300 m. The HD values for the comparison of the spatial patterns of soil properties for the scenarios No_DA and DA_SM_Par are shown in Fig. 9. The similarity of soil properties between background and reference is small due to the imposed bias, and the HD values were 580.83, 1186.53, 1203.93 for sand fraction, clay fraction and organic matter density, respectively. The scenario DA_SM_Par resulted in a spatial distribution of soil properties closer to the reference case, with HD values of 69.83 for sand fraction (580.83 for No_DA), 149.96 for clay fraction (1186.53 for No_DA) and 185.21 for organic matter density (1203.93 for No_DA)

HD values for sand fraction, clay fraction and organic matter density for the region with distance CRP- grid cell ≤ 100 m decreased by 80%, 82% and 64%, compared to No_DA, respectively. For the region with distance CRP- grid cell > 100 m and ≤ 200 m the decreases were 82%, 66% and 67%, respectively. Finally, for the region with a distance CRP- grid cell > 200 m and ≤ 300 m the decreases were 78%, 82% and 40%, respectively.

Figure 10 shows the HD values for the comparison of the spatial distribution of irrigation amounts with the reference. The figure includes comparisons for all different data assimilation scenarios. It is clear that the assimilation of CRP neutron intensity (Only_DA_SM) increased the similarity between the spatial distribution of estimated annual irrigation amount and the reference irrigation distribution. Soil properties estimation and soil moisture bias estimation increased the similarity in spatial irrigation pattern further. The HD values for the scenario DA_SM_Bias decreased 89% (distance CRP- grid cell ≤ 100 m), 81% (distance CRP- grid cell > 100 m and ≤ 200 m) and 82% (distance CRP- grid cell > 200 m and ≤ 300 m) compared to No_DA. The HD values for the scenario DA_SM_Par, for the same three distance classes and in the same order, decreased by 88%, 87% and 85% compared to No_DA.

[Insert Figure 9 here]

[Insert Figure 10 here]

The total annual ET for the different scenarios was also calculated, and also the contributions from ground evaporation, evaporation of intercepted water by the canopy and canopy transpiration. The ET for the reference scenario was 756.6 mm.

The ET for all other scenarios was very close to the reference. In case of the scenario No_DA, too much water was irrigated so that drought stress did not occur. The other scenarios with data assimilation resulted in less irrigation, but ET was also close to the reference value, indicating that less irrigation was not associated with plant stress. For all the scenarios, the ET values did not deviate much from the reference value. This is because in all cases an overestimation of the percentage of sand led to excessive irrigation and sustainment of potential ET. The largest contribution to the ET was the canopy transpiration. The irrigated grid cells were assumed to be fully covered by the vegetation, and therefore the ground evaporation was low. The low evaporation from the canopy intercepted water maybe related to the fact that the rain events occurred mainly in the spring and winter seasons.

5. Discussion

The proposed data assimilation and parameter estimation (or bias estimation) can improve the soil moisture and irrigation estimation. The joint state-parameter estimation is the best scenario, and reduced the RMSE values of soil moisture content more than 50%, the spatial similarity of irrigation amount was increased and the HD values were decreased by 86% on average. The novelty of this work was the assimilation of the new CRP data in combination with irrigation scheduling. In general, classical parameter estimation tends to focus on uncertainty in the parameter estimates only, while neglecting partial or all of the other uncertainty sources (Liu and Gupta, 2007). We did not aim to compare the parameter estimation methodology with

other methodologies in this study. The synthetic CRP neutron intensity observations were assimilated in CLM and the synthetic study potentially overestimated the performance of the proposed method.

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

operational implementation on-line parameter estimation was avoided and only states were updated. It is therefore possible that for on-line irrigation scheduling a conservative, potentially less successful strategy should be followed where only states are updated. This synthetic study showed that state updating only also would improve irrigation scheduling considerably. We believe therefore that although in a real-world case study results will be less favorable than in the synthetic study, data assimilation with updating states only, or joint updating of states and bias (or parameters) in case of a systematic bias, will improve irrigation scheduling compared to a scenario without data assimilation.

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

measured CRP neutron intensity. The soil spatial heterogeneity in the CRP footprint was introduced by adding a random spatially correlated noise. Heterogeneous land cover was not considered in this study. However, the spatial variability of ecosystem parameters could be a further confounding parameter influencing the results. Altogether, accounting for temporally variable biomass in the COSMIC operator does not seem a large limitation, but spatially variable soil moisture conditions within the cosmic ray probe footprint, are a serious challenge.

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

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

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

participation of farmers.

A further possible improvement is the consideration of irrigation below ET requirement, known as deficit irrigation (DI), which can reduce water demand to meet the maximum ET (Fereres and Soriano, 2007).

