Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

Using lagged dependence to identify (de)coupled surface and subsurface soil moisture values

Coleen D.U. Carranza[1], Martine J. van der Ploeg[1], and Paul J.J.F. Torfs[2]

[1]Soil Physics and Land Management Group, Wageningen University, the Netherlands
[2]Hydrology and Quantitative Water Management Group, Wageningen University, the Netherlands

Correspondence to: Coleen Carranza (coleen.carranza@wur.nl)

Abstract. Recent advances in radar remote sensing popularized mapping of surface soil moisture at different spatial scales. Surface soil moisture measurements are used in combination with hydrological models to determine subsurface soil moisture values. However, variability of soil moisture across the soil column is important for estimating depth-integrated values as decoupling between surface and subsurface can occur. In this study, we employed new methods to investigate the occurrence of (de)coupling between surface and subsurface soil moisture. Lagged dependence was incorporated in assessing (de)coupling with the idea that surface soil moisture conditions will be reflected at the subsurface after a certain delay. An exploratory step using residuals from a fitted loess function was performed as *a posteriori* information to determine (de)coupled values. The main approach was applying a distributed lag non-linear model (DLNM) to simultaneously represent both functional relation and lag structure. Both methods allow for a range of (de)coupled soil moisture values to be quantified. Results provide new insights on the decoupled range as its occurrence is not limited to dry conditions.

1 Introduction

Although recent decades have seen great advances in remote sensing applications for mapping surface soil moisture (Jackson, 1993; Njoku et al., 2003; Mohanty et al., 2017), most hydrological studies that make use of soil moisture data require integrated values over a certain soil depth (Brocca et al., 2017). Extrapolation of surface soil moisture from remote sensing techniques to depths beyond the sensor's capacity (up to 5cm) is not a trivial task given the spatio-temporal variability of soil moisture. The vertical distribution of soil moisture, which determines integrated soil moisture content over a soil column, is rarely uniform as more pronounced dynamics are expected closer to the surface compared to deeper in the soil (Hupet and Vanclooster, 2002). Currently, information derived from remote sensing are assimilated into hydrological models to obtain integrated soil moisture values (Houser et al., 1998; Das et al., 2008). However, Kumar et al. (2009) stressed that it is important to assess vertical variability, especially the strength of coupling between surface and subsurface soil moisture, for improvement of data assimilation results. While prevailing atmospheric conditions directly affect surface layers and control the temporal dynamics of soil moisture (Albertson and Montaldo, 2003; Koster et al., 2004), it is the downward movement of water from the surface that dictates the amount of subsurface soil moisture at a given time (Belmans et al., 1983; Rodriguez-Iturbe et al., 1999). Flow rates to the subsurface are driven by hydraulic properties, which are in turn controlled by physical soil characteristics such as texture, bulk density, and structure. Relative to changes in atmospheric conditions, soil physical properties change over longer timescales. Vegetation further modifies vertical soil moisture distribution by root water uptake (Yu et al., 2007) and by changing soil structure (Angers and Caron, 1998). Analyses of vertical soil moisture distributions also have important implications for modeling studies, as they could be used for calibration or validation of model parameters (De Lannoy et al., 2006). Given the variability along the soil column, during which conditions does surface soil moisture reflect subsurface soil moisture? Several studies have investigated this relation to address correspondence between surface and subsurface soil moisture content. One of the earliest studies is by Capehart and Carlson (1997) using modeling outputs for comparison with remote sensing measurements. Using very shallow depths of 5mm and 5cm, they observed deviations from linear correlation due to differences in drying rates which they referred to as "decoupling". Further assessment of decoupling from model-generated time series soil moisture data have been investigated using cross-correlation values (Martinez et al., 2008; Mahmood et al., 2012; Ford et al., 2014). High correlation to the subsurface was obtained using lagged values of surface soil moisture. However, cross-correlation is limited to providing a single value throughout the range of soil moisture encountered per lag. Furthermore, cross-correlation generally aims to evaluate the strength of lagged linear dependence between two variables (Shumway and Stoffer, 2010). However, lagged dependence between surface and subsurface soil moisture may not be linear given that non-linear processes determine water flow along the soil profile. Using in situ field measurements, Wilson et al. (2003) investigated spatial surface (0-6cm) and subsurface (0-30cm) soil moisture distribution by calculating statistical metrics and by means of a variogram. Decoupling between the two depths was observed which they suggested to be influenced by vegetation, especially root density at surface soil. Their results were also affected by the dry soil moisture range and emphasized the importance of distinguishing between surface and total soil moisture for future applications of remote sensing to atmospheric studies.

Based on previous studies, the term decoupling refers to a weak dependence between soil moisture contents at the surface and subsurface. Recognition of decoupling is important, however most studies have been limited to providing qualitative characterization of conditions when decoupling occurs (e.g. dry period). Only Capehart and Carlson (1997) identified a mid-range soil moisture ($\sim$0.3cm^{3}cm^{-3}) when the surface and very near surface begin to decouple. Their results, however, are limited to a thin layer of the soil column. In this paper, our main objective is to quantitatively identify a range of surface soil moisture values that is decoupled from the subsurface. Furthermore, we consider depths greater than those investigated by Capehart and Carlson (1997). The ability to quantify (de)coupled surface and subsurface soil moisture contents will contribute to more effective estimation of depth-integrated soil moisture data using remote sensing methods and improved data assimilation results in hydrological models.

We utilize in situ time series datasets at depths of 5cm and 40cm to represent surface and subsurface, respectively. Values outside the decoupled range are considered coupled since soil moisture is inherently bounded up a maximum value equal to soil's porosity. Investigation of (de)coupling is based on the idea that surface conditions will be reflected at the subsurface after a certain delay indicating strong coupling between the two zones, and vice versa. More focus is given to the decoupled soil moisture range since it has greater implications for extrapolation of surface soil moisture values to deeper soil layers. We applied statistical methods to identify conditions of decoupling with no prior assumptions on the type of functional relation between surface and subsurface. As an exploratory step, we first assessed dependence without considering lags using regres-

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

sion and residuals analysis. The main approach for assessing decoupling was application of distributed lag non-linear models (Gasparrini et al., 2010) to incorporate both lag structure and functional relation between surface and subsurface soil moisture. Applications of distributed lag models to econometrics and environmental epidemiology have been well documented (Almon, 1965; Zanobetti et al., 2002; Bhaskaran et al., 2013; Wu et al., 2013). However, their application to hydrological studies have rarely been explored.

2 Description of datasets and study sites

Four time series datasets from the Twente soil moisture and temperature monitoring network (Dente et al., 2011) were used in this study (fig.1). Datasets from 2014-2016 are available with only short periods of missing data. The stations are located in agricultural fields with sensors installed at 5cm, 10cm, 20cm, and 40cm depths. To investigate decoupling, only the 5cm and 40cm depths were considered because the largest possible distance was desired. Each station consists of EC-TM ECH2O capacitance probes (Decagon Devices, Inc., USA) that logged soil moisture data every 15 minutes. A calibration procedure using gravimetric measurements was applied prior to analysis (Dente et al., 2011).

