Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

Reducing soil moisture measurement scale mismatch to improve surface energy flux estimation

Joost Iwema[1], Rafael Rosolem[1,2], Mostaquimur Rahman[1], Eleanor Blyth[3], Thorsten Wagener[1,2]

[1]Department of Civil Engineering, University of Bristol, Bristol, UK
[2]Cabot Institute, University of Bristol, Bristol, UK
[3]Centre for Ecology and Hydrology, Wallingford, OX10 8BB, UK

Correspondence to: Joost Iwema (joost.iwema@bristol.ac.uk)

Abstract. At the so-called hyper-resolution scale (i.e. grid cells of 1 km^2) Land Surface Model (LSM) parameters are sometimes calibrated with Eddy-Covariance (EC) data and Point Scale (PS) soil moisture data. However, measurement scales of EC and PS data differ substantially. In our study, we investigated the impact of reducing the scale mismatch between surface energy flux data and soil moisture data by replacing PS soil moisture data with observations derived from Cosmic-Ray Neutron Sensors (CRNS) made at larger spatial scales. Five soil-evapotranspiration parameters of the Joint UK Land Environment Simulator (JULES) were calibrated against PS and CRNS soil moisture data separately. We calibrated the model for twelve sites in the USA representing a range of climatic, soil, and vegetation conditions. The improvement in surface energy partitioning for the two calibration solutions was assessed by comparing to EC data and to a version of JULES runs with default parameter values. We found that simulated surface energy partitioning did not differ substantially between the PS and CRNS calibrations, despite their differences in actual soil moisture observations. We concluded that potential differences due to distinct spatial scales represented by the PS and CRNS soil moisture sensor techniques were substantially undermined by the weak coupling between soil moisture and evapotranspiration within JULES.

1 Introduction

The land surface water and energy balances are coupled through the process of evapotranspiration. Soil moisture is one of the main water reservoirs near the land surface and can hence importantly control the surface water and energy balances. Soil moisture provides a first order (i.e. direct) control on evapotranspiration when there is insufficient water to meet the evaporative demand (Manabe et al., 1969; Budyko, 1956; Seneviratne et al., 2010). An indirect effect of soil moisture on surface energy flux partitioning is for instance the damping effect on soil and air temperature, which in its turn affects humidity, evapotranspiration, boundary-layer stability, and in some cases precipitation (Seneviratne et al., 2010). The coupling between soil moisture and surface energy flux partitioning is especially strong in transitional climate regions (Koster 2004).

Land Surface Models (LSMs) solve the surface mass (including water), energy, and momentum balances to provide the weather and climate prediction models with lower boundary conditions. The land surface has been shown to play an important role in atmospheric circulation (Koster et al., 2004). Because the soil moisture state and surface fluxes are so closely connected, it is

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

believed important to accurately simulate these simultaneously (Henderson-Sellers et al., 1996; Richter et al., 2004; Seneviratne et al., 2010; Dirmeyer, 2011; Dirmeyer et al., 2013).

LSMs have become more complex over the past decades. The first LSMs included only limited representations of soil hydrological and biophysical processes. The `second-generation' LSMs, however included the main physical processes within vegetation and in soil (Sellers, 1997; Seneviratne et al., 2010), and the later developed `third-generation' LSMs, additionally included plant physiological processes like photosynthesis and carbon assimilation (Seneviratne et al., 2010). This increasing complexity has brought with it an increasing number of parameters, with values not easily defined with in-situ measurements because of scale mismatch, e.g. stomatal resistance measured at leaf level is not the same as canopy stomatal resistance needed for LSMs (Blyth et al., 1993). Soil hydraulic parameter values (e.g. soil hydraulic conductivity) are often obtained from laboratory experiments on soil cores for cubic centimetres to cubic decimetres. The soil properties and processes at this scale can however differ from those at the LSM grid cell sizes, which are often as large as hundreds to tens of thousands of square kilometres (Pitman 2003). Due to the different governing processes upscaling the soil hydraulic properties from soil core scale to field scale is non-trivial (Vinnikov 1996; Crow et al., 2012).

Following discussion on how to move forward with hydrological models and LSMs in an effort to develop models that cover all land surface of the globe (Wood et al., 2011; Beven and Cloke, 2012), models are increasingly being applied at the finer `hyper-resolution scale' with grid cells of about 1 km^2. Typically, parameters are calibrated and validated at this hyper-resolution against in-situ measurements from sources such as Eddy-Covariance (EC) flux towers and point-scale (PS) soil moisture sensors (e.g. Time Domain Transmissivity; TDT and Time Domain Reflectometry; TDR) (Stockli et al., 2008; Richter et al., 2004; Blyth et al., 2010; Blyth et al., 2011; Rosolem et al., 2013). Such calibration/validation data has become more widely available at hyperresolution scale (Baldocchi, 2001; Smith, 2012). However, the horizontal footprints of different measurement techniques' vary from each other: EC surface energy flux data represent a downwind footprint of 100 m^2 to 1 km^2, while in-situ soil moisture probes link to much smaller surface areas by representing a support volume (soil volume represented by sensor; Blöschl and Sivapalan, 1995) of ~4 dm^3 only (Running et al., 1999; Kurc and Small, 2007; Vivoni et al., 2008). Soil moisture can be spatially non-uniform within the EC footprint due to heterogeneity in soil properties, vegetation, and topography. How variable the soil moisture content is under certain heterogeneous conditions depends on the wetness conditions; past research indicates the variability to be highest during soil wetting and drying periods. Ideally an EC footprint average soil moisture content which best (i.e. most effectively) connects to the surface exchange processes should be used within an LSM grid cell. If soil moisture is measured at only a single or few locations with limited support volume, like with PS sensors, the measured soil moisture content might be different from the effective soil moisture state that controls the surface exchange processes. This poses a potential scale mismatch issue, as depicted in Figure 1. The question which then rises is,

d*oes reduced observation scale mismatch improve LSM energy partitioning estimation?*

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

In recent years, new soil moisture measurement techniques have been developed that have, compared to in-situ soil moisture sensing techniques (TDT and TDR), a reduced scale mismatch with EC surface energy flux measurements. Improvement in wireless technology and remote data collection technology have made the development of PS soil moisture sensor networks more feasible (Cardell-Oliver, 2005; Ritsema et al., 2009; Trubilowicz, 2009; Bogena et al., 2010; Robinson et al., 2008). It is usually assumed that, due to spatial variability in soil physical characteristics, a network of PS sensor profiles is more informative than a single or a couple of PS sensor profiles for studying land-surface processes and constraining model parameters at scales of >~100 m^2 (Robinson et al., 2008).

Newer soil moisture sensor techniques, for instance one which makes use of the Global Positioning System (GPS; Larson et al 2008, 2010), and the Cosmic-Ray Neutron Sensor (CRNS; Zreda et al., 2008) have the advantage that only a single, above ground sensor is needed, which is easier to install than a PS network.

The CRNS (Zreda et al., 2008) is an above-ground passive sensor which utilises natural cosmic-ray neutron radiation to estimate soil moisture content. Networks of CRNS have been established in various countries (e.g. COSMOS (Zreda et al., 2012), COSMOS-UK (Evans et al., 2016), CosmOZ (Hawdon et al., 2004)). The CRNS passively measures fast neutron intensity, which can be non-linearly related to the soil moisture content in the top 10-70 cm, from an area with a radius of about 100 to 300 m surrounding the above-ground sensor (Figure 2; Desilets and Zreda, 2013; Kohli et al., 2015). The CRNS has been shown to provide similar soil moisture information as co-located in-situ sensor networks (Franz et al., 2012).

Unlike wireless in-situ sensor networks, which consist of a set of point measurements, both the GPS and CRNS technology provide an integrated soil moisture measurement over the entire support volume (Larson et al., 2008; Zreda et al., 2008). We chose to answer our research question using the CRNS technology because the COSMOS network provides publicly available data for multiple years at a range of sites co-located with Ameriflux/FLUXNET EC towers (ORNL-DAAC, 2015). We used twelve of these sites which provided sufficient LSM forcing data, PS soil moisture data, CRNS data, and EC LE and sensible heat flux (H) data. We phrased the following hypothesis for our research question: *Due to reduced scale mismatch, CRNS yields LSM surface flux estimates closer to eddy-covariance observations than in-situ soil moisture sensors.*

To investigate our research question we made the LSM simulated soil moisture content match the observed PS or CRNS data as closely as possible. We did this by calibrating parameters of LSM Joint UK Land Environment Simulation Model (JULES; Best et al., 2011) that importantly affect the soil moisture state in the model against the PS and CRNS data separately. We subsequently validated the results against EC observed data over the same periods. To assess the change in soil moisture and surface energy after calibration we compared the calibrated runs against a default run with parameter values computed from a widely used soil properties database. Before our modelling exercise we first compared the PS and CRNS data. The outcomes of this data analysis were mainly used to see if the results from the calibration and validation yielded larger differences in surface energy flux estimation at sites where the two soil moisture observation products showed higher deviation from each other.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

2 Data and Methods

2.1 Calibration and Validation data: PS, CRNS, and EC data

PS soil moisture and CRNS neutron count data from twelve Ameriflux/COSMOS sites were used (Figure 2). These twelve sites covered eight of twenty Ecoclimatic domains of the US National Ecological Observatory Network (NEON; www.neonscience.org) (Figure 2). These twelve sites hence represent a variety of climates and land cover types, but also different soil types (Table 1; for full site names see Appendix 1, Table A1.1).

Hourly PS data for nine sites were obtained from the publicly available Ameriflux Level 2 data source (ORNL). Data for the three California Climate Gradients sites (DC, CS, and SO) were obtained at http://www.ess.uci.edu/~california/ (data version 3.4; Goulden et al., 2015). The number of PS profiles, the installation depths, and sensor types differed between the twelve study sites (Table A1.1 in Appendix 1). We used data from the upper soil layers (up to 30 cm) only, because the CRNS data represents the contribution from the upper layers only. Quality control was applied to filter out spurious and unrealistic data points due to sensor errors. The PS data was then interpolated to the JULES soil layer on which the model was calibrated.

Hourly CRNS neutron count data were obtained from the COSMOS network website (www.cosmos.hwr.arizona.edu). Corrections were applied as by Zreda et al. (2012). Water vapour corrections (Rosolem et al., 2013) were applied with respect to dry air (Bogena et al., 2013). Similar quality control approach used for the PS analysis was also applied to CRNS neutron count data series to remove unrealistic points. Snow cover periods were also removed for the analysis. A 5-hour moving average window was applied to the observed CRNS neutron counts (following Shuttleworth et al., 2013; Rosolem et al., 2014). We compared PS soil moisture with CRNS integrated soil moisture computed from vertically homogeneous soil moisture values obtained from the observed neutron counts using the COsmic-ray SoiI Moisture Interaction Code (COSMIC; Shuttleworth et al., 2013).

We used the exact same data for the soil moisture data comparison as for the calibration (providing that both PS and CRNS were available in the same period). As validation data, latent heat (LE) and sensible heat (H) flux hourly data from Ameriflux Level 2 data source was used for nine sites while for the three California Climate Gradients sites data version 3.4 (Goulden et al., 2015). Quality control was applied to the LE and H flux data to remove outliers and unrealistic data points.

2.2 JULES forcing data and initial conditions

JULES requires precipitation, air temperature, atmospheric pressure, wind speed, specific humidity, downward shortwave radiation, and downward longwave radiation as meteorological forcing data. Quality controlled hourly data was obtained from Ameriflux Level 2 and the three California Climate Gradients sites. At some of the sites however, certain specific forcing data was not available from Ameriflux and hence data from different sources was used (Table A1.2 of Appendix 1).

Model input data was gap-filled following Rosolem et al. (2010) because JULES requires continuous time series except for precipitation where gaps were set to zero. For all sites gaps smaller than 3 hours were filled by linear interpolation, while larger gaps of up to 31 days were gap-filled using the average diurnal pattern of the preceding and following 15 days. In addition,

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

some remaining gaps in Downward Shortwave Radiation and Downward Longwave Radiation at Wind River (WR) were filled using linear least squares relationships with NLDAS-2 data (http://disc.sci.gsfc.nasa.gov/uui/datasets?keywords=NLDAS). At Tonzi Ranch (TR) NLDAS data were used to gap fill air temperature. Gap filling at Santa Rita Creosote was done with data from the nearby Sahuarita site followed by the gap-filling procedure described above.

2.3 Soil moisture data comparison methodology

To compare PS and CRNS soil moisture data we computed the Mean Squared Deviation (MSD) and its decomposition (Gupta et al., 2009) into:

(1) The squared difference between the means (structural bias)

(2) The squared difference between the standard deviations (indicates different seasonality)

(3) A term relating to the coefficient of correlation (r; indicates different dynamics): $2 \cdot \sigma_{ps} \cdot \sigma_{crns} \cdot (1\text{-}r)$

We however scaled the relative contributions of the three MSD components to the Root Mean Squared Deviation (RMSD) instead of MSD, to keep units of m^3m^3. We then ranked the sites from lowest to highest RMSD. According to our hypothesis, we would expect to see a larger difference in simulated surface energy fluxes after calibration when the two soil moisture time series differ most. The CRNS soil moisture values computed with COSMIC represent the integrated signal computed from vertically homogeneous profiles, whereas in reality heterogeneous profiles are common across different sites and conditions.

2.4 Calibration and validation methodology

2.4.1 Joint UK Land Environment Simulator (JULES)

We used the Joint UK Land Environment Simulator (JULES; Best et al., 2011; Clark et al., 2011) in this study. JULES can be coupled as lower boundary condition to the UK Met Office Unified Model (Cullen, 1993). Within JULES choices can be made (e.g. canopy radiation model type) and certain modules (e.g. vegetation dynamics) can be switched on or off to operate at different levels of complexity. In addition, we chose the UK Variable resolution configuration (UKV) because it is the standard setting when JULES is run coupled with the UK Met Office Unified Model. The UKV land grid cell size is 1 km by 1 km. However, our study focused on JULES standalone simulations at the 12 grid points located at the sites investigated. The UKV setting employs the multi-layer canopy radiation module with surface heat capacity and snow beneath the canopy, the single canopy layer `big leaf' approach for leaf-level photosynthesis (which computes radiation absorption with Beer's law). Soil heat conductivity was calculated using the approach of Dharssi et al. (2009). We used the default JULES-UKV soil layering (Figure 3). The hydraulic bottom boundary condition in JULES is free drainage.

JULES computes the transport of water through the soil using a finite difference representation of the Richards equation. The vertical fluxes are computed with the Buckingham-Darcy equation. JULES-UKV uses the Mualem-Van genuchten (Van Genuchten, 1980; Mualem, 1983) soil water retention equations. The Van Genuchten equation calculates soil water content θ (m^3m^{-3}) from soil hydraulic pressure head ψ (m):

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

$$\frac{\theta-\theta_{res}}{\theta_{sat}-\theta_{res}} = \frac{1}{[1+(\alpha\Psi)^n]^{1-1/n}}, \quad (1)$$

with shape parameter n (-), α (m^{-1}) representing the inverse of the water entry pressure, θ_{res} (or smres; m^3m^{-3}) is the empirical residual soil moisture content (without physical meaning), and θ_{sat} (or smsat; m^3m^{-3}) is the saturated soil moisture content. In JULES parameter n is defined as b = 1/(n-1) and sathh=α^{-1} (m).

The Mualem equation computes the unsaturated hydraulic conductivity K:

$$K = K_{sat}\frac{\theta-\theta_{res}}{\theta_{sat}-\theta_{res}}\left[1-\left(1-\frac{\theta-\theta_{res}}{\theta_{sat}-\theta_{res}}^{1/(1-\frac{1}{n})}\right)^{1-\frac{1}{n}}\right]^2, \quad (2)$$

where K_{sat} (or satcon; $mm\cdot s^{-1}$) is the saturated hydraulic water conductivity.