6. Conclusions

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

The results show that assimilating CRP neutron intensity can improve joint soil moisture and soil properties estimation, and irrigation scheduling. Data assimilation schemes that remove soil moisture bias or update soil properties on the basis of CRP neutron intensity outperform data assimilation without bias or parameter estimation.

The joint soil moisture and soil parameter estimation with simple multiscale assimilation strategy of CRP neutron intensity can potentially be used for irrigation scheduling in the future. The main challenges for the real world application are: model calibration to remove the bias, forward modeling of cosmic-ray neutron intensity under high vegetation coverage, precise weather forecasts and cooperation of farmers.

References:

Aravéquia, J.A. et al., 2011. Evaluation of a Strategy for the Assimilation of Satellite Radiance Observations with the Local Ensemble Transform Kalman Filter. Mon Weather Rev, 139(6): 1932-1951.

Baatz, R. et al., 2015. An empirical vegetation correction for soil water content quantification using cosmic ray probes. Water Resour Res, 51(4): 2030-2046.

Baek, S.J., Hunt, B.R., Kalnay, E., Ott, E., Szunyogh, I., 2006. Local ensemble Kalman filtering in the presence of model bias. Tellus A, 58(3): 293-306.

Ballester, C., Castel, J., Intrigliolo, D.S., Castel, J.R., 2011. Response of Clementina de Nules citrus trees to summer deficit irrigation. Yield components and fruit composition. Agr Water Manage, 98(6): 1027-1032.

Bateni, S.M., Entekhabi, D., 2012. Surface heat flux estimation with the ensemble Kalman smoother: Joint estimation of state and parameters. Water Resour Res, 48.

Bauser, G. et al., 2012. A comparison study of two different control criteria for the real-time management of urban groundwater works. J Environ Manage, 105: 21-9.

Bogena, H.R., Huisman, J.A., Baatz, R., Hendricks Franssen, H.J., Vereecken, H., 2013. Accuracy of the cosmic-ray soil water content probe in humid forest ecosystems: The worst case scenario. Water Resour Res, 49(9): 5778-5791.

Bosilovich, M.G., Radakovich, J.D., da Silva, A., Todling, R., Verter, F., 2007. Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. J Meteorol Soc Jpn, 85A: 205-228.

Bouyoucos, G.J., 1962. Hydrometer Method Improved for Making Particle Size Analyses of Soils1. Agron J, 54(5): 464.

Chen, Y., Zhang, D.X., 2006. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Adv Water Resour, 29(8): 1107-1122.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resour Res, 14(4): 601-604.

Crow, W.T. et al., 2012. Upscaling Sparse Ground-Based Soil Moisture Observations for the Validation of Coarse-Resolution Satellite Soil Moisture Products. Rev Geophys, 50(2): n/a-n/a.

Crow, W.T., Kustas, W.P., Prueger, J.H., 2008. Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model. Remote Sens Environ, 112(4): 1268-1281.

De Lannoy, G.J.M., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007a. State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency. Water Resour Res, 43(6).

De Lannoy, G.J.M., Reichle, R.H., Houser, P.R., Pauwels, V.R.N., Verhoest, N.E.C., 2007b. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. Water Resour Res, 43(9): n/a-n/a.

Dee, D.P., 2005. Bias and data assimilation. Q J Roy Meteor Soc, 131(613): 3323-3343.

Desilets, D., Zreda, M., Ferre, T.P.A., 2010. Nature's neutron probe: Land surface hydrology at an elusive scale with cosmic rays. Water Resour Res, 46(11): W11505.

Entekhabi, D. et al., 2010. The Soil Moisture Active Passive (SMAP) Mission. P Ieee, 98(5): 704-716.

Fereres, E., Soriano, M.A., 2007. Deficit irrigation for reducing agricultural water use. J Exp Bot, 58(2):

147-159.

Franssen, H.J.H. et al., 2011. Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating. Water Resour Res, 47(2): n/a-n/a.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Jimenez-Bello, M.A., Castel, J.R., Testi, L., Intrigliolo, D.S., 2015. Assessment of a Remote Sensing Energy Balance Methodology (SEBAL) Using Different Interpolation Methods to Determine

Evapotranspiration in a Citrus Orchard. Ieee J-Stars, 8(4): 1465-1477.

Kerr, Y.H. et al., 2010. The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle. P Ieee, 98(5): 666-687.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Oleson, K. et al., 2013. Technical Description of version 4.5 of the Community Land Model (CLM). Ncar

Technical Note NCAR/TN-503+STR, National Center for Atmospheric Research, Boulder, CO, 422 pp.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Walker, J.P., Willgoose, G.R., Kalma, J.D., 2001. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. Adv Water

Resour, 24(6): 631-650.

Waller, J.A., Dance, S.L., Lawless, A.S., Nichols, N.K., Eyre, J.R., 2014. Representativity error for temperature and humidity using the Met Office high-resolution model†. Q J Roy Meteor Soc, 140(681): 1189-1197.