Land cover in the area varies from corn in one field (SM05), to grass in two fields (SM05 and SM13), to a forest area (SM20). Values at 40cm capture the root zone of vegetation for each site. In reality, rooting depths vary and depend on species composition, climate, and plant growth rate. However, the depth considered would still allow for approximation of root zone conditions. The landscape is characterized by flat to slightly sloping terrain. It is important to note that SM20 is located at the eastern foot of a small hilly terrain. Throughout the study period, either land cover remained unchanged or the same crop was planted. The soil types for the stations range from coarse sandy soils to weakly silty soils (Wosten et al., 2013). A summary of the land cover and relevant characteristics of the stations are summarized in Table 1.

Soil moisture values were averaged into daily values to match the available daily rainfall data from the Dutch national weather service (KNMI). For SM13 and SM20, there are some missing data from the beginning of 2014. The datasets from SM13 begins on April 25, 2014 while SM20 begins on May 2, 2015 (fig.2).

3 Methods

3.1 Regression and residuals analysis

As an exploratory step, the dependence between surface and subsurface soil moisture was initially visualized using scatter-plots. Conditional means for every 0.01cm^{3}cm^{-3} interval and vertical bars representing ± standard deviation were added to show changes in vertical variability across the soil moisture range. Longer standard deviation bars indicate higher vertical variability. For the rest of this paper, variability will refer to vertical soil moisture variability, unless otherwise stated. Points were colored per month to show impacts of seasonality. The effect of rainfall was included by adjusting the sizes of the points proportional to rainfall intensity measured from the nearest KNMI stations. For the overall measure of dependence, Spear-

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

man's rank correlation coefficient R_s was computed for every pair of ranked values in the time series. This was chosen as the assumption of linear dependence is not made.

A flexible non-parametric locally weighted regression function (commonly called a loess function, Cleveland and Devlin (1988)) was fitted along the soil moisture range. This was used to explore and identify trends across the range. A linear regression was also fitted only for comparison. Residuals were analyzed further for variability not captured by the fitted function.

The cumulative residual variance, which more clearly shows the changes in variability across the range, was further analyzed. A significant change in slope between two neighboring points is identified along the cumulative residual variance line. This change occurs at an intermediate soil moisture value, marked by θ_c, that divides the soil moisture range into two groups. The group with the steeper slope is interpreted as the decoupled range, and vice versa. Since variance is sensitive to sample sizes, a correlation coefficient was calculated to determine if there was significant dependence between the two variables. Residuals variance were first normalized from 0-1 because of the varying soil moisture range encountered at each station.

Results of the exploratory methods are considered *a posteriori* knowledge for analysis of lagged dependence and interpretation of results.

3.2 Analysis of Lagged Dependence

3.2.1 Cross correlation

Since decoupling is based on the strength of lagged dependence, the existence of lag between surface and subsurface soil moisture values was first determined. Cross-correlation is a quick and easy method to apply for this objective. Lagged values of surface soil moisture are correlated with instantaneous values at the subsurface. Maximum cross-correlation at negative lags would indicate that surface soil moisture is leading subsurface soil moisture, and vice versa (Shumway and Stoffer, 2010). A 10-day lag was deemed long enough to show the presence of lag-lead relations in the time series since the maximum correlation occurs within this period.

3.2.2 Distributed lag model

The main approach for investigating decoupling is based on lagged dependence between surface and subsurface values. Decoupling is inferred when subsurface values show weak lagged dependence to surface soil moisture values. For the analysis, we only considered lag in vertical flow as lateral movement is deemed negligible in flat to slightly sloping terrain (Table.1). The strength of lagged dependence is determined using a distributed lag non-linear model (DLNM, (Gasparrini et al., 2010)). DLNM simultaneously represents both exposure-response dependencies and delayed effects in the time series. We considered surface soil moisture as the exposure which produces a delayed response in the subsurface. A non-linear model is used to capture the non-linear dynamics of flow and transport along the soil profile (Mohanty and Skaggs, 2001; Kim and Barros, 2002). In the following paragraphs, we provide a concise description of the DLNM concept. The full mathematical explanation of distributed lag non-linear models is described in Gasparrini et al. (2010) and (Gasparrini et al., 2017).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

Hydrology and Earth System Sciences
Discussions

From the general model of a time series, outcomes Y_t with $t = 1, \cdots, n$ are described by:

$$g(\mu_t) = \alpha + \sum_{j=1}^{J} s_j(x_{tj};\boldsymbol{\beta}_j) + \sum_{k=1}^{K} \gamma_k \boldsymbol{u}_t k \tag{1}$$

where $\mu \equiv E(Y)$, which is assumed to be derived from a Poisson distribution, and g is a monotonic link function. The functions s_j denote relationships between the variables x_j and vector parameters $\boldsymbol{\beta}_j$. Other u_k variables with predictors are included in coefficients γ_k specifying their related effects. The relation between x and $g(\mu)$ is represented by $s(x)$ through a basis function. The complexity of this estimated relationship depends on the type basis function chosen and its dimension. In the presence of delayed effects, the outcome Y at any time t is explained by the past exposures x_{t-l} with l as the *lag* representing the elapsed time between exposure and response. The final goal of DLNM is to simultaneously describe the dependency along both the predictor space and lag dimension. This is achieved by selecting two sets of basis functions that are combined to obtain the cross-basis functions (Gasparrini et al., 2010).

Within the DLNM framework, a response Y_t at time $t = 1$ is based on lagged occurrences of predictor x_t, which is represented by vector $q_t = [x_{t-l_0}; \cdots; x_{t-L}]^T$. The minimum and maximum lags are given by l_0 and L_T, respectively. The function represents dependence through:

$$s(q,t) = s(x_{t,t-l_0}, \cdots, x_{t-L}) = \sum_{l=l_0}^{L} f \cdot w(x_{t-L}, l) \tag{2}$$

where $f \cdot w(x_{t-L}, l)$ represents the exposure-lag-response function which is composed of two marginal functions: the exposure-response function $f(x)$ and lag-response function $w(l)$ in the space of the lag. Parameterization of f and w is achieved by application of known basis functions to vectors q_t and l. The result can be expressed as matrices **R** and **C** with dimensions $(L - l_0 + 1) \times v_x$ and $(L - l_0 + 1) \times v_l$, respectively.