The values of the Mualem-Van Genuchten parameters need to be defined by the user for each grid cell/point based on soil characteristics.

In JULES soil moisture directly interacts with transpiration (through root water uptake) and bare soil evaporation as described hereafter (see also Figure A1.1 of Appendix 1). JULES first computes the potential photosynthesis, which is a function of three limiting factors: Rubisco limitation, radiation limitation, and photosynthetic product transport limitation. The potential photosynthesis is multiplied with a soil moisture reduction factor to obtain the actual photosynthesis. To obtain this soil moisture reduction factor the model first computes a limiting factor for each layer:

$$\beta_i = \begin{cases} 1, & \theta_i \geq \theta_{crit} \\ \frac{(\theta_i-\theta_{wilt})}{(\theta_{crit}-\theta_{wilt})}, & \theta_{wilt} < \theta_i < \theta_{crit}, \\ 0, & x \leq \theta_{wilt} \end{cases} \quad (3)$$

where θ_i is the unfrozen soil moisture content in layer i, θ_{crit} is the critical point soil moisture content below which soil moisture is limiting the root water uptake (matrix water potential -330 cm in JULES), and θ_{wilt} is the wilting point soil moisture content below which no root water uptake occurs (-15000 cm in JULES). These reduction factors are multiplied with the root density in the layer. These weighted reduction factors are then summed to obtain the root zone soil moisture reduction factor. From the actual photosynthesis the plant stomatal conductance is computed. Separately the bare soil surface conductance, which is a function of the soil moisture content in the upper soil layer and the critical soil moisture, is computed. The surface conductance is then computed as a function of the stomatal conductance and the bare soil surface conductance.

The potential evapotranspiration is also calculated separately. This variable is multiplied with the saturated land fraction to compute the free water evaporation (e.g. lake and canopy evaporation). The rest of the potential evapotranspiration is multiplied with the surface conductance to obtain the bare soil evaporation + plant transpiration. Together these fluxes are the actual evapotranspiration (water) or latent heat flux (energy). The amount of water drawn from the top soil layer through bare soil evaporation depends on the bare soil surface conductance. The distribution of the root water uptake between the layers depends on the weighted soil moisture limitation factor for each layer. The water extraction from the soil in its turn directly affects the soil moisture content in the different layers at the start of the next time step. These soil moisture contents then affect the soil moisture redistribution, surface runoff, and deep drainage.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

JULES-UKV also requires a number of initial conditions: the amount of unfrozen water and snow stored on the surface (on canopy and on soil surface; set to zero in this study), snow properties (set to JULES-UKV defaults), the surface temperature (set to the air temperature of the hour before the first simulation time step), the soil temperature of each layer (set to the soil temperature from Ameriflux data the hour before initial time step) the soil water content in each layer (set to the soil water content from the PS observed moisture content of the hour before the first simulation time step and applied homogeneously throughout the profile). Soil moisture was spun up by running a maximum of five cycles and stopped when soil moisture convergence was lower than or equal to 10% compared to the previous cycle.

2.4.2 Calibration approach

At each site we calibrated JULES against observed PS observed soil moisture and against CRNS observed neutron counts respectively. We chose to calibrate simulated neutron counts against CRNS observed neutron counts because it allowed us to take into consideration vertical heterogeneity in modelled soil moisture by computing neutron counts from modelled soil moisture profiles using COSMIC. We computed the Root Mean Squared Error (RMSE) between simulated and observed hourly time series. To better match the observed soil moisture/neutron count time series, we calibrated five JULES parameters that influence the model soil moisture state (Figure 3). These included three Mualem-Van Genuchten parameters: b, sathh, and satcon and soil moisture limitation factor β parameters *smcrit* and *smwilt*. We chose these parameters because they are, in theory, directly linked to the movement of moisture in the soil and to the effects of soil moisture on transpiration in JULES.

To assess the effects of calibration on soil moisture and surface energy flux simulation we compared the calibrated solutions against a default run at each site. The parameter values for the default case were derived from soil properties (percentages clay, loam, and organic matter) reported by the Harmonised World Soil Database (HWSD; FAO, 2009) for each of the twelve sites. These properties were used in the Wösten Pedotransfer function (Wösten PTF; Wösten et al., 1997) to obtain values for b, sathh, and satcon. Parameter values for smwilt and smcrit were subsequently obtained with the Van Genuchten formula.

The parameter calibration ranges were the same for all sites (Table 2). They were constructed by computing the minimum and maximum parameter values for the entire soil texture triangle (based on Wösten PTF), while also taking into consideration three organic matter contents (yielding three triangles), but excluding clay percentages above 70% to avoid extreme values for parameter b especially. The range for smcrit was set to 10-90 % of the saturated soil moisture content. The saturated soil moisture content was computed as a function of the dry soil bulk density (bd): smsat =1 – bd / 2.65 (Brady and Weil, 1996). Soil bulk density values obtained from the COSMOS network were used. This thus yielded different smcrit ranges in terms of soil moisture content (m^3m^{-3}) for different sites. The residual soil moisture content (defined implicitly in JULES) was set to zero because the Wösten PFT does not consider it.

JULES' remaining two ancillary parameters; the soil heat conductivity and soil heat capacity, were for each site computed as a function of soil properties (HWSD) with De Vries' (1963) method. The bare soil albedo was set constant at 0.38 (-) for all sites. Plant Functional Type (PFT) parameters were set to JULES defaults except for the e-folding rooting depth (depth above

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

which 86 % of plant roots are present) and the canopy height, at sites where more specific information was available from Ameriflux/COSMOS site information or from site specific literature.

We calibrated JULES using the BORG Multi-Objective Algorithm (BORG-MOEA or BORG; Hadka and Reed, 2013). This calibration tool was designed for multi-objective problems but also works for single-objective calibration. BORG employs multiple optimisation algorithms simultaneously to obtain convergence while also keeping the searched parameter space wide. The algorithm measures progress with the epsilon-progress technique, which uses the Objective Function (OF) space divided in boxes with sides of size epsilon. Epsilon is a user defined value for each OF (we used epsilon values of 0.001 m^3m^{-3} for PS and 1 cph for CRNS). Only if a new solution resides inside a box with a better OF value, BORG considers it is progress. If no progress was obtained after 200 runs, the algorithm had stagnated. In this case the BORG algorithm triggers a restart, which consists of (among other techniques) changing population size to maintain a diverse population and to escape local optima. We used a maximum number of 3,000 runs and an initial population size of 100 runs.

2.4.3 Validation approach

As main validation metric we chose RMSE between the observed and simulated Evaporative Fraction EF = LE / (LE + H). The EF tells how the net surface radiation is partitioned between all LE and H fluxes. Additionally, we computed the RMSE between observed and simulated LE. Because data quality issues with EC data are often observed during at night time (Goulden et al., 1996; 2006; 2012; Aubinet et al., 2010), we computed these metrics over day time hours only. We defined day time as downward shortwave radiation > 20 Wm^{-2}. To avoid extreme RMSE-EF values, we used hours with both observed and simulated H and LE values ≥ 1 Wm^{-2} only. Otherwise a few hours with small LE or H values would have dominated the RMSE values, while this would probably have been due to forcing or EC data inconsistencies and would not relate to soil moisture temporal variability.

3 Results and Discussion

3.1 Data analyses

3.1.1 Time series analyses

In Figure 4 comparison between PS and CRNS soil moisture time series shows that the seasonal trends of the two soil moisture observation products were similar. The two soil moisture products however also differed from each other, in distinct ways at different sites. At UM PS was peakier (peak height up to 0.10 $m^3 \cdot m^{-3}$ compared to <0.05 $m^3 \cdot m^{-3}$). At DC PS dry down was more gradual compared to CRNS and reached soil moisture contents at least 0.01 m^3m^3 above PS SM. At SO PS dried down quicker and was then systematically drier than CRNS. At KE PS was peakier (peak height up to 0.20 $m^3 \cdot m^{-3}$ compared to <0.10 $m^3 \cdot m^{-3}$ for CRNS) and systematically wetter. ME was characterised by a slower dry down during summer for PS compared to CRNS. At SR PS was systematically higher (~0.25 $m^3 \cdot m^{-3}$) than CRNS. At CS PS SM dried down quicker and

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

then stayed at a systematically wetter soil moisture content (0.1 m^3m^{-3} higher) than CRNS soil moisture. At MM the seasonality was the main difference, with PS being up to 0.1 m^3m^{-3} drier in summer and 0.05 m^3m^{-3} wetter during winter. At TR PS was peakier and during wet periods (winter) systematically higher (up to 0.2m m^3m^3) than CRNS. At AR PS was mostly wetter. Also at WR PS was systematically wetter (up to 0.10 $m^3 \cdot m^{-3}$). During spring PS dried down slower than CRNS. At MO PS was wetter as well (up to 0.20 $m^3 \cdot m^{-3}$) but also showed a higher seasonality (up to 1.5 times larger amplitude) than CRNS.

PS SM observations were systematically higher than CRNS SM observations at eight of twelve sites. In addition, the CRNS data showed high frequency variation, especially clear at SO, MM, TR, WR, and MO. This was an effect of inherent randomness in neutron radiation reaching the CRNS sensor element. This effect was more pronounced for lower neutron intensity, when there was relatively more hydrogen prevalence in the surroundings of the CRNS.

3.1.2 Similarity metrics

The differences seen in Figure 4 are also summarised in Figure 5, which shows a gradual site to site increase in RMSD between observed PS and CRNS and soil moisture data series (RMSD-SM_{obs}). However, between WR and MO there is a sudden ~60% increase in RMSD-SM_{obs}. The systematic bias mentioned in the soil moisture time series analysis for SR reflects here in the large (~90%) contribution from the difference in mean soil moisture. The difference in seasonality mentioned in the time series analyses for MM reflects in the large contribution (~50%) from seasonality (i.e. standard deviation). The higher soil moisture content for PS at TR during winter shows up in the relatively large contribution (~60%) from seasonality. The slower dry down at DC is reflected by the high contribution (~70%) from dynamics (coefficient of correlation). Overall, bias contributed to 50% or more of the total error at seven out of twelve sites.

Scatter plots of RMSD-SM_{obs} against mean and standard deviation of observed PS and CRNS soil moisture content (not shown) revealed that the difference between PS and CRNS soil moisture content is larger at wetter sites. A similar plot, with RMSD-SM_{obs} against mean and standard deviation of precipitation (also not shown) showed a weaker correlation. No correlation was found between MSE-SM_{obs} decomposition and wetness conditions. Dominant vegetation type seemed not to have an effect on the similarity between the two soil moisture data products: forested sites included those with relatively small RMSD-SM_{obs} (UM, SO, ME) and those with relatively high RMSD-SM_{obs} (MM, WR, MO). The same holds for grass sites (KE, TR, AR). Only bare/shrub covered sites (DC, SR, CS) were all below the sites' average RMSD-SM_{obs} of 0.049 m^3m^{-3}. These four sites were however also relatively dry. No correlation between dominant land cover and decomposition was found. The grassy sites had the most even distributions between the three decomposition terms. Soil type and soil bulk density were also investigated for correlations with RMSD-SM_{obs}, but no trends were discovered.

Larger differences between PS and CRNS soil moisture could be expected at sites with more heterogeneous soil or vegetation. Static satellite photos of the sites from the COSMOS website did not indicate systematically more heterogeneous conditions at the sites where PS and CRNS soil moisture differed more. Site info (e.g. topography, presence of rocks) from COSMOS and Ameriflux did also not clearly show more horizontally heterogeneous soil properties for those sites.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Concluding, the differences between the two soil moisture products could not be clearly related to physical site characteristics besides the mean soil wetness.

3.1.3 Discussion of data analyses results

The fact that the soil moisture time series of PS and CRNS differed from each other in various ways could be related to a number of issues. First, PS sensor types and numbers of sensors differed between the sites (Figure A1 of Appendix A). Secondly, the installation methods may also have been different between the sites. Thirdly, the exact installation locations of the PS sensors may in certain cases have been for instance next to a macropore, or near roots, while at other sites they were coincidently installed in a homogeneous soil patch.

The presence of hydrogen pools other than soil moisture (e.g. biomass, intercepted water, and water in litter layer) also affect the observed CRNS neutron count. Because different hydrogen pools are more present at certain sites than at others, the uncertainty on neutron count observations varies between sites. The results did however not show effects of land cover and soil properties on the similarity between the two soil moisture products. Another factor is that, the quality of the calibration of COSMIC could possibly be different for different sites. Finally, at multiple sites, the PS and CRNS soil moisture time series were quite similar. This could be expected at rather homogeneous sites. Moreover, Kohli et al. (2015) suggested the CRNS footprint to be around 150-200 m instead of 300 m as reported by Desilets and Zreda (2013). In that case the differences between the two soil moisture observation techniques could be smaller than initially thought.

3.1.4 Limitations of the data comparison methodology

We derived vertically constant CRNS soil moisture values from observed neutron counts with COSMIC. This method contains inherent uncertainty because in reality soil moisture is often not vertically homogeneous. Comparing these derived CRNS soil moisture values with PS soil moisture data integrated to certain layers (0-10cm and 10-35cm) is therefore a deviation from field conditions. PS installation depths differed between sites and CRNS measurement depth varied between sites and over time. The comparison might therefore have been `more valid' at some sites than at others.

The different lengths (between sites) of the time series used can have affected the metrics and sites' ranking. However, we used at least one year of data at all sites.

For these reasons these data analyses should be seen as a first simple comparison of the two data types only. The outcomes of the calibration (against PS and CRNS) and validation provide insight in the effects of the differences between the two soil moisture products on JULES' surface energy flux partitioning and latent heat flux simulation.

3.2 Calibration

Calibration reduced the objective function (RMSE-SM) values in all cases (Figure 6). However, the degree to which this happened differed between sites and the two calibration strategies (PS/CRNS), with decreases of 21% (AR-PS) to 93% (UM-CRNS). While the errors of the default runs existed mostly of systematic bias, after calibration the difference in dynamics was

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

the largest source of uncertainty and in 16 out of 24 cases this contribution actually increased in absolute terms. This happened because peaks and valleys became less extreme after calibration. Higher peaks yield higher correlation (with scatter plot slope > 1) because the relatively higher peaks within a time series differ more clearly from smaller (noise) peaks (e.g. SR-PS, WR in Figure 7). When the peaks are then reduced by calibration, the high peaks do not stand out as much anymore and the correlation decreases.

The RMSE values reduced relatively more for the PS calibration than for the CRNS calibration. The calibration method could possibly explain this. CRNS calibration was against observed neutron counts, while PS calibration was against observed soil moisture contents. Because neutron counts have an inverse relationship with soil moisture content, the PS calibration was possibly governed by avoiding larger errors occurring during a few brief soil moisture peaks. While focussing on getting the fitting for those peaks right, the PS calibration neglected the smaller errors during dry periods. This would then result in relatively smaller decrease in RMSE values than for the CRNS calibrations because those were fitted with heavier weights to the drier periods.

Figure 6 also shows that the relative improvement was not systematically lower or higher for sites with higher similarity between the two observed soil moisture time series (actually the largest improvement was for the CRNS calibration at UM). The quality of the default runs was hence maybe simply determined by the quality of the chosen default parameter values.

Figure 7 shows soil moisture time series for four selected sites: UM was chosen because PS and CRNS soil moisture were most similar there, SR was a site with moderate difference, and at WR and MO PS and CRNS soil moisture were most different. The soil moisture time series approached the observations better in all cases, especially at UM and WR, where the default runs overestimated soil moisture contents. The simulated soil moisture dynamics at WR approached those of PS and CRNS observed soil moisture closely, while these differed from each other substantially.