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

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

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

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

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

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

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

Table List

Table 1 Summary of perturbation parameters for atmospheric forcing data

Table 1 Summary of perturbation parameters for atmospheric forcing data

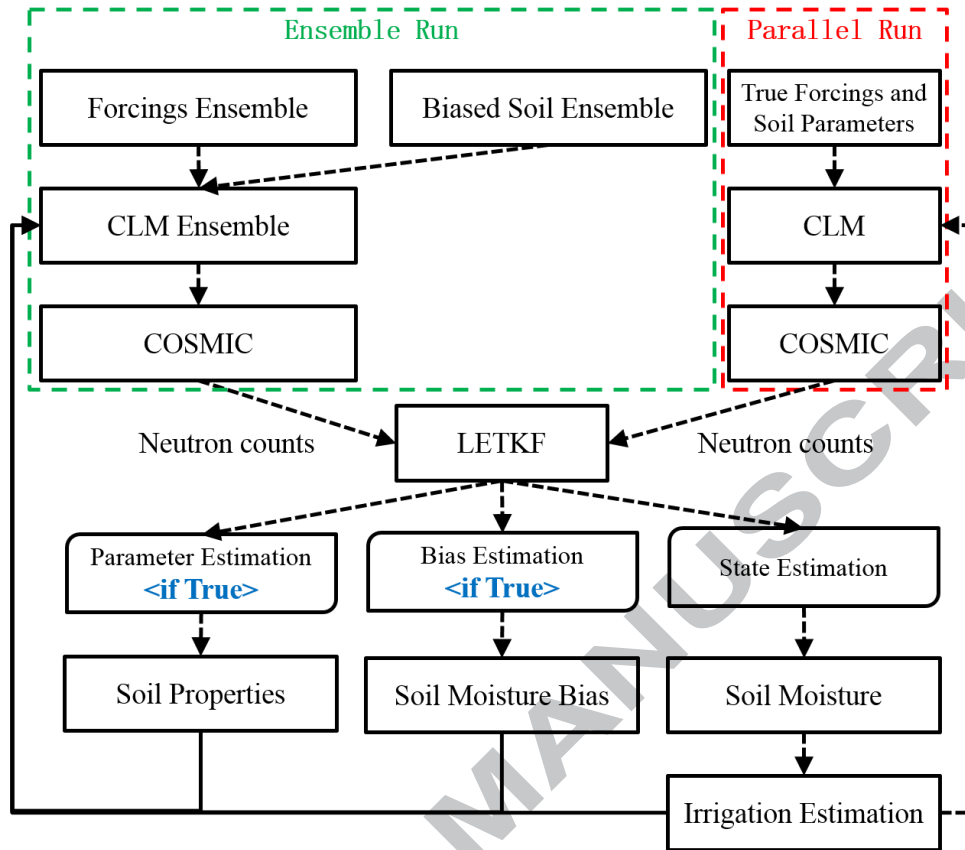| Variables	Noise	Standard deviation	Time Correlation scale	Forcing Cross Correlation
Precipitation	Multiplicative	0.5	24 h	[1.0,-0.8, 0.5, 0.0,
Shortwave radiation	Multiplicative	0.3	24 h	-0.8, 1.0,-0.5, 0.4,
Longwave radiation	Additive	20 W/m^2	24 h	0.5, -0.5, 1.0, 0.4,
Air temperature	Additive	1 K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

Figure List

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

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

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

Figure 4. Temporal evolution (collected every three days) of soil moisture bias for the first soil layer at the CRP location (scenario DA_SM_Bias), the true soil moisture bias was calculated from scenario No_DA is showed in blue

Figure 5. Temporal evolution (collected every three days) of saturated hydraulic conductivity K of soil (K_10cm) and the empirical parameter *B* of the Clapp–Hornberger parameterization (B_10cm) at the CRP location for the scenario DA_SM_Par

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)