The cross basis function s and parameterized coefficients $\boldsymbol{\eta}$ are given by:

$$s(x_{t,t-l_0}, \cdots, x_{t-L}; \boldsymbol{\eta}) = (1^T_{L-l_{0+1}} \boldsymbol{A}_t)\boldsymbol{\eta} = \boldsymbol{w}_t^T \boldsymbol{\eta} \tag{3}$$

The values of $\boldsymbol{w}$ are derived from $\boldsymbol{A_t}$, which is computed from the row-wise Kronecker product between matrices **R** and **C**. The dependence is expressed through $\boldsymbol{w}$ and parameters $\boldsymbol{\eta}$. The cross-basis function represents the integral of $s(x,t)$ over the interval $[l_0, L]$, summing the contributions from the exposure history. Estimated dependence to specific exposure values is determined by prediction of $\hat{\beta}$, called lag coefficients. The estimated $\hat{\beta}$ and covariance matrix $V(\hat{\beta})$ is given by:

$$\hat{\beta} = \boldsymbol{A}_x \hat{\boldsymbol{\eta}} \tag{4}$$

$$V(\hat{\beta}) = \boldsymbol{A}_x V(\hat{\boldsymbol{\eta}}) \boldsymbol{A}_x^T \tag{5}$$

The values of $\hat{\beta}$ indicate the strength of dependence between surface and subsurface soil moisture. Higher $\hat{\beta}$ values indicate stronger dependence or coupling between the two. Hence, we refer to $\hat{\beta}$ as the relative influence of surface soil moisture on subsurface values. A further extension to DLNM is the application of penalties for smoothness of the lag structure and shrinkage of lag coefficients to null at very high lags. These penalties are applied using a second-order difference (Wood, 2006b) and varying ridge penalties (Obermeier et al., 2015; Gasparrini et al., 2017), respectively. Application of penalties is based on the assumption that, at higher lags, the lag coefficients become smaller and approach the null value.

In order to identify a range that is decoupled, a threshold value ($\hat{\beta}_c$) must be specified. This value is comparable to the intermediate soil moisture θ_c identified from Section 4.1. The values of θ_c provided a suitable guide for identifying a threshold common to all four sites. The corresponding $\hat{\beta}$ values obtained at θ_c were very close to 1, therefore, setting the threshold $\hat{\beta}_c = 1$ seemed a reasonable choice. This was preferred over the exact $\hat{\beta}$ at each θ_c since the latter was defined using exploratory methods at $lag = 0$. Using the chosen $\hat{\beta}_c = 1$, surface soil moisture values with $\hat{\beta} < 1$ are considered decoupled while those with $\hat{\beta} \geq 1$ are coupled.

In assessing lagged dependence, event scale analysis of decoupling is of interest rather than large scale patterns within the time series (Wilson et al., 2004). This requires seasonal patterns to be addressed, which was done by fitting a loess function and then subtracting it from the time series data (Cleveland et al., 1990). Removal of seasonality was further justified by scatterplot results (see Section 4.1). The influence of seasonality on vertical soil moisture variability is indicated by clustering of observation points occurring within the same months (fig.3). De-seasonalized soil moisture values were used for determining decoupling from the time series datasets.

For consistency in modeling, the range of surface soil moisture values used was from 0-0.50cm^{3}cm^{-3}. This is based on the highest value encountered for surface soil moisture among the four sites. A lag value of up to 30 days was considered long enough to investigate delayed effects. This period also approximates the recurrence of heavy rainfall within the study sites. Spline function was the basis function chosen for both exposure-response and lag-response functions as it offers flexibility to capture non-linearities. In addition, contributions from daily rainfall data were used to incorporate current and past meteorological conditions. This was used as a covariate that was represented with an additional basis function. It is a covariate that is represented with an additional basis function. The analysis was performed in R software using *dlnm* (Gasparrini, 2011) and *mgcv* (Wood, 2006a) packages.

4 Results

4.1 Regression and Residuals analysis

The overall dependence between surface and subsurface given by Spearman's rank coefficient (R_s) range from 0.746 to 0.866 (fig.3). However, even with a high overall dependence, variability is not uniform across the soil moisture range (fig.3). Except for SM13, increased variability is observed towards drier soil moisture values. Furthermore, the degree of variability also differs among the four sites. Most pronounced variability is observed at SM13 and the least at SM05. Clustering of observation points occurring within the same months indicate that seasonality dictates soil moisture values and impacts soil moisture variability.

Rainfall events measured on the same day do not show a clear effect on surface and subsurface soil moisture dependence. Observations with higher rainfall intensities appear scattered. In addition, the said observation points do not necessarily fall along the fitted functions or at the wet soil moisture region of the scatterplots. As lag is not considered, the impact of rainfall on variability is not fully captured from scatterplots alone.

Assessment of the regression fit quality was performed by comparison using residual standard errors (RSE). The results for both linear and loess functions show highly similar values (fig.3). This indicates that, in this case, a linear function captures the relation between surface and subsurface values. Nevertheless, the more flexible loess function was preferred for further residuals analysis because of its slightly better model fit and, using only visual inspection of fig.3, it more closely approximates the calculated conditional mean.

Figure 4 shows the residual plots with lines of the cumulative residual variance. The change in slope of the line is a feature consistent for all sites regardless of the magnitude of residual variance. Furthermore, changes in variability are more clearly observed from the residuals than from the standard deviation bars in the scatterplots. The location of the change in slope, θ_c, is highlighted by the vertical dashed line. Decoupled soil moisture corresponds to the section of cumulative variance line with a steeper slope.

Specifically, the range of decoupled surface soil moisture values (in cm^3cm^{-3}) is 0.08-0.21 for SM05, 0.12-0.27 for SM09, 0.30-0.39 for SM13, and 0.08-0.12 for SM20. Except for SM13, the decoupled values are within the dry to intermediate soil moisture range. The cumulative variance line for SM13 appears to increase exponentially with increasing surface soil moisture. This differs from the other three sites which show a distinct decrease in slope at increasing soil moisture. For SM20, a second point is identified with a change in slope. The flat line starting from 0.24-0.28cm^3cm^{-3} indicates there is still lowered variance at the very wet soil moisture range.

The correlation between normalized variance and sample size yielded a value of -0.24 (fig.5). This low correlation magnitude confirms that the variance obtained for the soil surface moisture intervals was not strongly influenced by the sample size used.

4.2 Cross-correlation

Figure 6 shows cross-correlation values at the four sites. Maximum correlation occurs at -1 to -2 days lag, except at SM20. This translates to a 1-2 day lead of surface soil moisture values. For SM20, the maximum correlation occurs at positive lags. Correlation values from lag=0 to lag=10 are almost equal at SM20. Although this indicates leading subsurface values, it does not eliminate the possibility of having a lag between surface and subsurface values (see Section 5.2). Other factors may play a role in having leading subsurface values in the cross-correlation plots. Hence, SM20 was still analyzed for decoupling using DLNM.