3.3 Validation

While calibration errors decreased, the Evaporative Fraction estimation improved for eleven out of twenty-four calibrations (Figure 8a) and we observed similar results for latent heat flux (Figure A2.1 of Appendix 2). This basically means that an improvement in simulated soil moisture did not necessarily lead to better estimation of surface energy fluxes.

In fact, the magnitude of those improvements (EF: 3% to 30%; LE: 1% to 37%) were comparably to the deterioration from other calibration cases (EF: 2% to 28%; LE: 1% to 29%). The results also showed no systematically greater improvement across all sites for either of the two calibration strategies. Compared to the default case, CRNS calibration yielded lower RMSE-EF at four sites while PS calibration yielded better EF at eight sites. In Figure 8b 10% change in EF estimation was chosen to distinguish substantial change from non-substantial change, derived from the approximate error in EC sensible heat flux data (Finkelstein and Sims, 2001). It shows that in just three cases for PS calibration and three cases for CRNS calibration the improvement was actually substantial. At WR and CS, PS calibration improved EF most, while at these same sites CRNS calibration did not yield improvement. Finally, larger differences between the two observed soil moisture time series did yield

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

more different simulated surface energy flux time series at some sites (MO and WR compared to UM and DC) (e.g. UM compared to MO) while at other sites it did not (e.g. TR and AR compared to UM and DC)..

At sites DC, KE, SR, CS, MM, and AR integrated root zone soil moisture content ($SM_{root\ zone}$) was all the time (>93%, not shown) below the critical point soil moisture (smcrit). The surface energy partitioning time series and RMSE values of the other sites were additionally analysed over periods with $SM_{root\ zone}$ below the critical point. Only at SO did we see somewhat better performance during these periods compared to periods with $SM_{root\ zone}$>smcrit, after both PS and CRNS calibration.

We explored trends between relative improvement in surface energy flux estimation and soil wetness, precipitation, vegetation type, vegetation height, and soil characteristics. No clear trends were discovered. The only feature that could be distinguished was that for the two sites with a bare soil tile (DC and SR) PS and CRNS calibration improved EF and LE. Rooting depth did also not explain relative improvement in surface energy flux estimation.

In Figure 9 the monthly mean diurnal latent heat cycles of four sites (UM, SR, WR, and MO) are shown. The PS calibration for site UM yielded, compared to the default run, a higher overestimation of latent heat flux during winter and spring, while in summer and autumn no large change occurred. The CRNS calibration yielded a higher underestimation during summer. At SR, both PS and CRNS calibration improved latent heat flux during periods of low evapotranspiration, while during the other periods only the CRNS calibrations yielded better results. AT WR overestimation of LE decreased after PS calibration in all seasons, while CRNS calibration yielded too high LE, especially during summer. The PS calibration at MO yielded LE underestimation during summer while it improved during winter, whereas CRNS calibration did not change LE substantially during any season.

3.4 Were calibrated parameter values physically feasible?

If parameter values obtained after calibration were not physically feasible (e.g. representing a sandy soil while there was a clay soil) then, if model structure is assumed to represent biophysical processes sufficiently well, that could yield undesirable results. We investigated whether the parameter values obtained were `realistic' by seeing if they were substantially different from the soil conditions reported by HWSD (used to obtain default values), COSMOS, and Ameriflux. We analysed the calibrated parameter values from the single-objective calibrations for each site with the parallel coordinate plots in Figure A2.3 of Appendix 2. Calibration yielded values far from default for sensitive parameter b in two cases (MM-PS and TR-PS) only. The saturation hydraulic conductivity (satcon) took on substantially different values in eight cases. The critical point and wilting point soil moisture parameters assumed values far from the defaults in three and four cases respectively. The wilting point multiplier, which was actually calibrated, took a wide range of values, but the actual wilting point parameter stayed closer to the default values in most cases.

Although the parameter calibration range of sathh was rather wide compared to the range of soil types from HWSD at the twelve sites, the values were on, or very near to the boundaries at all sites for the default runs and in thirteen cases after the calibrations. The saturated hydraulic conductivity was on the edges for 22 runs. It should be noted however that, especially for

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

the saturated hydraulic conductivity, the parameter calibration range was non-linear, with a factor 1000 difference between the upper and lower boundaries.

COSMOS/Ameriflux soil information (Table 1) for UM suggests sandier conditions, which would yield higher saturation hydraulic conductivity, which was indeed obtained after PS calibration. At SO calibration yielded higher values for the saturation hydraulic conductivity and a lower value for b, which is in line with the thick organic layer occurring there. At the KE site, the soil was reported to be coarser (and also containing stones) than actually reported by the HWSD. This would mean higher saturation hydraulic conductivity, but a higher values was not obtained after calibration. At SR and MM the soil was reported to be of finer texture than the data from HWSD. Lower satcon values than the defaults were however not obtained. At TR a clay hardpan was found at 30-40 cm, which would impede drainage and calibration could hence be expected to yield lower satcon. This did however not happen after our calibrations. AR and WR were reported to be sandy, in contrast with the loam soils reported by HWSD. Higher satcon values were found after CRNS calibration at AR and after both calibrations at WR.

We could think of a range of reasons why apparently non-physically realistic parameter values were found, for instance:

- Horizontal and vertical soil heterogeneity (e.g. macro-pores, soil layering in organic matter and particle sizes, stones) was not (properly) accounted for in the model
- The vertical discretisation of the soil (layers of 10, 25, 65, and 200 cm) is not suitable for solving Richards' Equation (Beven and Germann, 2013)
- Richards' Equation applicability at larger horizontal scales is questionable for multiple reasons (Beven and Germann, 2013).
- JULES' bottom boundary condition (free drainage) may not have been suitable in all cases. We used Figure 3 from Fan et al. (2013) to get an idea of the water table depth. This did not indicate the presence of shallow groundwater tables (< 5m), which would have made free drainage unsuitable as bottom boundary condition. However, the resolution of the used map was rather coarse for our purpose. Moreover, other conditions, such as perched groundwater tables and shallow hard rock could be present and impede free drainage.
- The wilting point and critical point soil moisture contents were computed as a function of soil properties only; pressure heads of -16000 cm and -300 cm respectively were used regardless of vegetation type. More complex assumptions are however made in more complex models. For instance, a common soil water stress function, similar to the one used in JULES, is the Feddes function (Feddes et al., 1978). This function can take into account two different critical soil moisture content values; one for high potential transpiration and one for low potential transpiration. Moreover, the critical soil moisture pressure (-300 cm in JULES) can take different values for different vegetation types. Taylor and Ashcroft (1972) for instance reported values between -150 cm and -15000 cm for crops and grass. The results presented in Section 4.3 indicated these two parameters might govern the soil wetness importantly in JULES.
- Physically realistic parameters often provide unrealistic results (Gupta et al., 1998).

3.5 Two-objective calibration against soil moisture and latent heat flux

The single-objective calibration against PS and CRNS soil moisture did in thirteen of twenty-four cases not yield better surface energy fluxes and in only six cases was improvement substantial. This was not completely surprising because single-objective calibration is often insufficient to constrain parameters to simulate different states and fluxes well (Gupta et al., 1998). In addition, Vereecken et al. (2015) argued that calibrating soil hydraulic parameters against soil moisture only does not guarantee

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

better surface energy fluxes. To find out whether it was actually feasible to expect better surface energy flux simulation when soil moisture was improved we performed calibrations where we optimised the model for two objectives simultaneously. We employed the BORG algorithm to simultaneously optimise daytime latent leat flux RMSE and soil moisture RMSE (using PS and CRNS separately). Both optimisation experiments are defined as PS/LE and CRNS/LE.

The results from the PS/LE and CRNS/LE calibrations are shown in Figure 10 and Figure 11 respectively. From the results we learned that for five of the 12 PS calibrations and four of the 12 CRNS calibrations a substantially better LE estimation than obtained with the single-objective calibration could have been obtained while maintaining a similar soil moisture error (black points, representing all two-objective calibration runs, were present right below the single objective calibration solutions).

The shapes of the Pareto fronts did not suggest automatic improvement in LE estimation with improved soil moisture estimation except in a few cases (e.g. DC in higher RMSE-SM range and SO). To have obtained such automatic improvement, the lower edges of the point clouds should have shown positive correlations.

We plotted the compromise solutions (the runs which did relatively well for both RMSE-SM/N and RMSE-LE) on top (green triangles) to see if a two-objective calibration would have led to improved soil moisture and latent heat flux. This was the case for all PS/LE compromise solutions except at TR and for all CRNS/LE compromise solutions except at UM. This suggests that two-objective calibrations could have yielded improved soil moisture and latent heat flux. This is also clear from Figure 13, where the relative improvements for the compromise solutions with respect to the default solutions are shown.

The large differences in EF/LE simulation between the PS and CRNS single-objective calibrations at WR and MO seem, based on Figure 10 and Figure 11, mostly a coincidence of which best solution was chosen. At WR a CRNS-calibration solution with the same neutron count/soil moisture RMSE but better RMSE-LE could have been obtained, yielding LE performance similar to the PS-calibration. At MO this was true for the PS calibration. At other sites where we saw differences between PS and CRNS (e.g. MM), the Pareto fronts of Figure 10 and Figure 11 suggest these differences could as well have been quite limited. Figure 12 supports this conclusion because at all sites except SR (where the CRNS compromise solution was clearly better with respect to LE) differences in LE simulation between the compromise solutions were similar between PS/LE and CRNS/LE calibration.

Our results (Figure 10 and Figure 11; especially the rather horizontal lower edges of the point clouds in those figures), suggest that the coupling between soil moisture and latent heat flux was generally quite weak in JULES, even at sites where such coupling is expected to be relatively strong (e.g. DC, SR, KE, AR). Moreover, differences between surface energy flux simulation of PS and CRNS calibrations were minimal.

3.6 To which parameters were soil moisture and latent heat flux most sensitive?

We investigated if a change in some parameters had a relatively small effect on soil moisture, while at the same time having a large effect on latent heat flux. In such case calibration could yield better soil moisture, but the parameter value might be inappropriate for latent heat flux. The inverse (influential on SM but not on LE) could also occur. We explored this by

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

performing a sensitivity analysis with Morris' method (Morris, 1991) as implemented in the SAFE Toolbox (Pianosi et al., 2015) on the exact same parameter value ranges as used during the calibration. We computed the sensitivity indices (mean and standard deviations of the elementary effects) on the RMSE values (i.e. our Objective Functions; OFs) of simulated vs. observed PS soil moisture, simulated vs. observed CRNS neutron counts, and simulated vs. observed latent heat flux (day time only).

The results (shown for four sites in Figure A2.2 of Appendix 2) were consistent across most sites: all three OFs were most sensitive to changes in parameter b and least sensitive to the wilting point multiplier. Finch and Haria (2006) also found JULES parameter b to be most influential on soil moisture and latent heat, at a UK chalk site. The critical point soil moisture content was influential with respect to soil moisture / neutron counts but had at all sites less effect on latent heat flux. The lack of effect from the wilting point soil moisture content can probably be attributed to the use of the multiplier, despite being a common approach (Prihodko et al., 2008; Rosolem et al., 2012); a certain multiplier value could be good in combination with a certain value for the critical point but not for a different critical point value.

3.7 Could JULES model structure explain the limited improvement in surface energy flux estimation?

We found multiple causes for the lack of improvement in surface energy fluxes after calibration against soil moisture data and the quite weak coupling between JULES soil moisture and latent heat flux seen from the single and two-objective calibrations (PS-SM/LE and CRNS-N/LE). One cause was that even when root zone soil moisture changed considerably (not shown), soil moisture stress changed little (self-adjusting behaviour, Figure 13), and hence there was no considerable change in surface energy partitioning. We found this to be the main reason for KE, MO-CRNS, and TR-CRNS.

AR-CRNS, ME-CRNS, SR-PS, and SO-PS however did yield change in soil moisture stress (Figure 13), but other factors limited improvement in EF/LE. AR-CRNS, SR-PS, and SO-PS did yield different EF and LE timeseries, but because performance deteriorated during certain periods but improved during other periods, overall performance did not change. For instance, the CRNS calibration at AR yielded better EF and LE during wet periods with high LE, while during subsequent periods of drying the estimation was worse. EF and LE estimation after PS calibration at SR was better during dry periods while it was worse during wetter periods (monsoon). PS calibration at SO improved EF and LE during dry periods after summer while the simulation was worse in spring. Soil moisture stress at ME was relatively limited (beta mostly above 0.6) and during periods of soil moisture stress, the latent heat flux was dominated by different stress factors.

At UM, CS, MM, WR, and MO, EF and LE were substantially different between single-objective PS and CRNS calibrations. Figure 13 shows that in these cases the soil moisture stress was very different. At UM, CS, and MO, root zone soil moisture was similar (not shown) between the PS and CRNS calibration solutions, causing different stress due to different wilting point and critical point soil moisture, while at MM and MO, the root zone soil moisture actually changed considerably.

These results indicate the relatively large effect of the wilting point and critical point soil moisture parameters on the calibration results. The self-adjusting behaviour can yield similar soil moisture stress for different root zone soil moisture, while in other cases the wilting point and critical point values can be such that simulated soil moisture is close to the observations but root

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

zone soil moisture stress is different. With respect to the two-objective calibrations the self-adjusting behaviour contributed to the possibility of having highly different soil moisture but similar latent heat flux performance.

3.8 Discussion

Our results suggest a relatively weak coupling between soil moisture and evapotranspiration in JULES. In combination with the self-adjusting behaviour of wilting point and critical point soil moisture values this suggests that how soil moisture observations are used to calibrate JULES should be carefully considered beforehand. In the Land Surface Modeling, community it is known that the absolute value of soil moisture content has limited information content for the model (Dirmeyer et al., 2000; Koster et al., 2009). However, we found the wilting and critical point soil moisture parameters not only to move up or down after calibration, but also to move closer when the standard deviation in simulated root zone soil moisture content decreased. Our findings might also have implications for data assimilation if certain critical and wilting point soil moisture parameter values are chosen that do not match the magnitude of the assimilated soil moisture.

Finally, the potential differences in terms of contribution to model performance observed between the PS and CRNS sensors are undermined by the weak coupling in the current model structure of JULES. Namely, even at sites where the two observed soil moisture time series differed most, surface energy partitioning simulation did not differ substantially between PS and CRNS calibration. As mentioned in Section 2.3 the calibration approach differed for the two soil moisture products; we calibrated against PS soil moisture and CRNS neutron counts. We chose to calibrate against neutron counts because it is the way in which CRNS data could be used in the context of LSM when the COSMIC operator is used. If we had calibrated against representative CRNS soil moisture values, we might have found worse fits, more similar to those of the PS calibrations.

Previous research has indicated that soil moisture alone is insufficient to estimate soil hydraulic parameters (Vereecken et al., 2008). Vereecken et al. (2015) commented that even improved land surface models in combination with better soil moisture observations do not necessarily yield correct latent and sensible heat flux estimation. Our findings support this.

4 Conclusions

We investigated whether reducing the spatial scale mismatch between the surface energy flux data and soil moisture data could be reduced through the use of Cosmic-Ray Neutron Sensors (CRNS). Five soil-evapotranspiration parameters of LSM JULES were calibrated against Point Scale (PS) and CRNS soil moisture data separately, for twelve sites with different climate, land cover, and soil properties. Next, at each site, the improvement in surface energy partitioning and latent heat fluxes for the two calibration solutions was assessed by comparing the fit with EC data and a version of LSM JULES runs with default parameter values based on a widely used soil database. Before the calibrations we compared the observed soil moisture data from the two sensor types. These analyses showed the differences between PS and CRNS soil moisture happen in different ways at the investigated sites. While at certain sites there were mainly systematic biases, at other sites the seasonality was more different,

© Author(s) 2016. CC-BY 3.0 License.

while at other sites different time series dynamics were the main cause of differences between the two soil moisture observations. We found the difference between the two soil moisture products to be larger at wetter sites.