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

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

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

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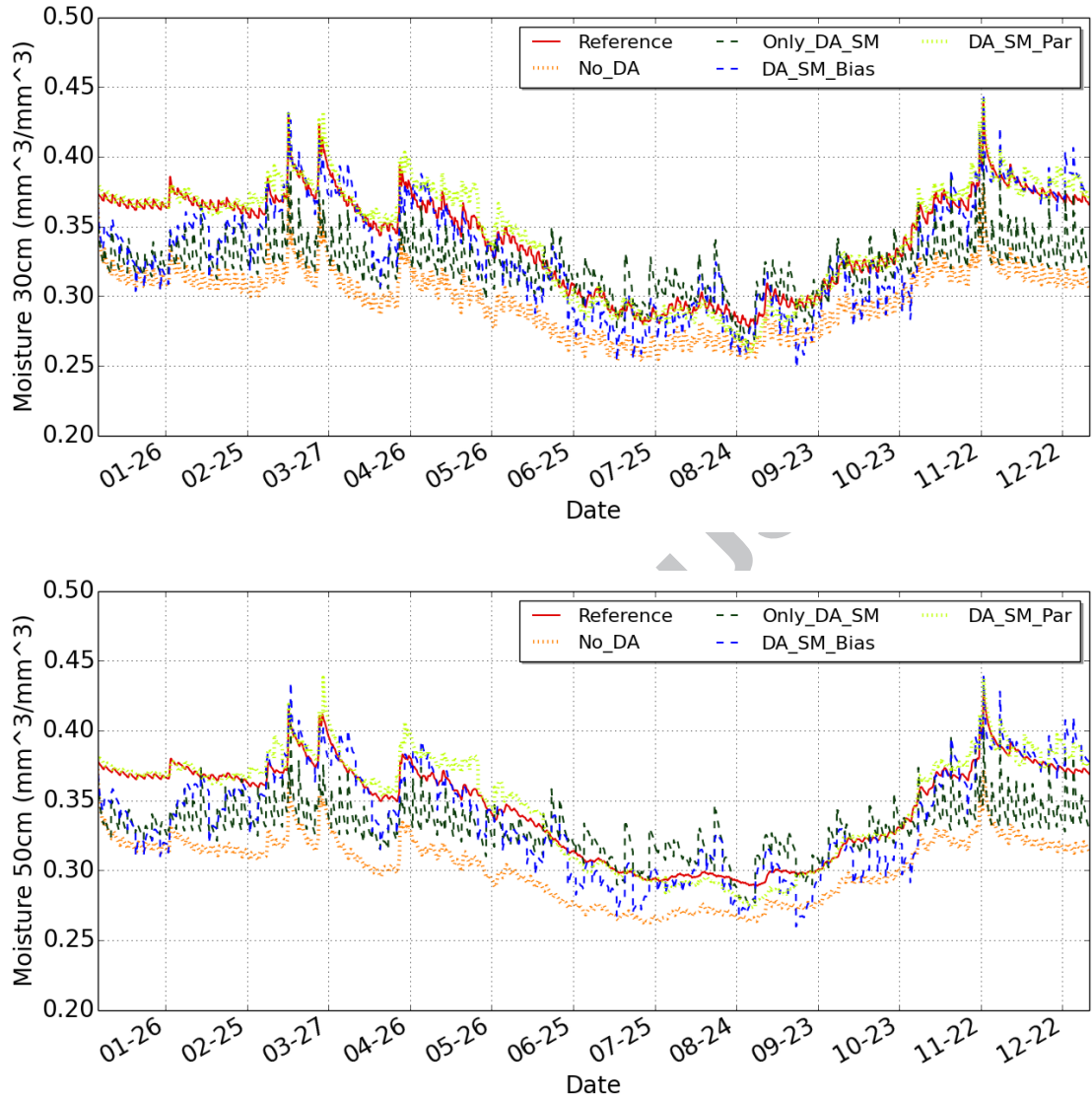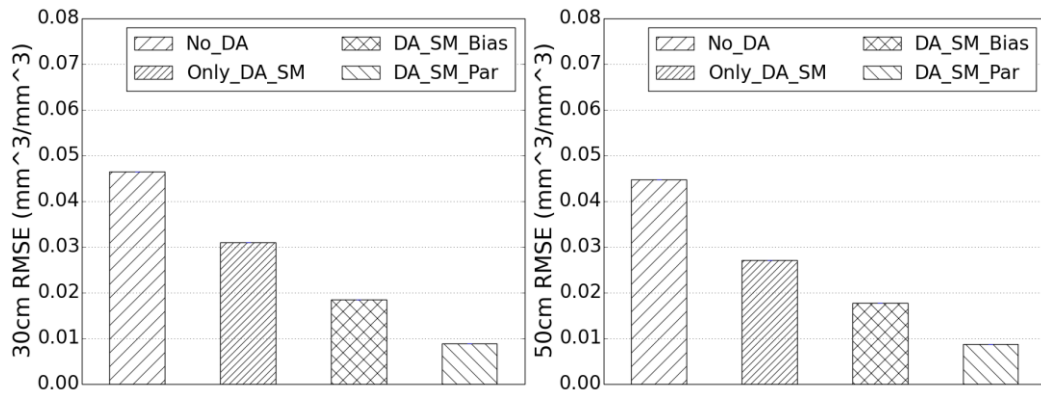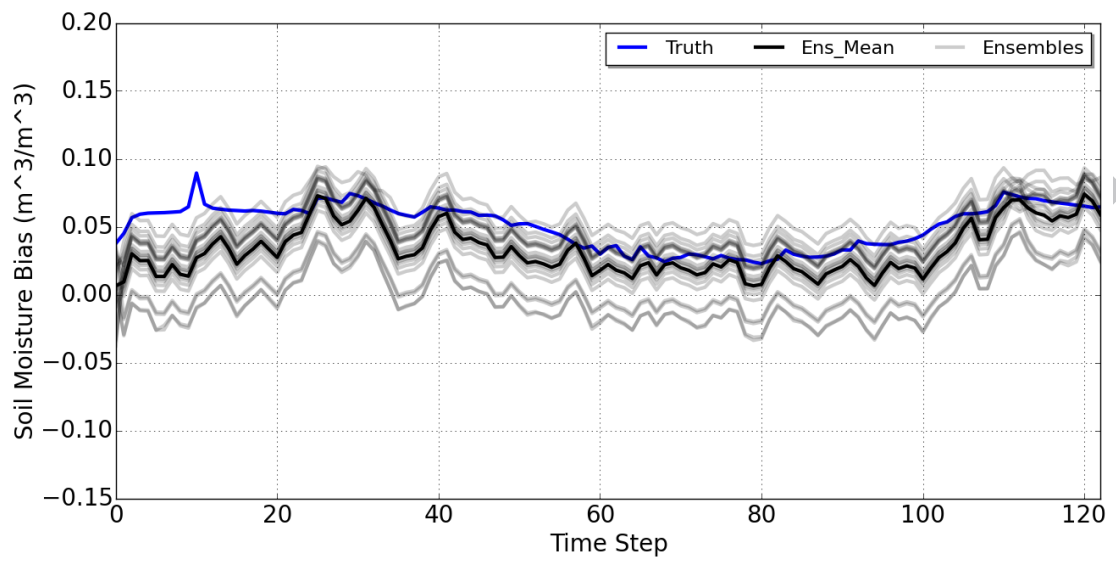
procedure