4.3 Distributed lag model

Figure 7 shows the overall $\hat{\beta}$ for each surface soil moisture value with 5% and 95% confidence intervals in shaded gray regions. Based on $\hat{\beta}_c$, the identified decoupled values are generally in the dry to intermediate soil moisture range (fig.7), except for SM13 where decouple values are at the wet range. Table 2 shows the decoupled values identified based on the selected $\hat{\beta}_c$.

Behavior and trends of $\hat{\beta}$ also differ for each station. For instance, at SM05 and SM09, there is a general increase in $\hat{\beta}$ from dry towards wet surface soil moisture values. However, the predicted $\hat{\beta}$ in the very dry soil moisture range (<0.10cm^{3}cm^{-3}) are different for the two sites as decoupling is predicted for only SM05. At SM13, $\hat{\beta}$ values mostly fluctuate around $\hat{\beta}_c$ and then proceed to decrease toward the decoupled wet range. The behavior of $\hat{\beta}$ for SM20 has an increasing trend over a limited range. Among the four sites, SM20 has the smallest soil moisture range, only reaching a maximum of 0.28cm^{3}cm^{-3}. The limited soil moisture range in SM20 leads to very high uncertainties in the predicted $\hat{\beta}$ for wet soil moisture conditions (>0.30cm^3/cm^3). Recall that the range used for DLNM was only for uniformity among the four study sites. The lack of or very few observations for very dry or very wet soil moisture conditions led to wider confidence intervals not only for SM20 but also for the other three sites.

5 Discussion

5.1 Decoupled soil moisture values

Regression and residuals analysis show that there is an inherent vertical variability between surface and subsurface soil moisture values based on the lack of 1:1 correspondence between the two (fig.3). This inherent variability is also not uniform as higher variability is observed in certain soil moisture ranges. The cumulative residual variance plots (fig.4) clearly indicate the soil moisture values where vertical variability starts to become consistently larger. The increase in variability further translates to weak lagged dependence which we observe as low $\hat{\beta}$ values from DLNM. The increase in vertical variability and weakening of lagged dependence is what we consider to be decoupling between the surface and subsurface.

Both residuals analysis and DLNM were successful in identifying a decoupled soil moisture range and there is good agreement between the results from both. Three out of four sites show decoupled values in the dry to intermediate soil moisture range (fig.4 and Table 2). These results agree with the known range where decoupling is expected (Capehart and Carlson, 1997; Hirschi et al., 2014; Wilson et al., 2003). For SM05 and SM09, the intermediate soil moisture value,θ_c that marks when decoupling begins (Table 2) is close to that identified by Capehart and Carlson (1997). They obtained a value of 0.3cm^{3}cm^{-3} as the point below which decoupling begins. However, results for SM13 do not conform to the traditional concept of decoupling. This result is significant as it implies that decoupling can occur at any value and is not confined to dry soil moisture range.

The physical characteristics at SM13 allow some insights into the potential controls for decoupling. For instance, further inspection of the soil profile showed a slight increase in silt content at 40cm compared to 5cm. The sensor is also located close to a small shallow ditch (Table 1). These factors contribute to increased water retention and soil water content at the 40cm depth which may affect decoupling at the site. The identified decoupled values occur during colder winter periods (fig.3). One factor that could potentially induce changes to soil physical characteristics during winter periods is the presence of burrowing animals that are in hibernation. Site inspection confirmed the presence of burrows at SM13. These burrows create macropores which eventually alter hydraulic properties of the soil (Kodešová et al., 2006; Beven and Germann, 2013).

Site-specific characteristics at each station control the magnitude of variability as well as the range where decoupling is observed. However, the occurrence of decoupling is independent of the magnitude of variability since it was observed from

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

SM05 where variability is least up to SM13 where it is greatest. The methods applied in this study only identify conditions when decoupling occurs but do not explicitly determine its controls. Identification of controls for decoupling requires a separate analysis where mechanistic models or statistical approaches can be applied.

5.2 Assessing the use of lagged dependence for identification of decoupling

To assess the applicability of the methods applied, we further discuss their strengths and weaknesses. We also present opportunities for further studies as well as foreseen limitations for other sites.

Strengths: he residuals analysis and DLNM methods allow quantification of a range of soil moisture values where decoupling occurs. This provides further extension to previous studies where decoupling is only described qualitatively. As seen from the results at the four sites, decoupling can occur at any soil moisture value, and is not confined to dry periods or ranges. Furthermore, by making no initial assumptions on data distributions and the type of functional relation and lag structure, the methods applied were considered robust. Non-linear functions were applied as they conform to the nonlinearity of water flow in the unsaturated zone. They can also handle a variety of bivariate dependence, even in cases where the relation is linear, as shown by the highly similar fit of the loess and linear functions in Section.4.1.

Weaknesses: The first aspect that needs to be further investigated is the selected $\hat{\beta}$ value used for identifying the decoupled soil moisture range. Although the selection in this study was based on trends identified from time series datasets, the methods applied should be tested further using other datasets to confirm the suitability of $\hat{\beta}_c = 1$ for other depths and soil types. Another aspect is the use of cross-correlation for confirming the presence of leading surface soil moisture values. Results from SM20 show maximum correlation at positive lags which indicate leading subsurface values (fig.6). The weakness of using cross-correlation as a test for the presence of lag can be two-fold. First, cross-correlation can also capture the effect of subsurface dynamics such as groundwater influence and lateral flow. We infer that in SM20, subsurface dynamics dominates and masks the lag relation sought. An additional covariate representing subsurface dynamics was not included in the DLNM analysis since a dominant downward vertical flow was assumed. This assumption was based on the flat slopes encountered at SM20 (Table 1). Therefore, the occurrence of subsurface lateral flow or groundwater influence pose limitations to the applicability of DLNM for assessing decoupling. Second, cross-correlation is limited to evaluating linear lagged dependence. Incorporating non-linear lagged dependence can make the test more robust. Equivalent methods exist (e.g. mutual information content (Qiu et al., 2014)) but they are much more computationally demanding when the goal is simply to check for the existence of lag-lead relation.

Opportunities: In relation to utilizing remote sensing techniques, our results imply that the accuracy of estimating subsurface values from surface soil moisture can be greatly affected by vertical coupling. Lower variability and hence lower uncertainties are expected in the coupled soil moisture range. Assessment of decoupling can be used in combination with modeling studies as a preliminary method to determine the range where variability is expected to be higher. Furthermore, it can be helpful in assessing whether simulation results capture the variabilities observed in both the coupled and decoupled ranges. Taking decoupling into account can also assist in evaluating the necessity of complex models for simulating vertical soil moisture content.