The calibration of JULES parameters did not necessarily yield an improvement in model surface energy partitioning and latent heat flux simulation. Single-objective (soil moisture) and multi-objective calibrations (soil moisture and latent heat flux) did not yield substantial differences between PS sensor and CRNS calibrations in simulating latent heat flux at eleven of the twelve sites. Moreover, at sites where the observed soil moisture time series from the two observation techniques diverged more, the differences between the resulting surface energy fluxes were not larger. The assumed benefits of a larger footprint of the Cosmic-Ray Neutron Sensor were not clear to JULES due to its weak coupling behaviour observed in this study. We found a few causes of the lack of improvement: (1) the wilting point and critical point soil moisture moved up or down in a similar way as the root zone soil moisture did, yielding similar transpiration for different soil moisture conditions; (2) calibration against soil moisture improved surface energy fluxes during certain periods, but deteriorated surface energy fluxes during other periods, yielding similar overall performance; (3) soil moisture was not the main surface energy partitioning controlling factor. Because our findings indicate that the coupling between model soil moisture and evapotranspiration in JULES is somewhat weak, we recommend to improve the representation of the related land surface processes at sub-kilometre scale; the spatial scale which is consistent with Cosmic-Ray Neutron Sensor and Eddy Covariance flux data.

Acknowledgements

This research was supported by the Queen's School of Engineering (University of Bristol) Ph.D. scholarship. Additional support for this work was also provided by the Natural Environment Research Council (A MUlti-scale Soil moisture-Evapotranspiration Dynamics study (AMUSED); grant number NE/M003086/1). Funding for Ameriflux data resources was provided by the US. Department of Energy's Office of Science. We would like to thank the investigators of the sites used for providing us with site information through personal communication. The authors would like to thank Jos van Dam for providing constructive comments.

References

Anderson, R.G., Goulden, M.L.: Relationship between climate, vegetation, and energy exchange across a montane gradient. J. Geophys. Res., 116, G01026, doi:10.1029/2010JG001476, 2011.

Aubinet, M., C. Feigenwinter, B. Heinesch, C. Bernhofer, E. Canepa, A. Lindroth, L. Montagnani, C. Rebmann, P. Sedlak, and Van Gorsel, E.: Direct advection measurements do not help to solve the night-time CO2 closure problem: Evidence from three different forests, Agric. For. Meteorol., 150(5), 655–664, doi:10.1016/j.agrformet.2010.01.016, 2010.

Baldocchi, D., Eva Falge, Lianhong Gu, Richard Olson, David Hollinger, Steve Running, Peter Anthoni, Ch Bernhofer, Kenneth Davis, Robert Evans, Jose Fuentes, Allen Goldstein, Gabriel Katul, Beverly Law, Xuhui Lee, Yadvinder Malhi,

© Author(s) 2016. CC-BY 3.0 License.

Tilden Meyers, William Munger, Walt Oechel, K. T. Paw, Kim Pilegaard, H. P. Schmid, Riccardo Valentini, Shashi Verma, Timo Vesala, Kell Wilson, and Steve Wofsy, 2001: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. Bull. Amer. Meteor. Soc., 82, 2415–2434, doi: 10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Menard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description - part 1: Energy and water fluxes. Geoscientific Model Development 4(3), 677-699, 2011.

Beven, K. J., and Cloke, H. L.: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al., Water Resour. Res., 48, W01801, doi:10.1029/2011WR010982, 2012.

Beven, K., and Germann, P.: Macropores and water flow in soils revisited, Water Resour. Res., 49, 3071–3092, doi:10.1002/wrcr.20156, 2013

Bloeschl, G., Sivapalan, M.: Scale issues in hydrological modelling: A review. Hydrological Processes 9 (3-4), 251{290. URL http://dx.doi.org/10.1002/hyp.3360090305, 1995.

Blyth, E. M., Dolman, A. J. and Wood, N.: Effective resistance to sensible- and latent-heat flux in heterogeneous terrain. Q.J.R. Meteorol. Soc., 119: 423–442. doi:10.1002/qj.49711951104, 1993.

Blyth, E., Gash, J., and Lloyd, A.: Evaluating the JULES Land Surface Model Energy Fluxes Using FLUXNET Data. Journal of Hydrometeorology, 11, 509-519, 2010.

Blyth, E., Clark, D.B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M, and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale. Geosci. Model Dev., 4, 255-269, doi: 10.5194/gmd-4-255-2011, 2011.

Bogena, H.R., J.A. Huisman, H. Meier, U. Rosenbaum, and A. Weuthen: Hybrid wireless underground sensor networks: Quantification of signal attenuation in soil. Vadose Zone J. 8. 3. 755-761, 2009.

Bogena, H. R., Huisman, J. A., Baatz, R., Hendricks Franssen, H. J., Vereecken, H.: Accuracy of the cosmic-ray soil water content probe in humid forest ecosystems: The worst case scenario. Water Resources Research 49, 5778-5791, doi:10.1002/wrcr.20463, 2013.

Brady, N.C., and Weil, R.R.: The nature and properties of soils (11th ed.). Prentice Hall, New York, 1996.

Budyko, M.I.: Тепловой баланс земной поверхности (Heat Balance of the Earth's Surface). Gidrometeoizdat, Leningrad. 255 pp, 1956.

Cardell-Oliver, R., Kranz, M., Smettem, K. and Mayer, K.: A reactive soil moisture sensor network: Design and field evaluation. International journal of distributed sensor networks, 1(2), 149-162, 2005.

Cavanaugh, M.L., Kurc, S.A., Scott, R.L.: Evapotranspiration partitioning in semiarid shrubland ecosystems: a two-site evaluation of soil moisture control on transpiration. Ecohydrology 4 (5), 671e681, 2011.

© Author(s) 2016. CC-BY 3.0 License.

Chen, X., Rubin, Y., Ma, S., Baldocchi, D.: Observations and stochastic modelling of soil moisture control on evapotranspiration in a Californian oak savanna. Water Resources Research, 44, W08409, doi:10.1029/2007WR006646, 2008.

Clark, D. B., Mercado, L. M. , Sitch, S., Jones, C.D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Menard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Blyth, E., Boucher, O. , Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description - part 2: Carbon fluxes and vegetation dynamics. Geoscientific Model Development 4(3), 677-699, 2011.

Cullen, M. J. P.: The unified forecast/climate model, Meteorol. Mag., 122, 81–94, 1993.

Crow, W.T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., Rosnay, P., Ryu, D., Walker, J. P.: Upscaling sparse groundbased soil moisture observations for the validation of coarse-resolution satellite soil moisture products. Rev. Geophys. 50 (RG2002), 1-20. URL http://dx.doi.org/10.1029/2011RG000372, 2012.

De Vries, D.A.: Thermal properties of soils. In D.A. De Vries, and Afgan, N.H. (eds.): Heatand mass transfer in the biosphere. I. Transfer processes in plant environment, 5-28, Washington, DC: Scripts, 1963.

Dharssi, I., Vidale, P. L., Verhoef, A., Macpherson, B., Jones, C., and Best, M.: New soil physical properties implemented in the Unified Model at PS18, Met Office Technical Report 528, Met Office, Exeter, UK, 2009.

Dirmeyer, P.A., Zeng, F.J., Ducharne, A., Morrill, J.C., and Koster, R.D.: The Sensitivity of Surface Fluxes to Soil Water Content in Three Land Surface Schemes. J. Hydrometeor., 1, 121–134, doi: 10.1175/1525-7541(2000)001<0121:TSOSFT>2.0.CO;2, 2000.

Dirmeyer, P.A.: The terrestrial segment of soil moisture-climate coupling. Geophysical Research Letters, 38 (16), 2011.

Dirmeyer, P.A., Jin, Y., Singh, B., Y, X.: Trends in Land-Atmosphere Interactions from CMIP5 Simulations. Journal of Hydrometeorology, 14, 829-849, doi: 10.1175/JHM-D-12-0107.1, 2013.

Evans, J. G., Ward, H. C., Blake, J. R., Hewitt, E. J., Morrison, R., Fry, M., Ball, L. A., Doughty, L. C., Libre, J. W., Hitt, O. E., Rylett, D., Ellis, R. J., Warwick, A. C., Brooks, M., Parkes, M. A., Wright, G. M. H., Singer, A. C., Boorman, D. B., and Jenkins, A.: Soil water content in southern England derived from a cosmic-ray soil moisture observing system – COSMOS-UK. Hydrol. Process., doi: 10.1002/hyp.10929, 2016

FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009. Harmonized World Soil Database (version 1.1). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Fan, Y., Li. H., and Miguez-Macho, G.: Global Patterns of Groundwater Table Depth. 339(6122, 940-943, doi:10.1126/science.1229881, 2013.

Feddes, R.A., Kowalik, P.J., and Zaradny, H.: Simulation of field water use and crop yield. Simulation Monographs. Wageningen: Pudoc, 1978.

Finch, J.W. and Haria, A: The representation of chalk soils in the JULES/MOSES soil hydrology model. Rept., NERC/Centre for Ecology and Hydrology, 2006.

Franz, T. E., Zreda, M., Rosolem, R., Ferre, T. P. A.: Field validation of a cosmic-ray neutron sensor using a distributed sensor network. Vadose Zone Journal 11 (4), doi:10.2136/vzj2012.0046, 2012.

© Author(s) 2016. CC-BY 3.0 License.

Goulden, M. L., Munger, J.W., Fan, S.M., Daube, B.C., and Wofsy, S.C.:, Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy, Global Change Biol., 2(3), 169–182, doi:10.1111/j.1365-2486.1996.tb00070.x, 1996.

Goulden, M. L., Winston, G. C., McMillan, A.M.S., Litvak, M. E., Read, E. L., Rocha, A. V., and Elliot, J. R.: An eddy covariance mesonet to measure the effect of forest age on land-atmosphere exchange, Global Change Biol., 12(11), 2146–2162, doi:10.1111/j.1365-2486.2006.01251.x, 2006.

Goulden, M.L., Anderson, R.G., Bales, R.C., Kelly, A.E., Meadows, M., and Winston, G.C.: Evapotranspiration along an elevation gradient in California's Sierra Nevada. Journal of Geophysical Research, 117, G03028, doi:10.1029/2012JG002027, 2012.

Goulden, M.L.: Measurement of Energy, Carbon and Water Exchange Along California Climate Gradients. Website: http://www.ess.uci.edu/~california/, last visited: July 2016. 2015.

Gupta, H.V., and Sorooshian, S.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4), 751-763, 1998.

Hawdon, A., D. McJannet, and J. Wallace.: Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia, Water Resources Research, 50(6), 5029-5043. doi:10.1002/2013WR015138, 2014.

Hadka, D., and Reed P.: Borg: an auto-adaptive many-objective evolutionary computing framework. Evolut. Compt., 21(2), 231-259, 2013.

Henderson-Sellers, A., McGuffie, K., Pitman, A.J.: The Project for Intercomparison of Land-Surface Parametrization Schemes (PILPS): 1992 to 1995. Clim. Dyn. 12, 849–859, 1996.

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., and Zacharias, S.: Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons. Water Resources Research 51, doi:10.1002/2015WR017169, 2015.

Koster, R.D., Dirmeyer, P.A., Guo, Z.C., Bonan, G., Chan, E., Cox, P., Gordon, C.T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y.C., Taylor, C.M., Verseghy, D., Vasic, R., Xue, Y.K., Yamada, T.: Regions of strong coupling between soil moisture and precipitation. Science 305, 1138–1140, 2004.

Koster, R.D., Guo, Z., Yang, R., Dirmeyer, P.A., Mitchell, K., Puma, M.J.: On the nature of soil moisture in land surface models. J. Climate 22 (16), 4322–4335, 2009.

Kurc, S. A. and E. E. Small: Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland, Water Resources Research, 43, W06416, doi:10.1029/2006WR005011, 2007.

Larson, K.M., Small, E.E., Gutmann, E.D., Bilich, A.L., Braun, J.J., Zavorotny, V.U.: Use of GPS receivers as a soil moisture network for water cycle studies. Geophysical Research Letters, 35, L24405, doi:10.1029/2008GL036013, 2008.

Larson, K.M., Braun, J.J., Small, E.E., Zavorotny, V.U., Gutmann, E.D., Bilich, A.L.: GPS Multipath and Its Relation to Near-Surface Soil Moisture Content. IEEE J. Sel. Topics in Appl., Earth Obs. And Rem. Sens., 3(1), 91-99, 2010.

© Author(s) 2016. CC-BY 3.0 License.

Manabe, S.: Climate and ocean circulation. Part I: the atmospheric circulation and the hydrology of the earth's surface. Mon. Weather Rev. 97, 739–774, 1969.

Morris, M.D.: Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33(2), doi:10.2307/1269043, 1991.

Mualem, Y.: A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resour. Res., 12(3), 513–522, doi:10.1029/WR012i003p00513, 1976.

ORNL-DAAC, 2015. Ameriflux Web Page. http://ameriflux.ornl.gov.

Pianosi, F., Sarrazin, F., and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis. Environmental Modelling and Software 70, 80-85, doi:10.1016/j.envsoft.2015.04.009, 2015.

Pitman, A.J.: The evolution of, and revolution in, land surface schemes designed for climate models. Int. J. Climatol. 23, 479–510, 2003.

Richter, H., Western, A. W., and Chiew, F. H. S.: The Effect of Soil and Vegetation Parameters in the ECMWF Land Surface Scheme. J. Hydrometeor., 5, 1131–1146, doi: 10.1175/JHM-362.1, 2004.

Ritsema, C.J., Kuipers, H., Kleiboer, L., Van Den Elsen, E., Oostindie, K., Wesseling, J.G., Wolthuis, J.W. and Havinga, P.: A new wireless underground network system for continuous monitoring of soil water contents. Water resources research, 45(4), 2009.

Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., Wendroth, O.: Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. Vadose Zone Journal 7 (1), 358{389. URL http://dx.doi.org/doi:10.2136/vzj2007.0143, 2008.

Rosolem, R., W. J. Shuttleworth, X. Zeng, S. R. Saleska, and T. E. Huxman: Land surface modeling inside the Biosphere 2 tropical rain forest biome, J. Geophys. Res., 115, G04035, doi: 10.1029/2010JG001443, 2010.

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Gonalves de Gonalves, L. G., Zeng, X.. Towards a comprehensive approach to parameter estimation in land surface parameterization schemes. Hydrological Processes 27 (14), 2075{2097, 2013b.

Rosolem, R., W. J. Shuttleworth, M. Zreda, T. E. Franz, X. Zeng, and Kurc. S. A.: The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System. J. Hydrometeor., 14, 1659–1671, doi: 10.1175/JHM-D-12-0120.1.

R Rosolem, T Hoar, A Arellano, JL Anderson, WJ Shuttleworth, X Zeng, TE Franz: Assimilation of near-surface cosmic-ray neutrons improves summertime soil moisture profile estimates at three distinct biomes in the USA. Hydrol. Earth Syst. Sci. Discuss., 11(5), 5515-5558, doi: 10.5194/hessd-11-5515-2014, 2014.

Running, S.W., Baldocchi, D.D., and Turner, D.: A global terrestrial monitoring network, scaling tower fluxes with ecosystem modelling and EOS satellite data. Remote Sensing, 70, 108-127, 1999.

Skott, R.L., Cable, W.L., Hultine, K.R.: The ecohydrologic significance of hydraulic redistribution in a semiarid savanna. Water Resources Research, 44 (2), 1990.

© Author(s) 2016. CC-BY 3.0 License.