Figure 2. Soil moisture content at 30 cm (upper graph) and 50 cm (lower graph) depth at the CRP location for the different simulation scenarios

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

Figure 4. Temporal evolution (collected every three days) of soil moisture bias for the first soil layer at the CRP location (scenario DA_SM_Bias). The true soil moisture bias was calculated from the scenario No_DA and is shown in blue. The unit of x-axis is for time steps of 3 days.
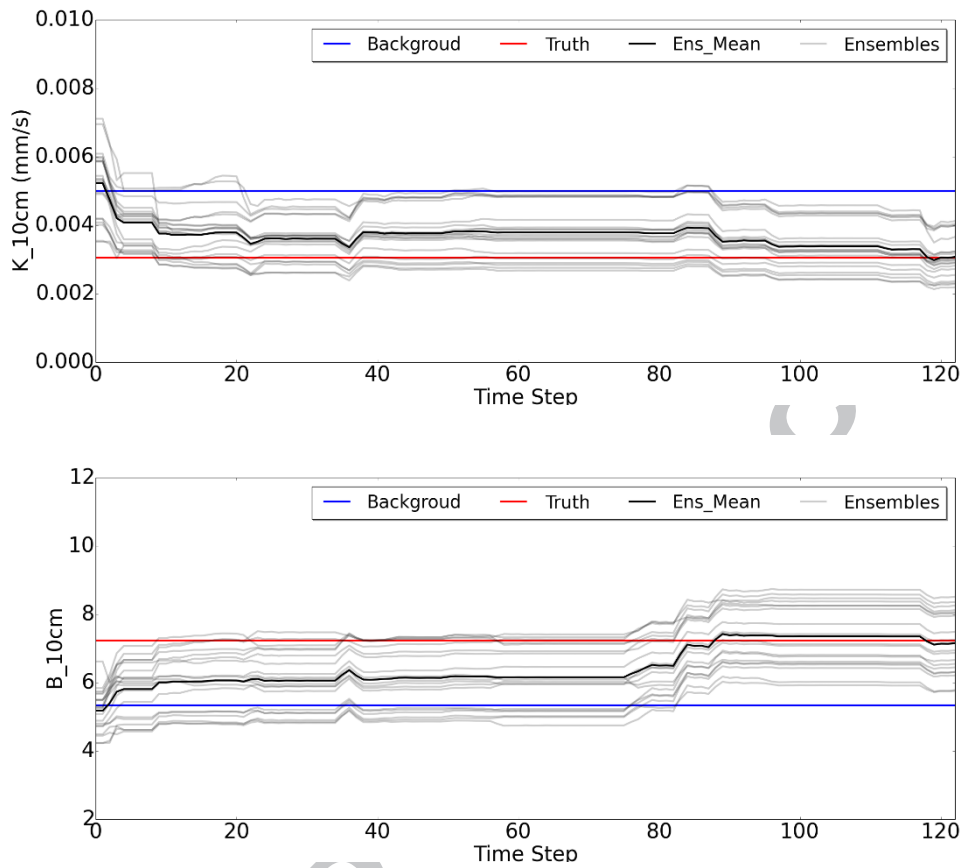
Figure 5. Temporal evolution (collected every three days) of saturated hydraulic conductivity *K* of soil (K_10cm) and the empirical parameter *B* of the Clapp–Hornberger parameterization (B_10cm) at the CRP location for the scenario DA_SM_Par. The unit of x-axis is for time steps of 3 days.
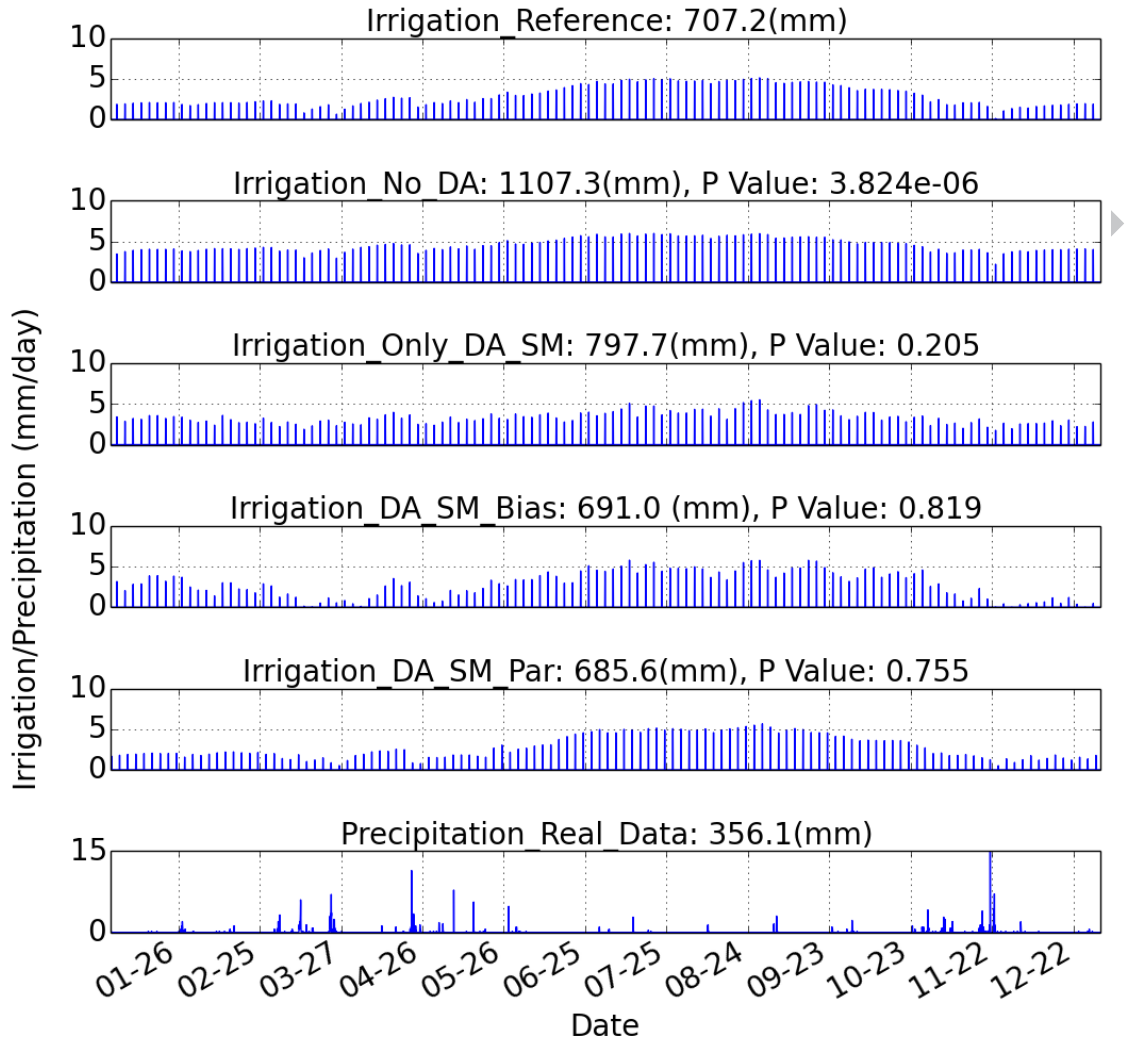