Limitations: In this study, only meteorological factors were incorporated in the DLNM analysis since vertical movement was assumed to be the dominant flow mechanism. However, the subsurface can also be influenced by lateral movement or groundwater by capillary rise. In such scenarios, decoupling will not be limited to changes in surface conditions. For this, SM20 provides an excellent example. This station is located at the foot of a small hill (fig.2) where the occurrence of lateral subsurface movement is highly probable. This shows that although the analysis would be limited to smaller scales, or even a single point, recognition of regional setting is important for interpretation of results. In addition, subsurface dynamics can also be affected by capillary rise in areas with shallow groundwater. For future applications, the effect of both capillary rise and lateral movements to subsurface dynamics should be assessed and included in the DLNM analysis but caution should be exercised when interpreting results. Assessment of decoupling with DLNM is deemed more applicable to areas where the subsurface has insignificant groundwater influence and where vertical downward movement is the dominant flow mechanism.

6 Conclusion

The methods applied in this study allow for investigation of vertical soil moisture variability. More importantly, application of DLNM allowed for decoupled soil moisture range to be quantitatively identified. The results also reveal that decoupling is not confined to dry soil moisture range as implied by previous studies. The reasons for decoupling are manifold and controls for the dry soil moisture range may differ from those for the wet range. The results of this study have implications for remote sensing and data assimilation methods, especially for uncertainties related to the use of surface soil moisture to obtain integrated soil moisture values.

Data availability. The datasets for soil moisture were obtained from the Water Resource Department of ITC-Twente University. At the moment, the datasets are not publicly available. Access to the datasets may be granted upon request from the institute thru Prof. Rogier van der Velde, PhD (r.vandervelde@utwente.nl).

Author contributions. Coleen Carranza and Martine van der Ploeg initially conceptualized the idea for investigating the relation between surface and subsurface soil moisture values. Paul Torfs provided significant contributions to the statistical analysis applied. All three authors contributed to writing and editing of the manuscript.

Competing interests. No competing interest present

Acknowledgements. The authors are grateful for the Water Resources Department of ITC-Twente University, the Netherlands for sharing the datasets from their network. This work is part of the research programme Optimizing Water Availability through Sentinel-1 Satellites

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

Hydrology and Earth System Sciences
Open Access
Discussions

(OWAS1S) with project number 13871 which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO). We also thank Demie Moore for reviewing the manuscript.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

References

Albertson, J. D. and Montaldo, N.: Temporal dynamics of soil moisture variability: 1. Theoretical basis, Water Resources Research, 39, 2003.

Almon, S.: The distributed lag between capital appropriations and expenditures, Econometrica: Journal of the Econometric Society, pp. 178–196, 1965.

Angers, D. A. and Caron, J.: Plant-induced changes in soil structure: processes and feedbacks, in: Plant-induced soil changes: processes and feedbacks, pp. 55–72, Springer, 1998.

Belmans, C., Wesseling, J., and Feddes, R. A.: Simulation model of the water balance of a cropped soil: SWATRE, Journal of hydrology, 63, 271–286, 1983.

Beven, K. and Germann, P.: Macropores and water flow in soils revisited, Water Resources Research, 49, 3071–3092, 2013.

Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L., and Armstrong, B.: Time series regression studies in environmental epidemiology, International journal of epidemiology, 42, 1187–1195, 2013.

Brocca, L., Ciabatta, L., Massari, C., Camici, S., and Tarpanelli, A.: Soil Moisture for Hydrological Applications: Open Questions and New Opportunities, Water, 9, 140, 2017.

Capehart, W. J. and Carlson, T. N.: Decoupling of surface and near-surface soil water content: A remote sensing perspective, Water Resources Research, 33, 1383–1395, 1997.

Cleveland, R. B., Cleveland, W. S., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, Journal of Official Statistics, 6, 3, 1990.

Cleveland, W. S. and Devlin, S. J.: Locally weighted regression: an approach to regression analysis by local fitting, Journal of the American statistical association, 83, 596–610, 1988.

Das, N., Mohanty, B., Cosh, M., and Jackson, T.: Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04, Remote Sensing of Environment, 112, 415–429, 2008.

De Lannoy, G. J., Verhoest, N. E., Houser, P. R., Gish, T. J., and Van Meirvenne, M.: Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE 3), Journal of Hydrology, 331, 719–730, 2006.

Dente, L., Verkedy, Z., Su, Z., and Ucer, M.: Twente Soil Moisture and Soil Temperature Monitoring Network, Tech. rep., University of Twente—ITC: Enschede, The Netherlands, http://www.itc.nl/library/papers_2011/scie/dente_twe.pdf, 2011.

Ford, T., Harris, E., and Quiring, S.: Estimating root zone soil moisture using near-surface observations from SMOS, Hydrology and Earth System Sciences, 18, 139–154, 2014.

Gasparrini, A.: Distributed lag linear and non-linear models in R: the package dlnm, Journal of statistical software, 43, 1, 2011.

Gasparrini, A., Armstrong, B., and Kenward, M.: Distributed lag non-linear models, Statistics in medicine, 29, 2224–2234, 2010.

Gasparrini, A., Scheipl, F., Armstrong, B., and Kenward, M.: A penalized framework for distributed lag non-linear models, Biometrics, 2017.

Hirschi, M., Mueller, B., Dorigo, W., and Seneviratne, S.: Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability, Remote sensing of environment, 154, 246–252, 2014.

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., and Goodrich, D. C.: Integration of soil moisture remote sensing and hydrologic modeling using data assimilation, Water Resources Research, 34, 3405–3420, 1998.

Hupet, F. and Vanclooster, M.: Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field, Journal of Hydrology, 261, 86–101, 2002.

Jackson, T. J.: III. Measuring surface soil moisture using passive microwave remote sensing, Hydrological processes, 7, 139–152, 1993.

Kim, G. and Barros, A. P.: Space–time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data, Remote sensing of environment, 81, 393–403, 2002.

Kodešová, R., Kodeš, V., Žigová, A., and Šimnek, J.: Impact of plant roots and soil organisms on soil micromorphology and hydraulic properties, Biologia, 61, S339–S343, 2006.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., Lawrence, D., et al.: Regions of strong coupling between soil moisture and precipitation, Science, 305, 1138–1140, 2004.

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., and Peters-Lidard, C. D.: Role of subsurface physics in the assimilation of surface soil moisture observations, Journal of hydrometeorology, 10, 1534–1547, 2009.

Mahmood, R., Littell, A., Hubbard, K. G., and You, J.: Observed data-based assessment of relationships among soil moisture at various depths, precipitation, and temperature, Applied Geography, 34, 255–264, 2012.

Martinez, C., Hancock, G., Kalma, J., and Wells, T.: Spatio-temporal distribution of near-surface and root zone soil moisture at the catchment scale, Hydrological Processes, 22, 2699–2714, 2008.

Mohanty, B. and Skaggs, T.: Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation, Advances in water resources, 24, 1051–1067, 2001.