Sellers, P. J., W. J. Shuttleworth, J. L. Dorman, A. Dalcher, and Roberts, J.M.: Calibrating the Simple Biosphere Model for Amazonian tropical forest using field and remote sensing data. Part I: Average calibration with field data. J. Appl. Meteor., 28, 727–759, 1989.

Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F., Berry, J., Collatz, G., Denning, A., Mooney, H., Nobre, C., Sato, N., Field, C., Henderson-Sellers, A.: Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. Science 275 (5299), 502–509, 1997.

Seneviratne, S. I. and Corti, T. and Davin, E. L. and Hirschi, M. and Jaeger, E. B. and Lehner, I. and Orlowsky, B. and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: A review, Earth-Sci. Rev., 99, 125—161, doi="10.1016/j.earscirev.2010.02.004, 2010.

Shuttleworth, J., Rosolem, R., Zreda, M., and Franz, T.: The COsmic-ray Soil Moisture Interaction Code (COSMIC) for use in data assimilation, Hydrol. Earth Syst. Sci., 17, 3205-3217, doi:10.5194/hess-17-3205-2013, 2013.

Smith, A. B., J. P.Walker, A. W.Western, R. I.Young, K. M.Ellett, R. C.Pipunic, R. B.Grayson, L.Siriwardena, F. H. S.Chiew, and H.Richter: The Murrumbidgee soil moisture monitoring network data set, Water Resour. Res., 48, W07701, doi:10.1029/2012WR011976, 2012.

Stöckli, R., Lawrence, D.M., Niu, G.Y., Oleson, K.W., Thornton, P.E., Yang, Z.L., Bonan, G.B., Denning, A.S., Running, S.W.: Use of FLUXNET in the community land model development. J. Geophys. Res 113 (G1), G01025, 2008.

Taylor, S.A., and Ashcroft, G.M.: Physical Edaphology. Freemand and Co., San Francisco, California, 434-435, 1972.

Trubilowicz, J., K. Cai, and M. Weiler 2009, Viability of motes for hydrological measurement, Water Resour. Res., 45, W00D22, doi:10.1029/2008WR007046, 2009.

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. America J., 44, 892–898, 1980.

Vereecken, H., Huisman, J.A., Bogena, H., Vanderborght, J., Vrugt, J.A., and Hopmans, J.W.: On the value of soil moisture measurements in vadose zone hydrology: A review. Water Resources Research, 44, doi:10.1029/2008WR006829, 2008.

Vereecken, H., Huisman, J.A., Hendricks Franssen, H.J., Bruggemann, N., Bogena, H.R., Kollet, S., Javaux, M., van der Kruk, J., and Vanderborght, J: Soil hydrology: recent methodological advaces, challenges, and perspectives. Water Resources Research, 51, 10.1002/2014WR016852, 2015.Appendices

Vinnikov, K. Y., Robock, A., Shuang, W., Schlosser, A.: Scales of temporal and spatial variability of midlatitude soil moisture. Journal of Geophyisical Research 101, 7163{7174. URL http://dx.doi.org/10.1029/95JD02753, 1996.

Vivoni, E.R., Moreno, H.A., Mascaro, G., Rodriguez, J.C., Watts, C.J., Garatuza-Payan, J. and Scott, R.L.: Observed relation between evapotranspiration and soil moisture in the North American monsoon region. Geophysical Research Letters, 35(22), 2008.

Wood, E.F., et al.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, W05301, doi:10.1029/2010WR010090, 2011.

© Author(s) 2016. CC-BY 3.0 License.

Wösten, J. H. M.: Pedotransfer functions to evaluate soil quality. In Gregorich, E. G. and Carter, M. R. (eds.) Soil Quality for Crop Production and Ecosystem Health. Elsevier, Amsterdam. pp. 221–246, 1997.

Zreda, M., Desilets, D., Ferre, T. P. A., Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic- ray neutrons. Geophysical Research Letters 35, 1-5, doi:10.1029/2008GL035655, 2008.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., Rosolem, R.: Cosmos: the cosmic-ray soil moisture observing system. Hydrology and Earth System Sciences 16, 4079-4099, doi:10.5194/hess-16-4079-2012, 2012.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Table 1: Site characteristics. Altitude from COSMOS website, land cover percentages from Ameriflux and publications. Harmonised World Soil Database (HWSD) data was used to define default model parameter values (here only soil categories shown).

Site	Altitude (meter above sea level; COSMOS)	Land cover (%)		HWSD dominant soil type	Site soil info (Ameriflux, COSMOS, literature)	Data sources
		Dominant	Remaining			
UM	220	100% broadleaf		Loamy sand	Deep well drained soils	COSMOS, ORNL-DAAC (2015)
DC	1300	46% shrubs	46% bare, 8% needleleaf	Sandy loam		Anderson and Goulden (2012)
SO	1160	63% needleleaf	37% shrubs	Loam	0-10 cm thick organic litter layer, bedrock at 1-2 m	COSMOS, Goulden et al. (2012)
KE	1531	100% C4-grass		Loam	Coarse-loamy, limestone fragments	ORNL-DAAC (2015)
ME	1253	100% needleleaf		Loamy sand	Sandy, minimal organic	COSMOS, ORNL-DAAC (2015)
SR	989	76% bare	24% shrubs	Loam	Silty clay loam	COSMOS, Cavanaugh et al. (2011)
CS	457	100% shrubs		Loam		Anderson and Goulden (2012)
MM	288	100% broadleaf		Loam	Well drained clay loam	COSMOS, ORNL-DAAC (2015)
TR	177	60% C3-grass	40% broadleaf	Loam	Sandy clay loam, clay hardpan at 30-40 cm	COSMOS, ORNL-DAAC (2015), Chen et al. (2008)
AR	314	100% C4-grass		Loam	Sandy	COSMOS, ORNL-DAAC (2015)
WR	371	100% needleleaf trees		Loam	5-10 cm organic layer, silty sand	COSMOS, ORNL-

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

					DAAC (2015)
MO	219	100% broadleaf trees	Loam	Silty loam	COSMOS, ORNL-DAAC (2015)

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Table 2: Calibrated parameter definitions and calibration ranges.

JULES parameter name	Unit	Role	Range	
			Minimum	Maximum
b	-	Mualem-Van Genuchten parameter (b=1/(n-1))	0.63	24.43
sathh	m	Mualem-Van Genuchten parameter (sathh=α^{-1})	0.09	28.01
satcon	mm s^{-1}	Mualem saturated hydraulic conductivity	$3 \cdot 10^{-5}$	$4.3 \cdot 10^{-1}$
smcrit	m^3m^{-3}	Critical soil moisture content	0.1·smsat	0.9·smsat
smwilt	m^3m^{-3}	Wilting soil moisture content	0.1·smcrit	0.9·smcrit

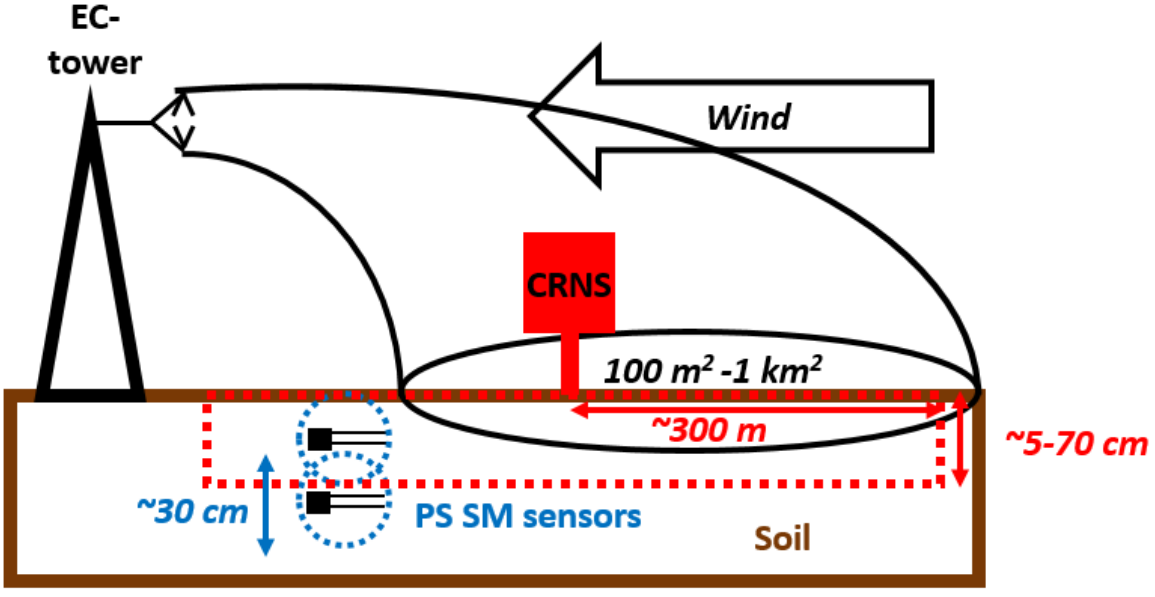


Figure 1: Schematic depicting the scale mismatch between Point Scale (PS) soil moisture sensor footprint (highlighted in dashed blue) and Eddy-Covariance tower footprint (highlighted with solid black circle on the ground). The Cosmic-Ray Neutron Sensor (CRNS; footprint highlighted in dashed red) may help fill the gap this scale mismatch leaves.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

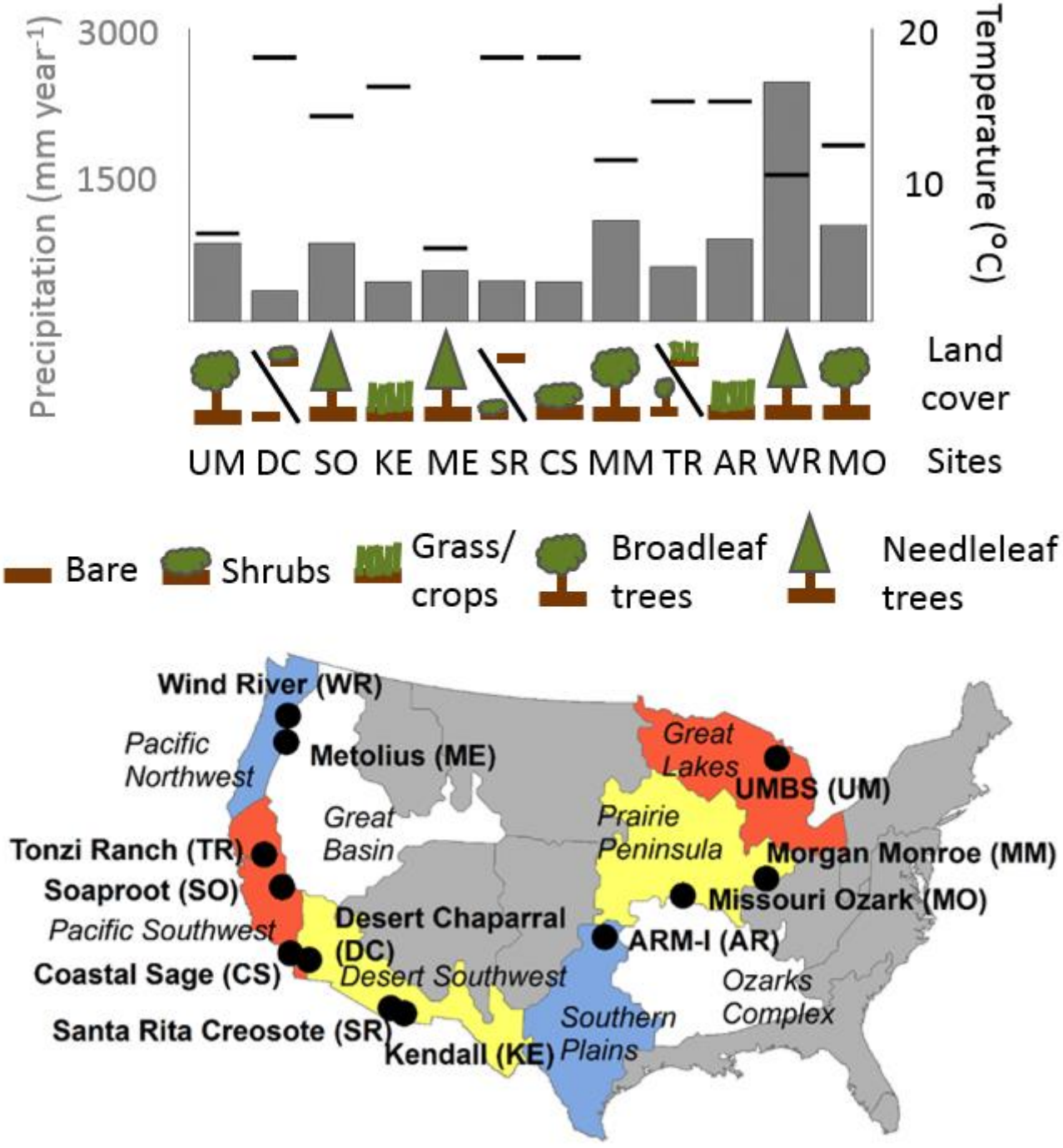


Figure 2: The upper figure shows the yearly mean precipitation, air temperature, and dominant land cover types for the twelve Ameriflux/COSMOS sites used. At sites DC, SR, and TR two different land cover types were shown because they covered similar areas in size. The map below shows the locations of the twelve sites within eight NEON Ecoclimatic Domains. Data sources: COSMOS, ORNL-DAAC (2015), Goulden et al. (2012), Anderson and Goulden (2012), Scott et al. (1990), Chen et al. (2008).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