Figure 6. Irrigation requirement as function of time at the CRP location for the different scenarios; t-test statistics (p-value) with significance level 0.05 for comparing the calculated irrigation distribution with reference irrigation are also shown for the different scenarios (large p-values indicate high similarity)
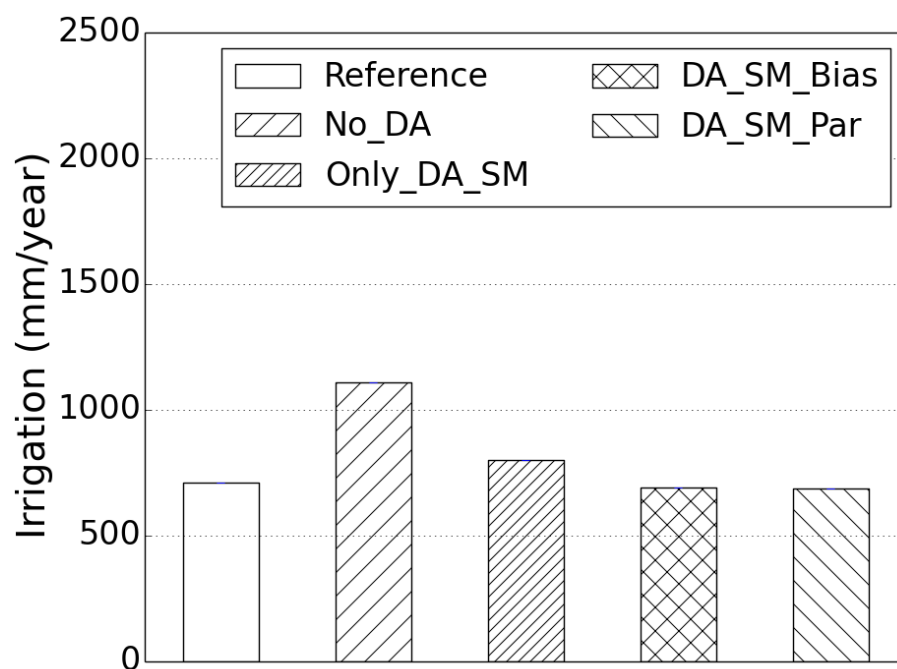
Figure 7. Annual irrigation requirement according to the different scenarios at the
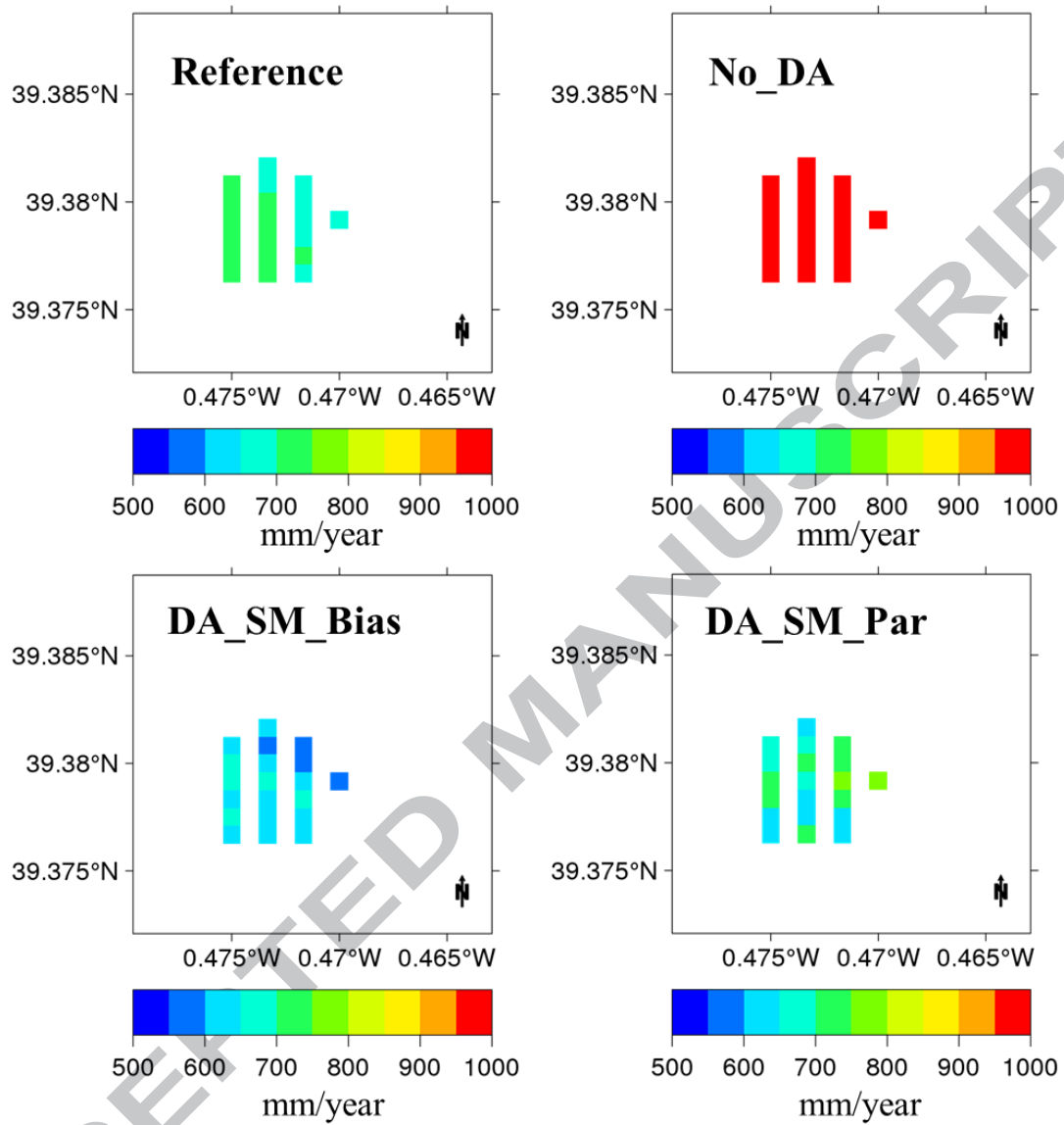
CRP location