Mohanty, B. P., Cosh, M. H., Lakshmi, V., and Montzka, C.: Soil Moisture Remote Sensing: State-of-the-Science, 16, https://doi.org/10.2136/vzj2016.10.0105, 2017.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, IEEE transactions on Geoscience and remote sensing, 41, 215–229, 2003.

Obermeier, V., Scheipl, F., Heumann, C., Wassermann, J., and Küchenhoff, H.: Flexible distributed lags for modelling earthquake data, Journal of the Royal Statistical Society: Series C (Applied Statistics), 64, 395–412, 2015.

Qiu, J., Crow, W. T., Nearing, G. S., Mo, X., and Liu, S.: The impact of vertical measurement depth on the information content of soil moisture times series data, Geophysical Research Letters, 41, 4997–5004, 2014.

Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., and Cox, D.: Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation, Proceedings of the Royal Society of London Series A, 455, 3789, 1999.

Shumway, R. H. and Stoffer, D. S.: Time series analysis and its applications: with R examples, Springer Science & Business Media, 2010.

Wilson, D. J., Western, A. W., Grayson, R. B., Berg, A. A., Lear, M. S., Rodell, M., Famiglietti, J. S., Woods, R. A., and McMahon, T. A.: Spatial distribution of soil moisture over 6 and 30cm depth, Mahurangi river catchment, New Zealand, Journal of Hydrology, 276, 254–274, 2003.

Wilson, D. J., Western, A. W., and Grayson, R. B.: Identifying and quantifying sources of variability in temporal and spatial soil moisture observations, Water Resources Research, 40, 2004.

Wood, S.: Generalized Additive Models: An Introduction with R, CRC Press, 2006a.

Wood, S. N.: Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models, Biometrics, 62, 1025–1036, https://doi.org/10.1111/j.1541-0420.2006.00574.x, 2006b.

Wosten, J., de Vries, F., Hoogland, T., Massop, H., Veldhuizen, A., Vroon, H., Wesseling, J., Heijkers, J., and Bolman, A.: BOFEK2012, de nieuwe bodemfysische schematisatie van Nederland, Tech. rep., Alterra, 2013.

Wu, W., Xiao, Y., Li, G., Zeng, W., Lin, H., Rutherford, S., Xu, Y., Luo, Y., Xu, X., Chu, C., et al.: Temperature–mortality relationship in four subtropical Chinese cities: a time-series study using a distributed lag non-linear model, Science of the Total Environment, 449, 355–362, 2013.

Yu, G.-R., Zhuang, J., Nakayama, K., and Jin, Y.: Root water uptake and profile soil water as affected by vertical root distribution, Plant Ecology, 189, 15–30, https://doi.org/10.1007/s11258-006-9163-y, https://doi.org/10.1007/s11258-006-9163-y, 2007.

Zanobetti, A., Schwartz, J., Samoli, E., Gryparis, A., Touloumi, G., Atkinson, R., Le Tertre, A., Bobros, J., Celko, M., Goren, A., et al.: The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement, Epidemiology, 13, 87–93, 2002.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

Hydrology and Earth System Sciences
Open Access
Discussions

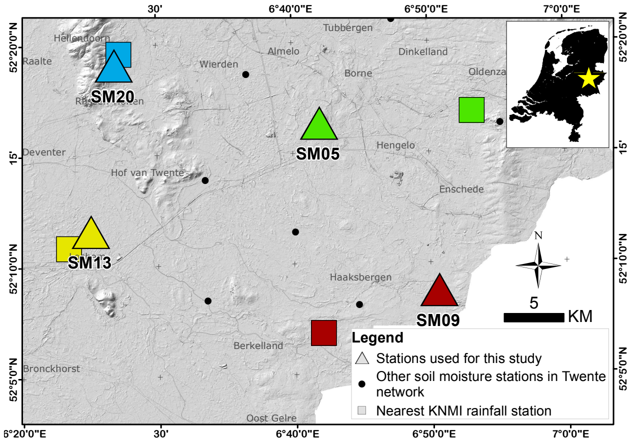

Figure 1. Location of study site in the eastern part of the Netherlands (inset). Triangles represent stations in the Twente soil moisture and temperature monitoring network (Dente et al., 2011). Squares represent meteorological stations. Symbols with similar colors indicate the pair of measurements used for the analysis..

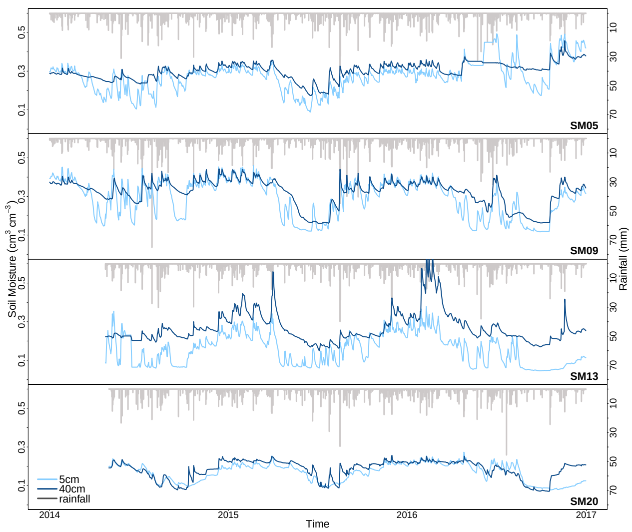

Figure 2. Time series plots of surface (5cm in light blue) and subsurface (40cm in dark blue) soil moisture. Vertical black bars at the top show daily precipitation data from the nearest KNMI station.

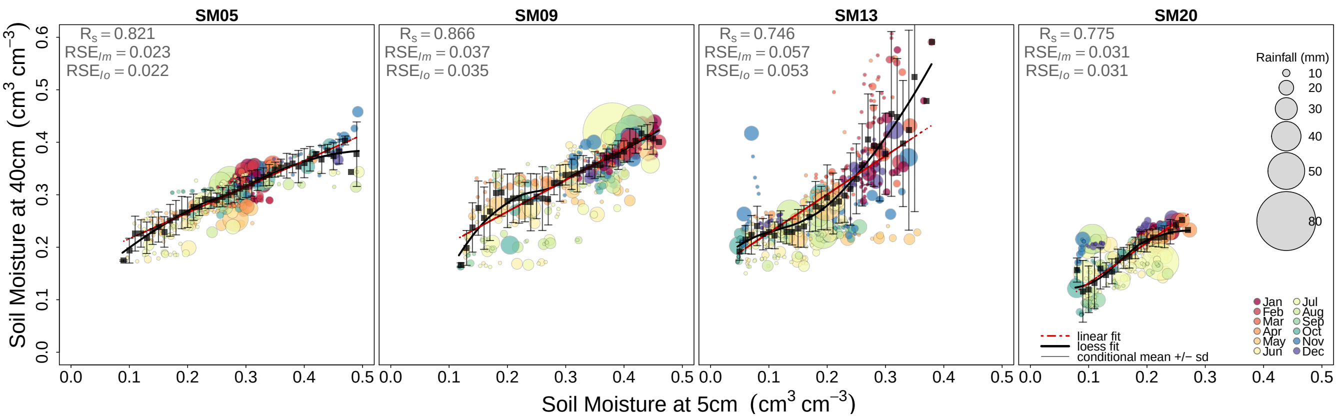

Figure 3. Scatter plots of 5cm vs 40cm soil moisture values at lag=0. Colors correspond to the months in a year and sizes of points are proportional to rainfall intensity. Trends along the soil moisture range shown with the fitted loess function (black line). Vertical variability across the soil moisture range is indicated by the lengths of standard deviation bars. A linear function (red line) is also fitted for comparison. The overall dependence using Spearman's rank correlation R_s is given in the upper left corner each plot. Residual standard errors (RSE) for loess (lo) and linear (lm) fits are also shown.