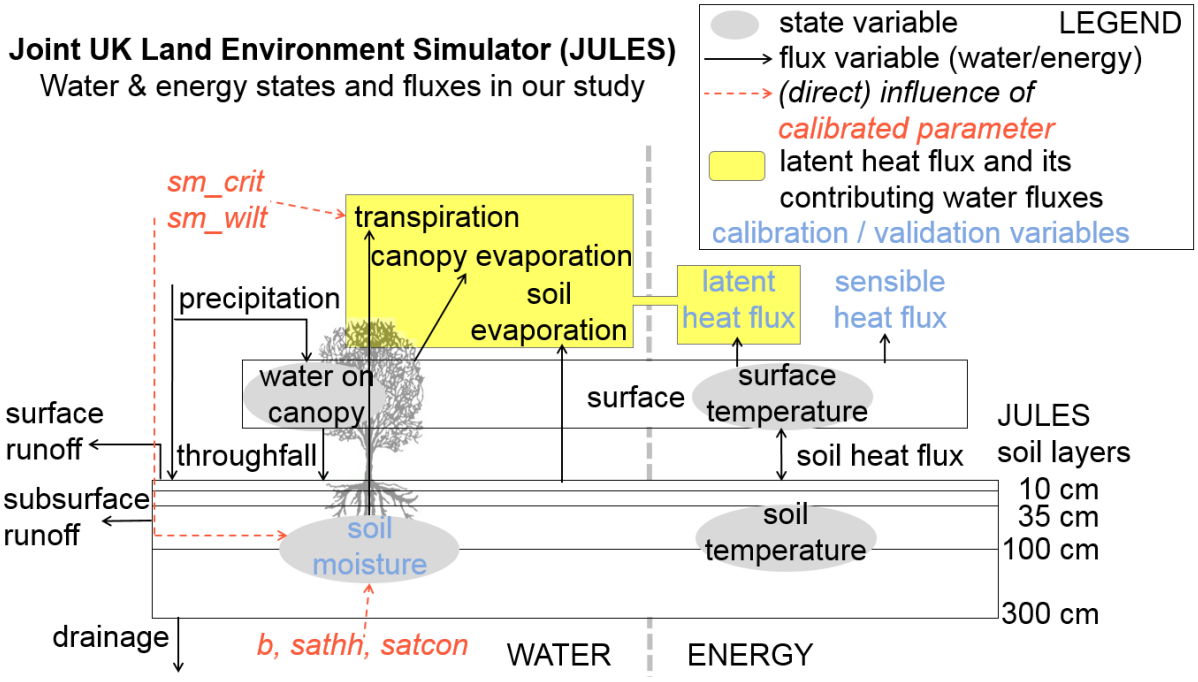
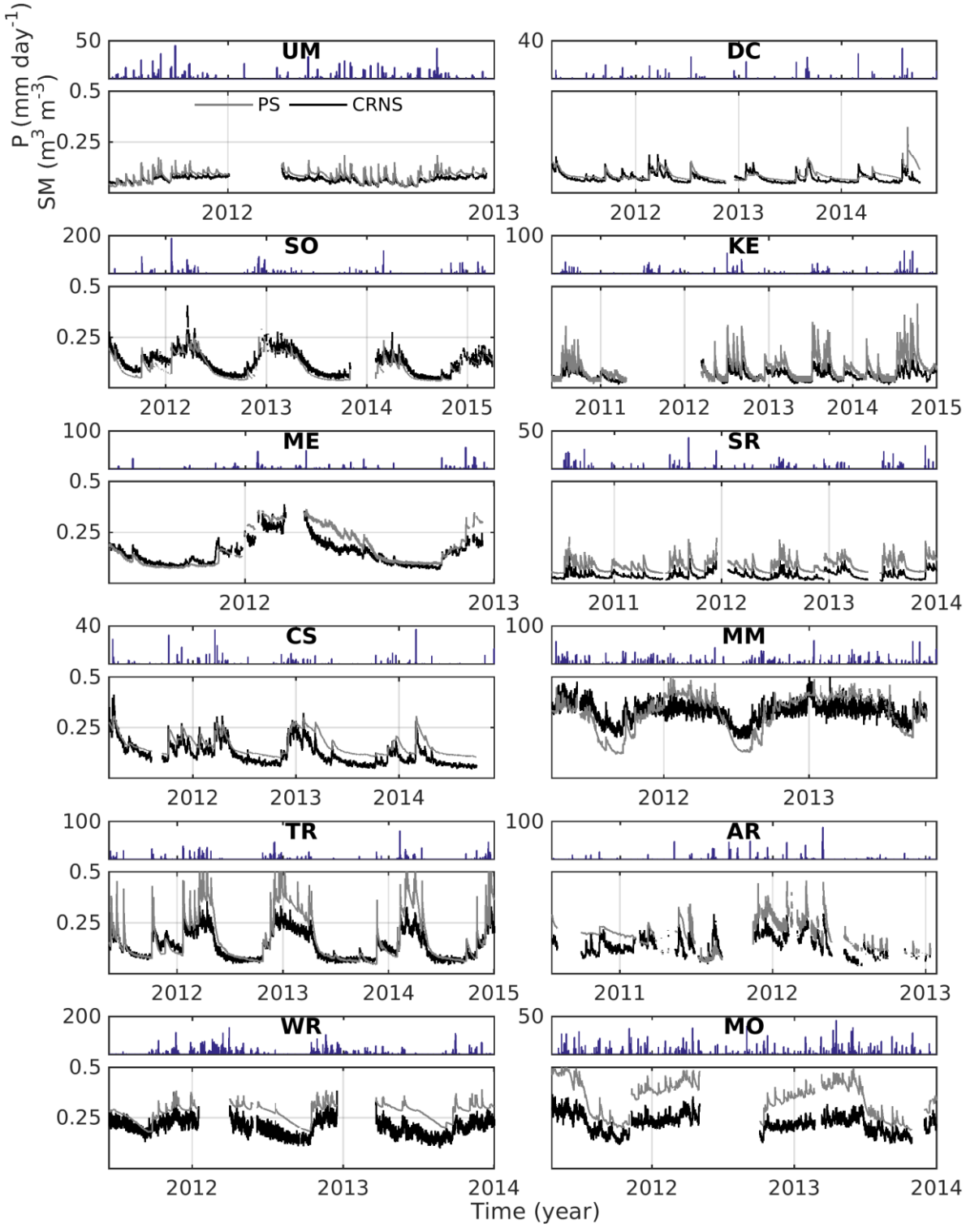


Figure 3: JULES water and energy states and fluxes most relevant to our study. Calibrated parameters and their direct effects shown. Please, refer to Table 2 for detailed description of parameters.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions



Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

Figure 4: PS and CRNS observed soil moisture time series for the twelve study sites. Notice the PS soil moisture time series have been linearly interpolated from individual measurement depths to the corresponding JULES soil layers. CRNS soil moisture was obtained using COSMIC while assuming vertically homogeneous soil moisture.

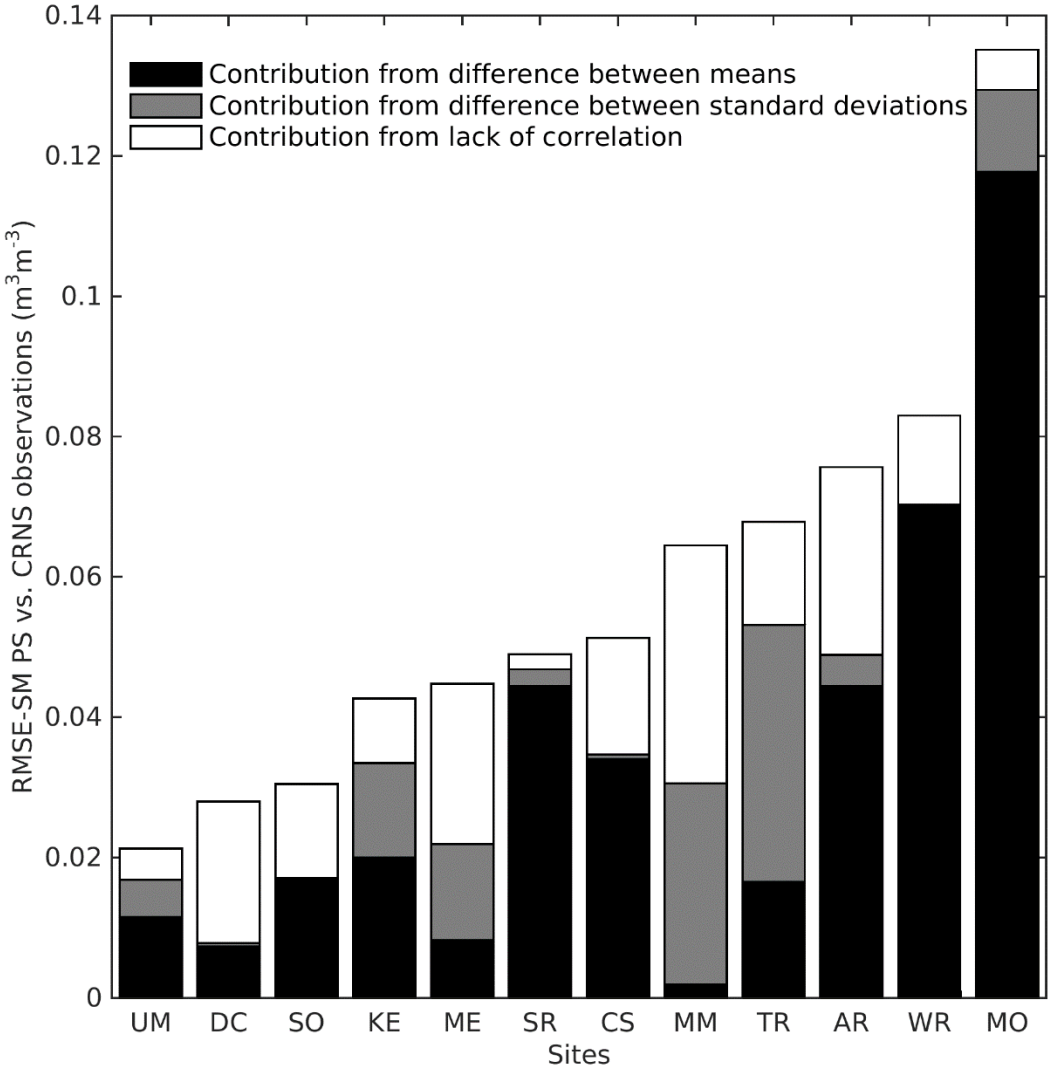


Figure 5: Root Mean Squared Error (RMSE) between observed PS and CRNS soil moisture. MSE decomposition (Gupta et al., 2009) was calculated and the fractions were then applied to the RMSE values. Sites are ranked from most similar to most different.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

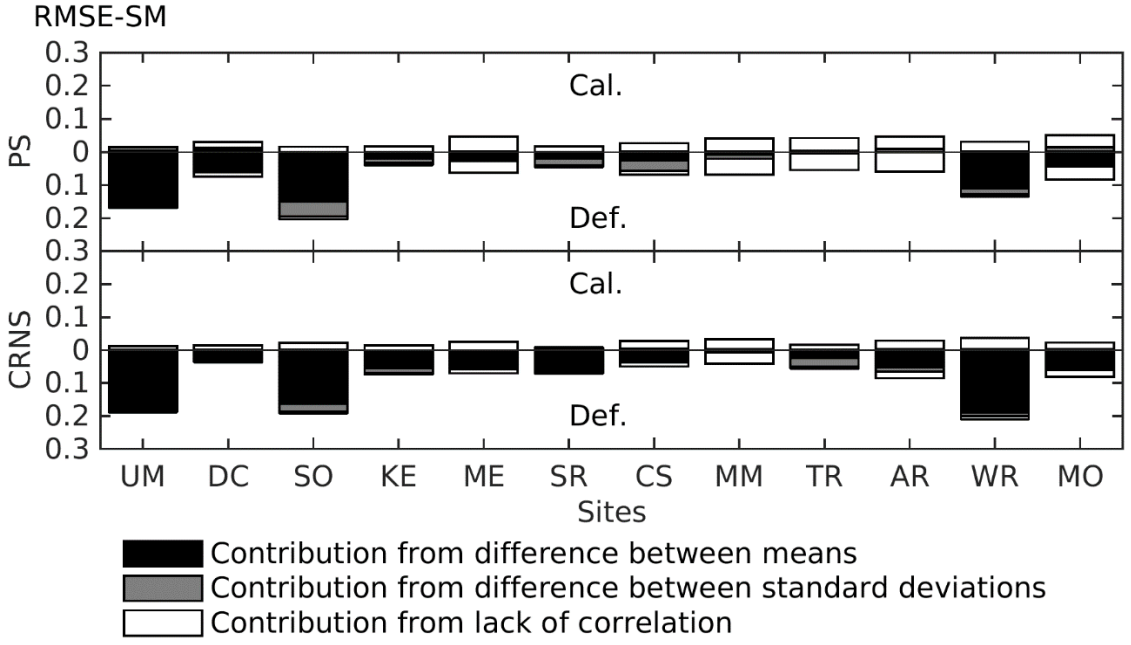


Figure 6: Objective function (Root Mean Square Error; RMSE (m^3m^{-3})) values between observed PS/CRNS soil moisture and JULES simulated soil moisture. For each calibration the RMSE of the default run is shown from the horizontal axis down, and the result after calibration is shown in the upward directions. The different error contributions from the MSE decomposition are shown as stacked bar plots.

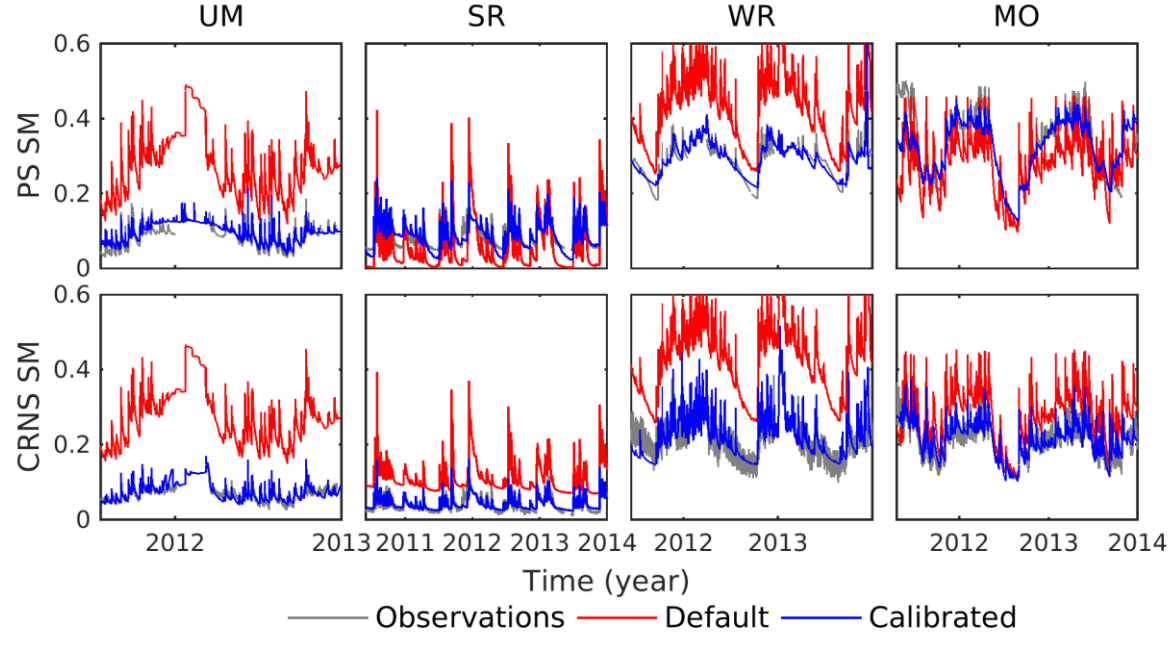


Figure 7: Hourly soil moisture time series of default and PS/CRNS calibrated runs for four of twelve sites.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