Figure 8. Annual irrigation calculated for different simulations scenarios and compared to the reference scenario
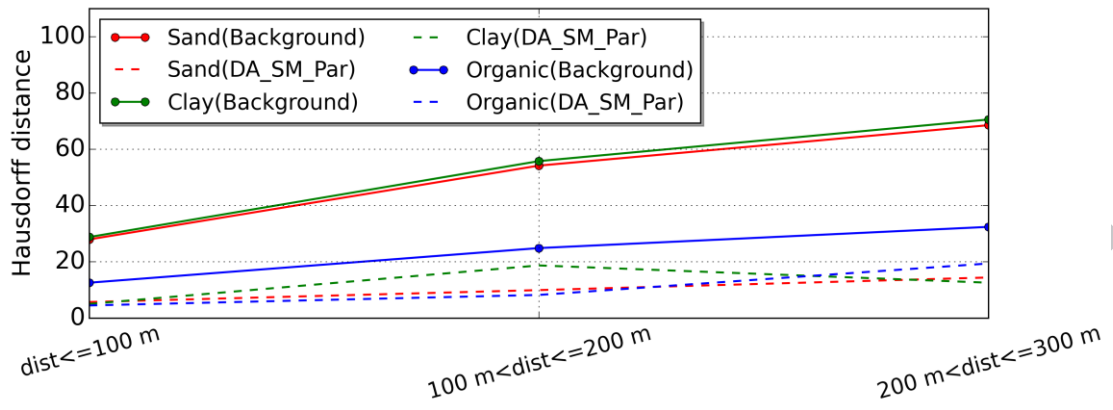
Figure 9. Hausdorff distance values for background soil properties (sand fraction, clay fraction and organic matter density) and estimated soil properties (scenario DA_SM_Par). Results are plotted as function of the distance between model grid cells and the CRP location
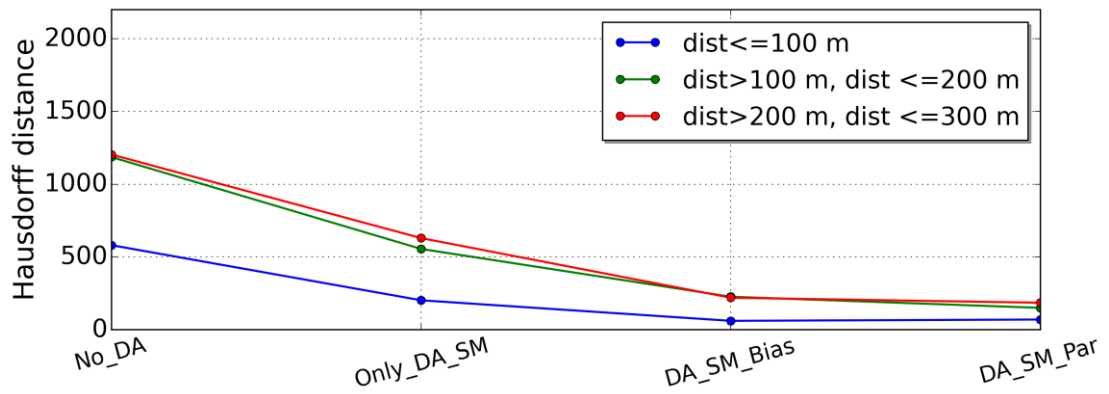
Figure 10. Hausdorff distance values of calculated annual irrigation requirement, compared to reference irrigation, for different scenarios. Results are plotted as function of distance between model grid cells and CRP location

Highlights

- Joint assimilation, parameter and bias estimation, and irrigation were evaluated
- Synthetic cosmic-ray neutron intensity data were used to update the soil moisture
- Horizontal and vertical weighting was introduced in data assimilation
- Soil moisture, soil properties and model bias were updated jointly in assimilation
- Joint state-parameter resulted in the best estimation