Table 1. Summary of land cover descriptions at each station covering the period of 2014-2016. Soil descriptions and codes are based on BOFEK 2012 (Wosten et al., 2013). Slope and distance to nearest ditch were determined from 5m resolution DEM. Datasets from 2016 were obtained from the publicly available national topographic database of the Netherlands (TOPNL)

Station No.	Land cover	BOFEK Soil description	Slope (degrees)	Aerial distance to nearest ditch (m)
SM05	Grass	Loamy sandy soils with a thick cultivated layer(317)	2.22	18.97
SM09	Corn	Weakly loamy sand soils with a thick cultivated layer(311)	2.70	1.41
SM13	Grass	Weak silty soils (podsols)(304)	1.0	17.09
SM20	Forest	Coarse sand (podsols)(320)	2.30	875.26

Table 2. List of surface soil moisture values (SSM in cm^{3}cm^{-3}) obtained from fig.4 and fig.7. SSM at θ_c in fig.4 were used to determine $\hat{\beta}$ in fig.7. A common threshold $\hat{\beta}_c$ was used for all sites since $\hat{\beta}$ are all close to 1. The resulting decoupled SSM values are shown in the fourth column.

Station No.	SSM at θ_c	$\hat{\beta}$ at θ_c	Decoupled values using threshold $\hat{\beta}_c = 1$
SM05	0.21	0.90	<0.24
SM09	0.27	0.97	<0.28
SM13	0.30	1.18	>0.34
SM20	0.12	0.94	0.16>SSM>0.23

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

Hydrology and Earth System Sciences
Open Access
Discussions

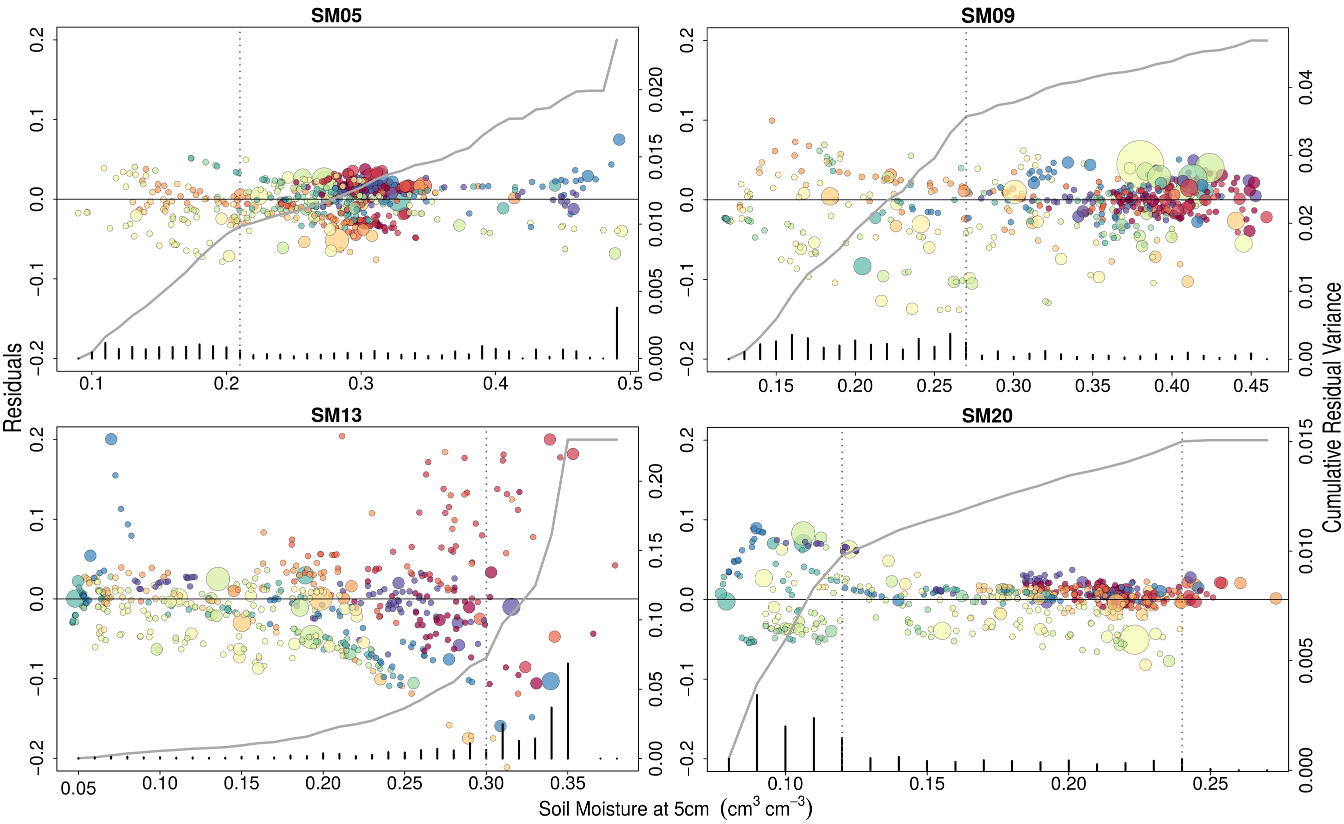

Figure 4. Residual variance plots from the fitted loess function. Vertical bars at the bottom of each plot represent the variance for every 0.1cm^{3}cm^{-3} interval. Cumulative residual variance (gray line) shows a change in slope at an intermediate soil moisture value (referred to as θ_c, indicated by the vertical black dashed line). The change in slope separates a range with higher variance (steeper slope) and lower variance (gentler slopes). The range with higher variance is considered decoupled, and vice versa.