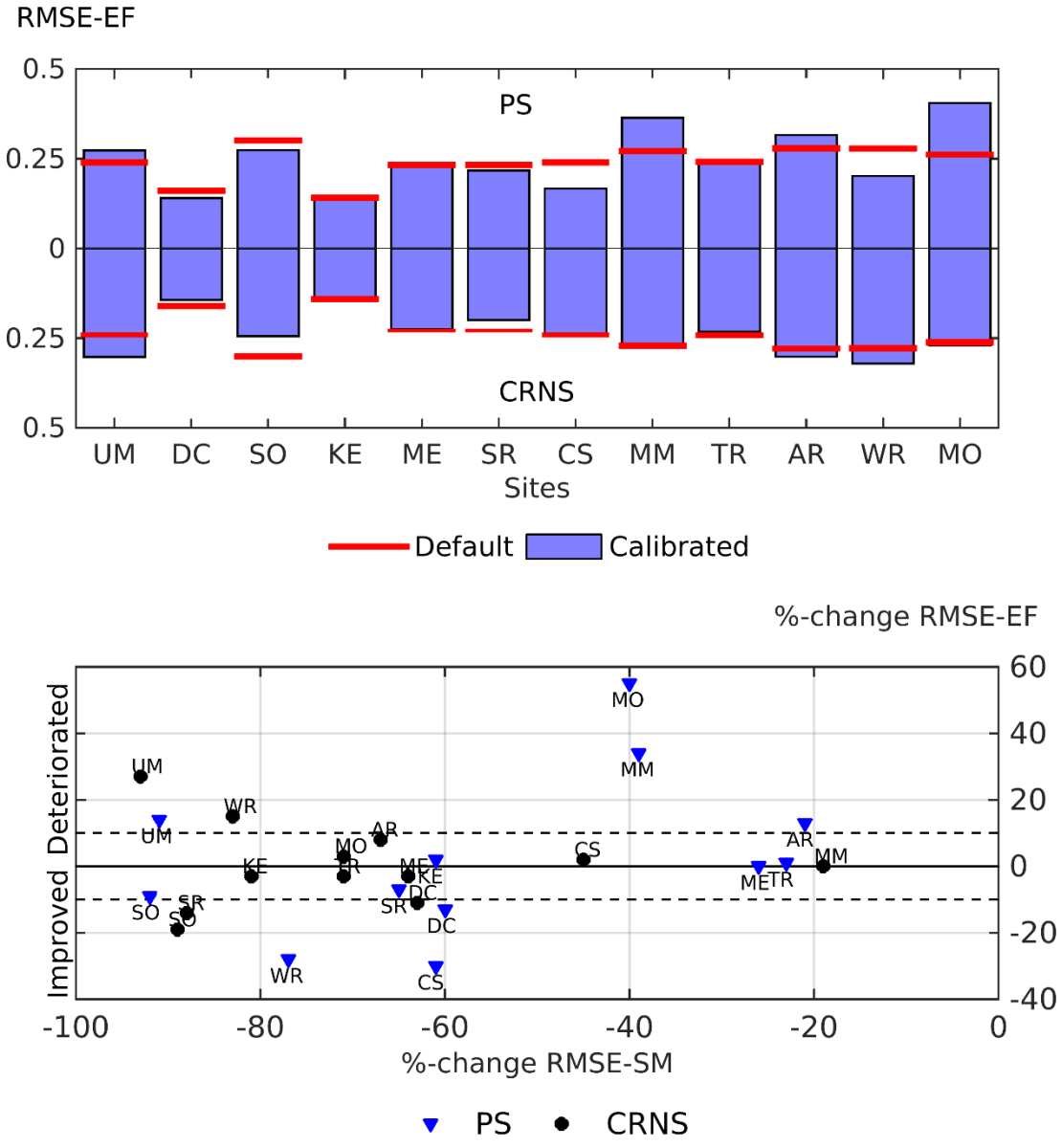


Figure 8: 8a) Root Mean Square Error (RMSE) values between observed and simulated Evaporative Fraction EF=LE/(LE+H), per site, for default and calibrated (PS/CRNS) runs. 8b) Relative change in RMSE-EF and RMSE-SM after calibration plotted against each other.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

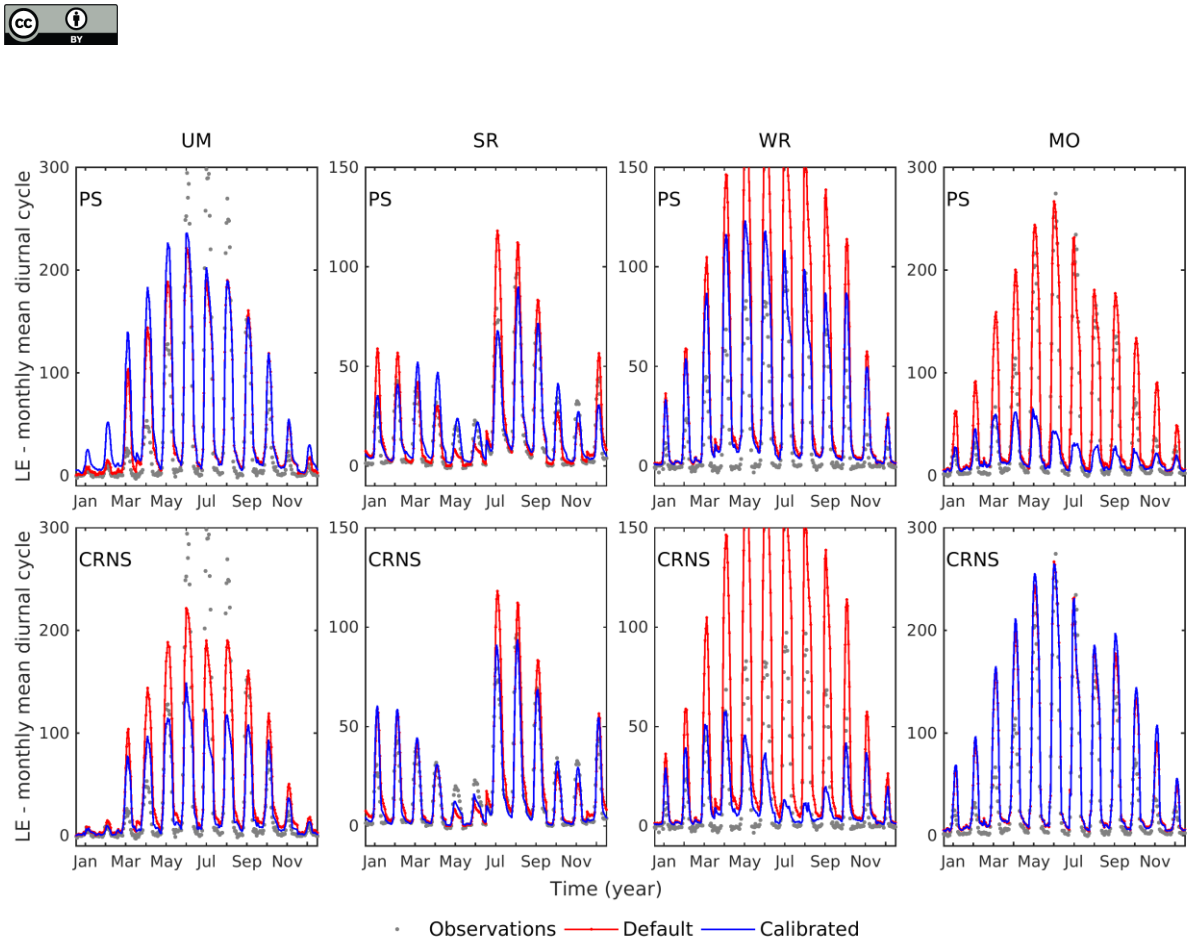


Figure 9: Monthly mean diurnal latent heat flux (LE) cycles. The upper row contains PS calibrated solutions, the lower row shows the CRNS calibrated solutions. Monthly mean diurnal latent heat flux cycles. The upper row contains PS calibrated solutions, the lower row shows the CRNS calibrated solutions.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

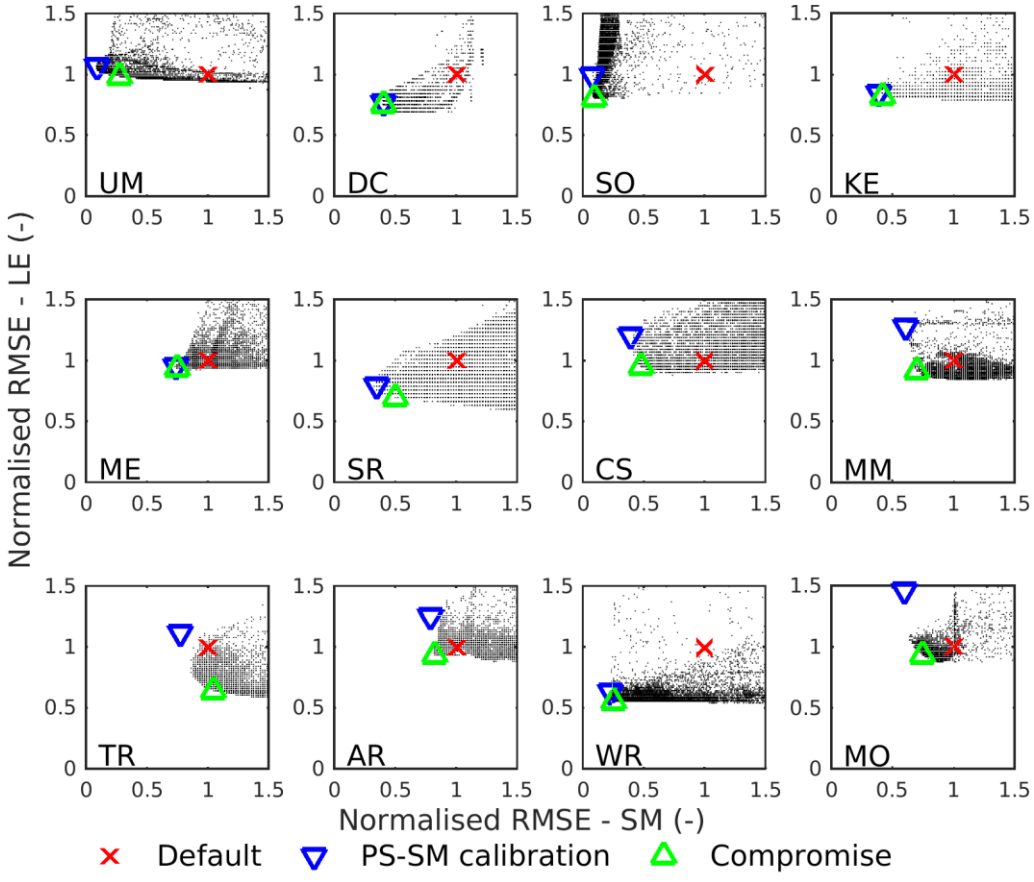


Figure 10: All runs from the two-objective calibrations against PS soil moisture and day time hourly latent heat flux (LE) at each site. Default run, PS-SM single objective calibrated run, and compromise solution runs are shown on top. All values were normalised to the values of the Default run. The plots were zoomed in to the area of 0 to 1.5 times the Default values.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

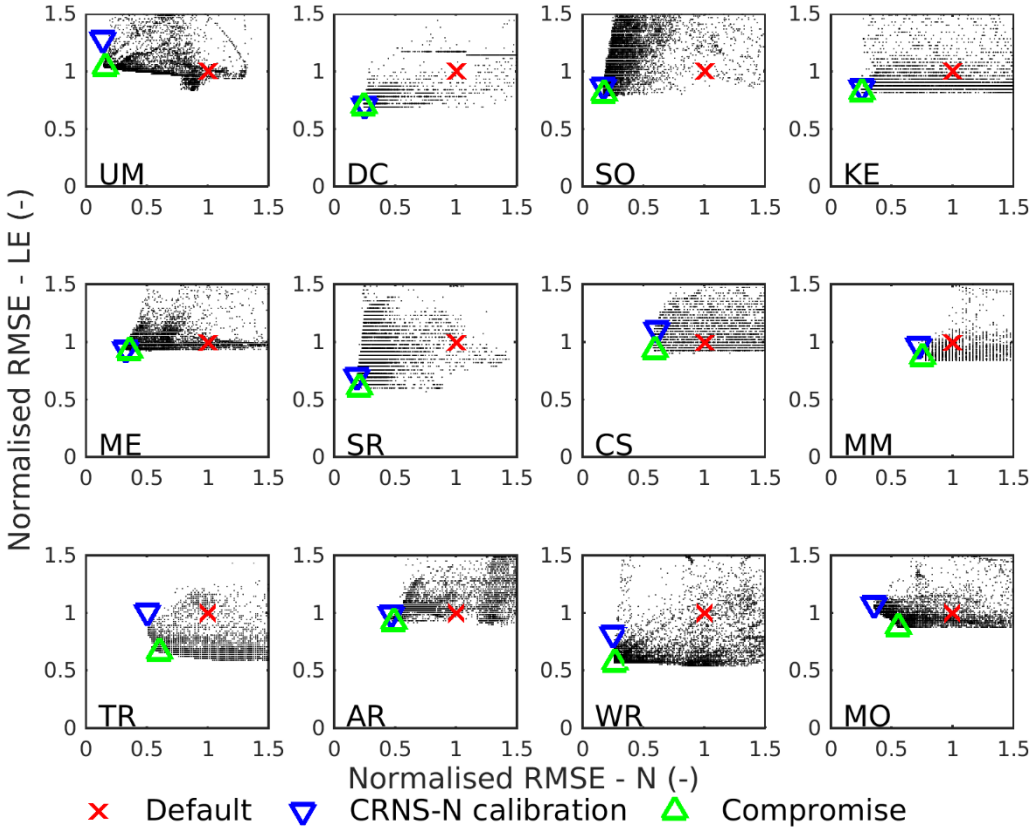


Figure 11: Same as Figure 10 but for the two-objective calibrations against CRNS neutron counts (representing soil moisture) and day time hourly latent heat flux (LE) at each site.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

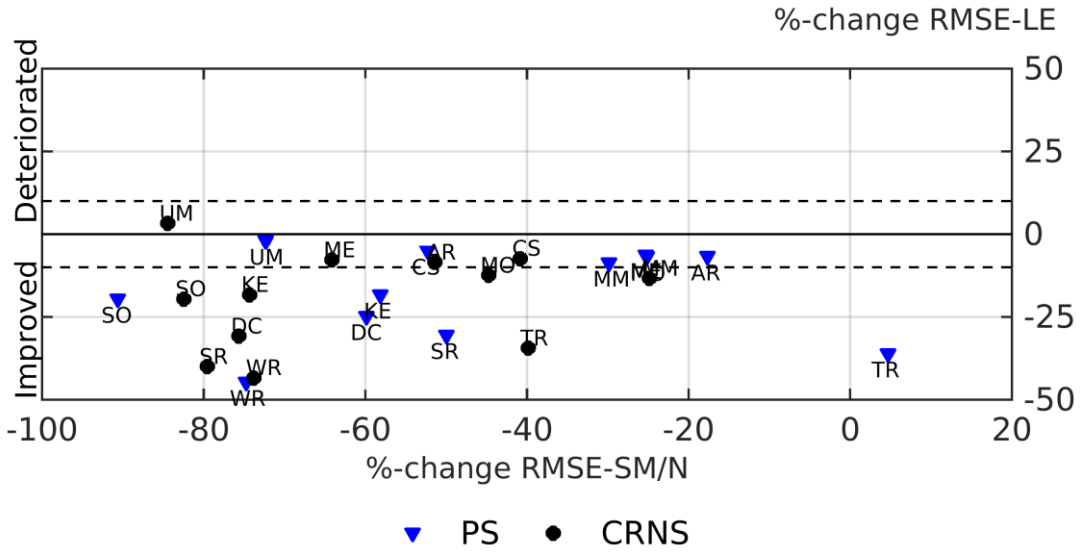


Figure 12: Relative change in performance (RMSE) of multi-objective compromise solutions compared to JULES simulations with default parameter values. The relative change in RMSE of soil moisture (SM; for PS) and neutron counts (N; for CRNS) are shown on the horizontal axis. The relative change in RMSE of latent heat is shown on the vertical axis.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