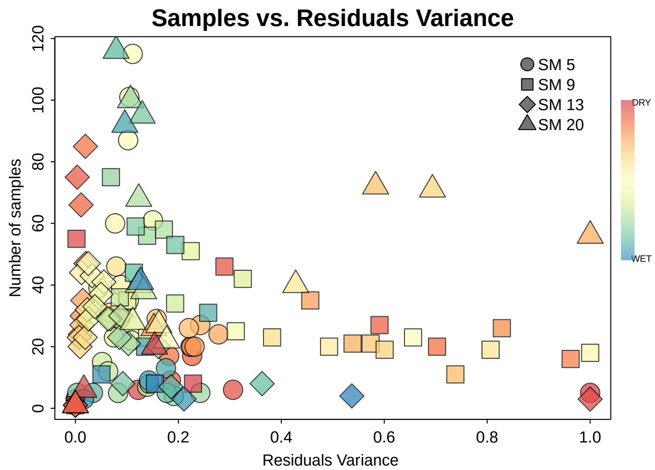

Figure 5. Scatterplot of sample size vs. normalized residual variance calculated for each $0.01cm^3cm^{-3}$ interval. Colors indicate soil moisture conditions at each point. The plot of points indicate very weak linear dependence quantified by a -0.24 correlation coefficient.

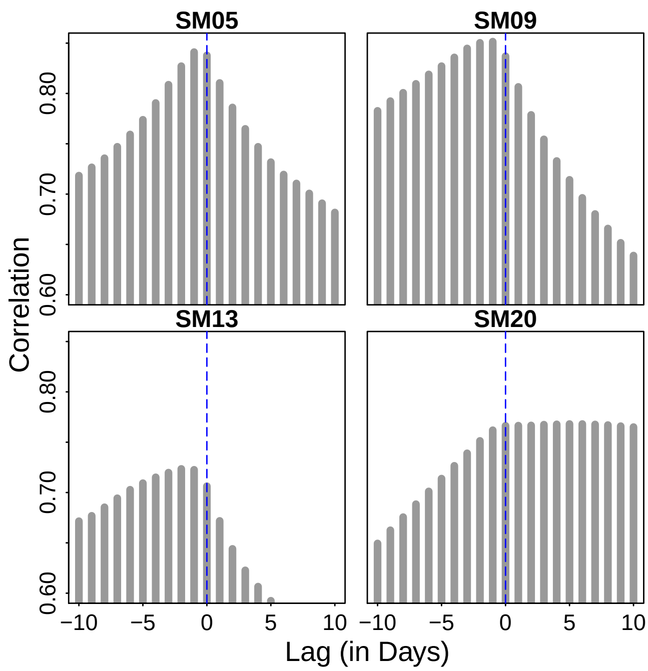

Figure 6. Cross-correlation of 5cm vs 40cm soil moisture values. Lagged 5cm values are correlated with 40cm values. A 1-2 day lead of surface soil moisture is observed except for SM20. This is indicated by having maximum values at lag -1 to -2 days. At SM20, maximum correlation occurs at positive lags.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-651
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 13 November 2017

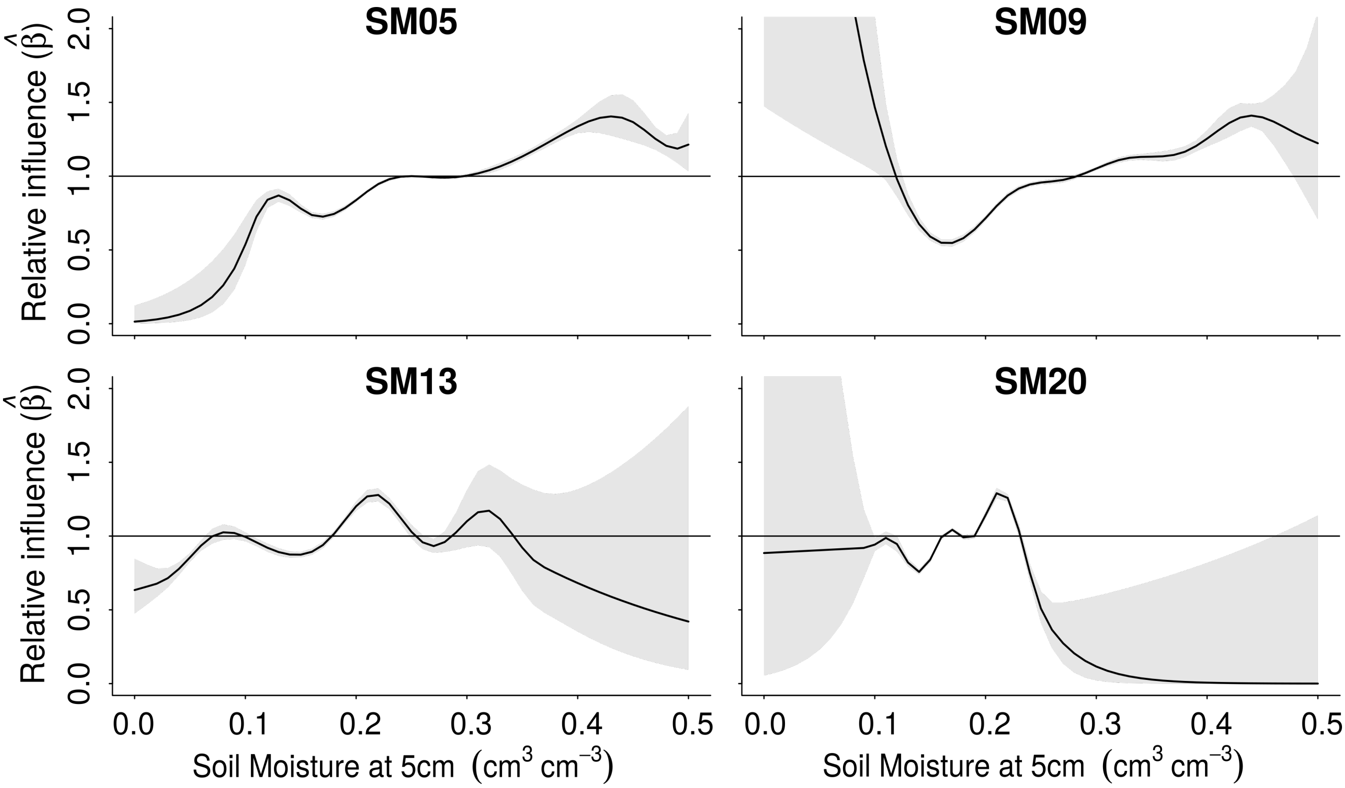

Figure 7. The relative influence of surface soil moisture on subsurface values obtained by summing the predicted $\hat{\beta}$ along the 30-day lag. The threshold value ($\hat{\beta}_c$) used to identify the decoupled range is indicated by the horizontal line. Surface soil moisture values below $\hat{\beta}_c$ are considered decoupled. The 5% and 95% confidence intervals of the predicted values are shown as shaded regions.