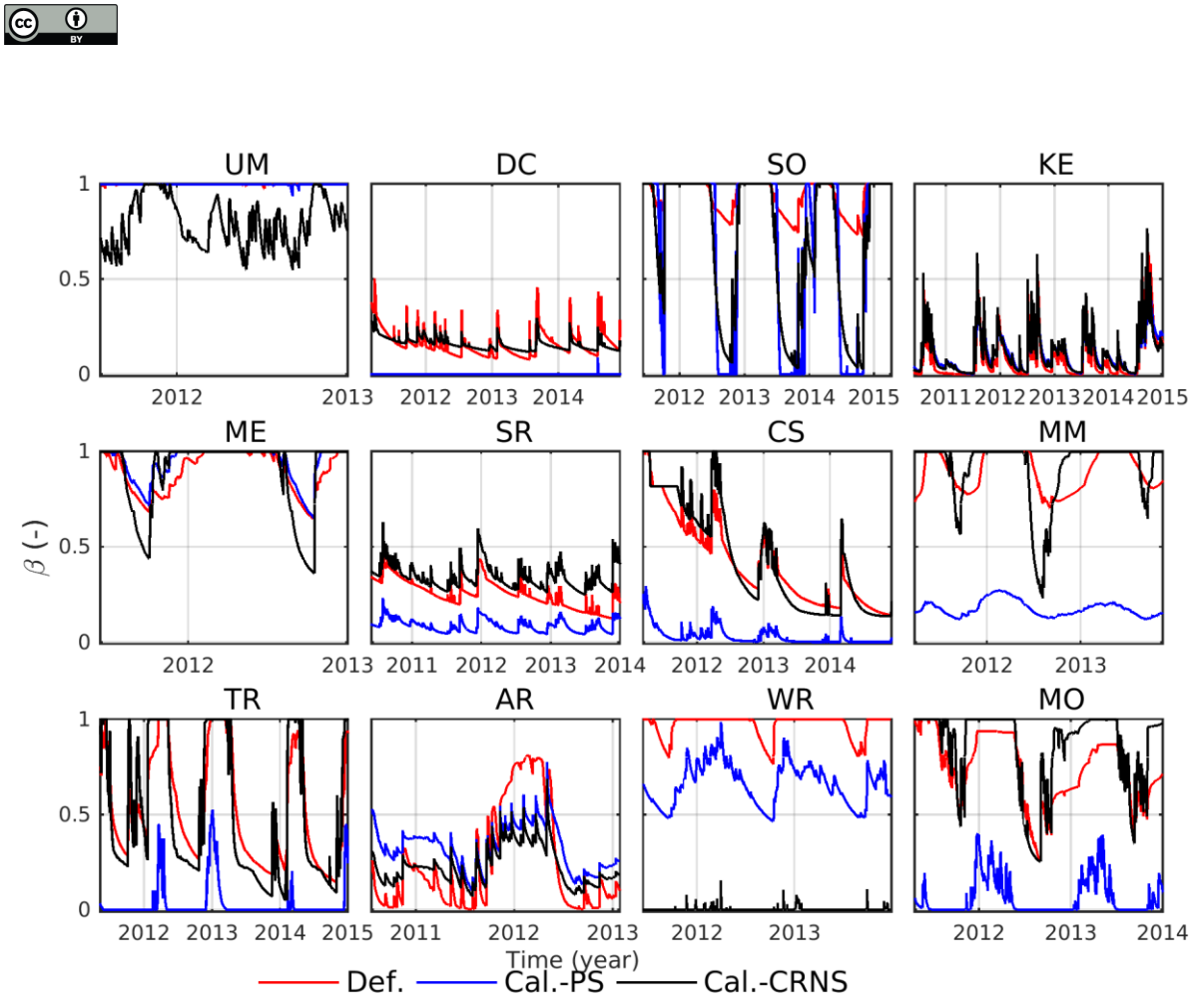


Figure 13: Soil moisture stress factor beta timeseries for all default runs and single objective (SM) calibration solutions.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

Appendices

Appendix 1: Additional Data and Methods tables and figures

Table A1.1: PS types, installation depths, number of profiles. In the last column it is shown to which JULES layer the observations were linearly interpolated.

Site	Type	Installation layers (cm below surface)	Number of profiles	Interpolated to JULES layer (cm)
UMBS	Campbell Scientific CS615 / CS616 (reflectometer)	0-30 average	1	0-10
Desert Chaparral UCI	Campbell Scientific CS616 (reflectometer)	0-30 average	4	0-10
Soaproot	Campbell Scientific CS615 (reflectometer)	0-30 average	4	0-10
Kendall	Stevens Water Hydra Probe	5,15	1	0-10
Metolius	Campbell Scientific CS615 (reflectometer)	0-30 average	1	0-10
Santa Rita Creosote	TDR	2.5,12.5	6	0-10
Coastal Sage UCI	Campbell Scientific CS616 (reflectometer)	0-30 average	4	0-10
Morgan Monroe	Campbell Scientific CS615 (reflectometer)	0-30 average	1	0-10
Tonzi Range	ThetaProbe (ML2)	0,(20)	2	0-10,(10-35)
ARM-1	Decagon Echo2 EC-20 (capacitance)	10,25	2	10-35
Wind River	Campbell Scientific CS615 (reflectometer)	0-30 average	1	0-10
Mozark	Delta-T	10	1	0-10

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

Table A1.2: Data used per site for downward longwave radiation (L_{win}), net radiation (R_{net}), and atmospheric pressure (P_{atm}). A is Ameriflux Level 2, COSMOS is COSMOS website, and NLDAS is National Land Data Assimilation Systems Forcing data Phase 2 (http://disc.sci.gsfc.nasa.gov/uui/datasets?keywords=NLDAS).

Site	$L_{w,in}$	R_{net}	P_{atm}
UM	A		A
DC	A		A
SO	A		A
KE		A	COSMOS
ME	NLDAS		COSMOS
SR	A		A
CS	A		A
MM	NLDAS		A
TR	NLDAS		COSMOS
AR	A		A
WR	A		A
MO	A		A

5

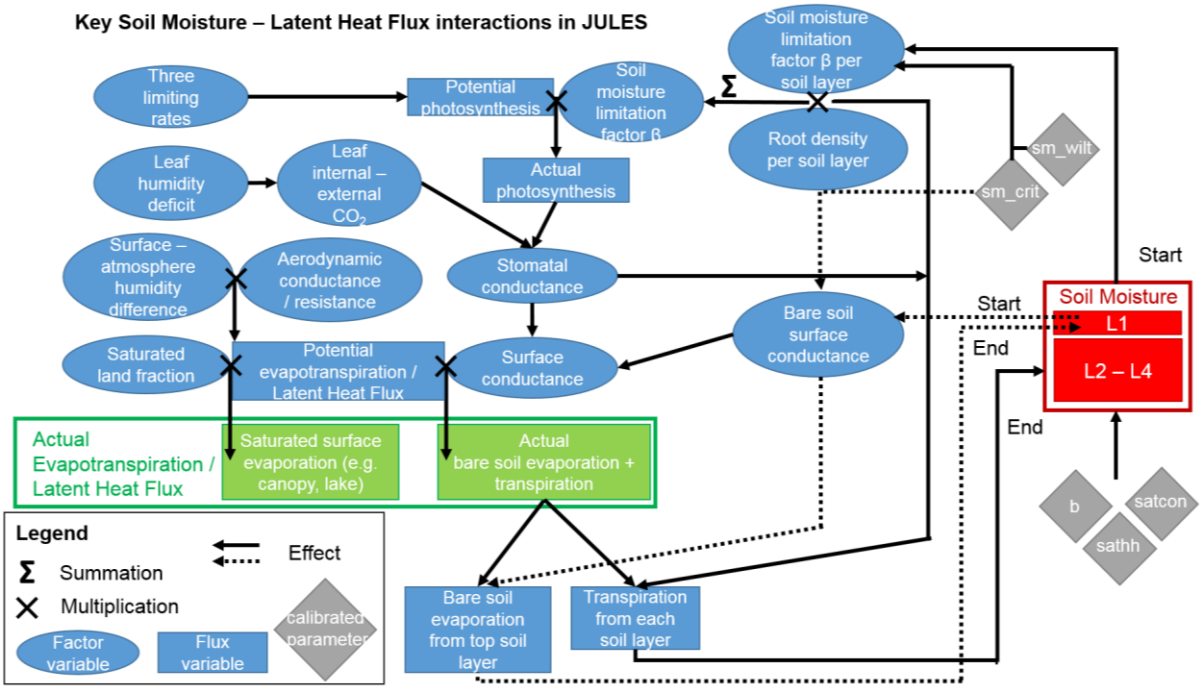


Figure A1.1: Key interactions between soil moisture and latent heat flux in JULES. The parameters calibrated in this study are indicated.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Discussions

Appendix 2: Additional Results and Discussion tables and figures

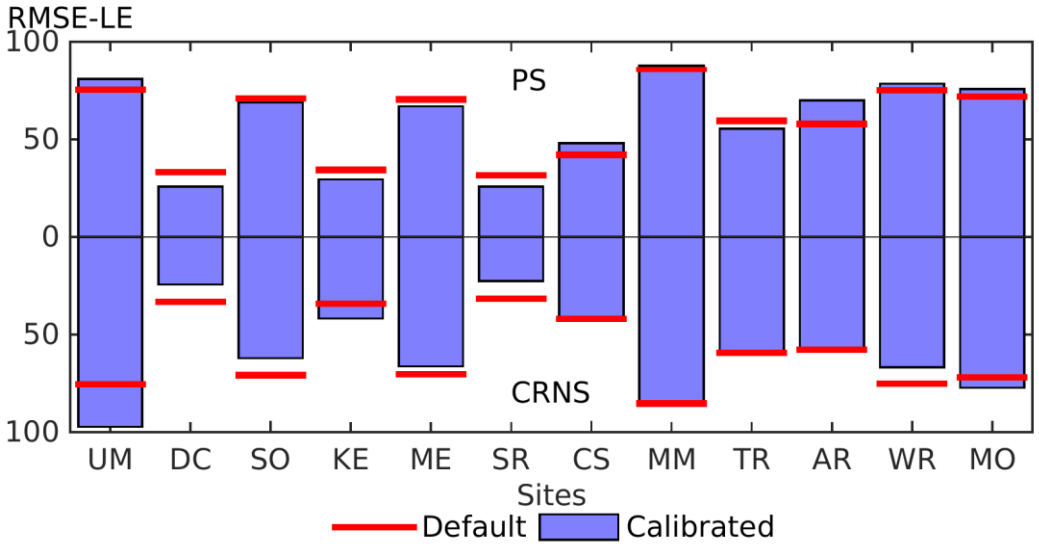
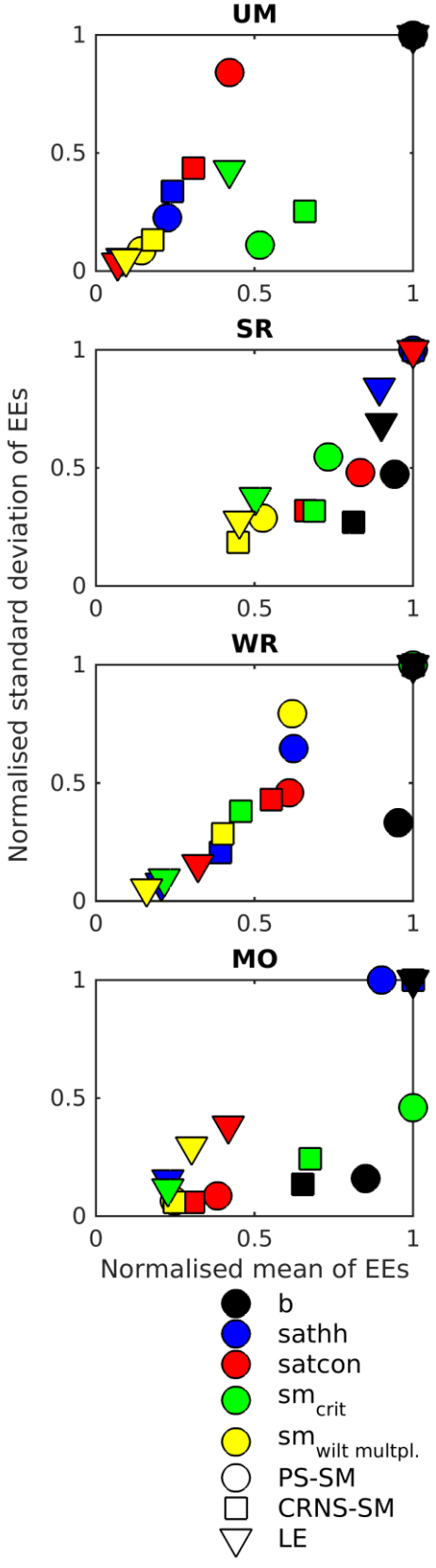


Figure A2.1: Root Mean Squared Error of simulated versus EC observed latent heat flux (w m^{-2}). The metric was computed over hourly day time data only. The upward bars represent PS-calibrated solutions, the downward bars represent CRNS-calibrated solutions. The Default runs are represented by the red horizontal lines, displayed twice for each site, to compare with both calibration strategies.

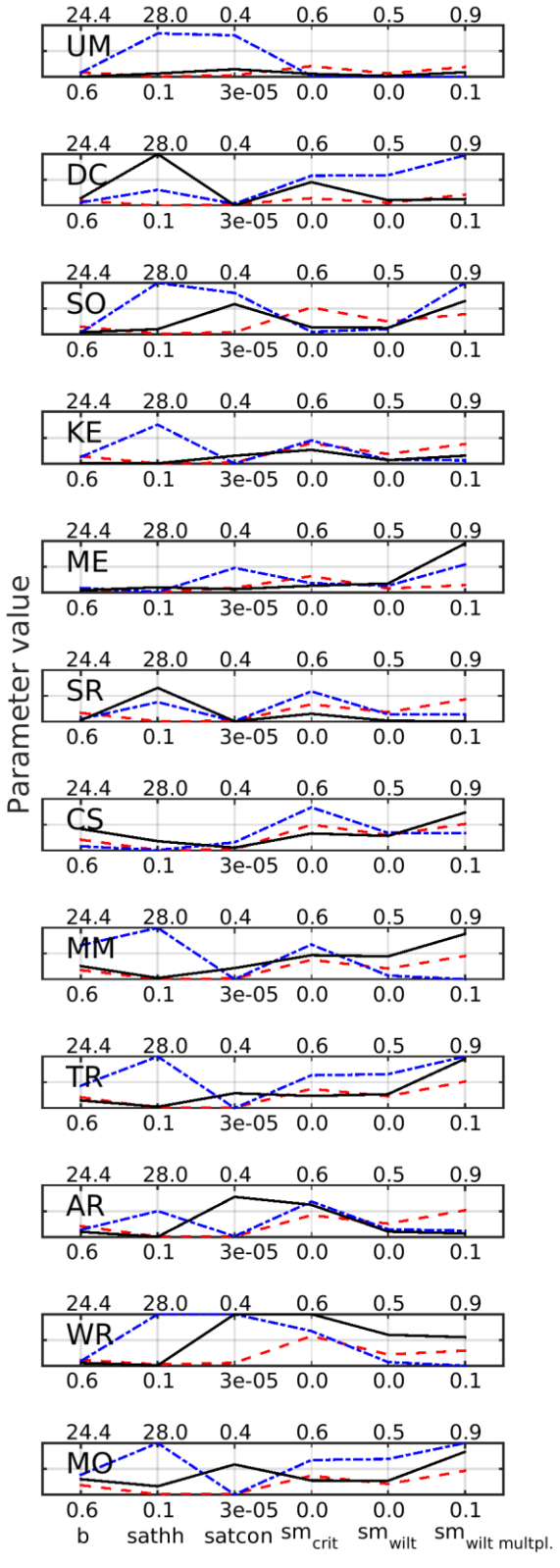
Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.



Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Figure A2.2: Morris sensitivity indices for the five calibrated parameters at four sites: UM, SR, WR, and MO. The three Objective Functions are the RMSEs between simulated and observed PS soil moisture (PS-SM), CRNS neutron counts (CRNS-N), and latent heat flux (LE). The mean elementary effects (EEs), representing the main effects, are displayed along the horizontal axis. The standard deviations of the elementary effects, representing parameter interactions, are shown along the vertical axis. Values were normalised to the most influential parameter for each Objective Function.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.



Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-558, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 8 November 2016
© Author(s) 2016. CC-BY 3.0 License.

Figure A2.3: Parameter values of the Default, PS-calibrated, and CRNS-calibrated solutions for each site shown in parallel coordinate plots. Parameter b is Mualem-Van Genuchten parameter 1/(n-1), sathh is Mualem-Van Genuchten parameter α^{-1}, satcon is the saturated soil hydraulic conductivity in the Mualem function, sm_{crit} is the critical point soil moisture content, sm_{wilt} is the wilting point soil moisture content, and is the multiplier (multiplied with sm_{crit}) which was actually calibrated instead of sm_{wilt}. The upper and lower bound values of the calibration ranges are shown for each parameter and were the same for all sites