Soil moisture: variable in space but redundant in time

Mirko Mälicke[1], Sibylle K. Hassler[1], Theresa Blume[2], Markus Weiler[3], and Erwin Zehe[1]

[1]Institute for Water and River Basin Management, Karlsruhe Institute of Technology (KIT), Germany
[2]GFZ German Research Centre for Geosciences, Section Hydrology, Germany
[3]Hydrology, Faculty of Environment and Natural Resources, University of Freiburg, Germany

Correspondence: Mirko Mälicke (mirko.maelicke@kit.edu)

Abstract.

"Soil moisture at the headwater scale exhibits a huge spatial variability and single or even distributed TDR measurements yield non-representative data" (Zehe et al., 2010, p.874 l.11-14). This suggests that even a huge amount of observation points would not be able to capture soil moisture variability. Here we ask whether spatial variability is the dead end to spatially
5 distributed point sampling - or whether point networks yield representative data on dynamic changes nevertheless?

We present a measure to capture the spatial dissimilarity, or dispersion, and its change over time. Statistical dispersion among observation points is related to their distance to describe spatial patterns. We analyzed the temporal evolution and emergence of these patterns and use Mean shift clustering algorithm to identify and analyze clusters. We found that soil moisture observations from the Colpach catchment in Luxembourg to cluster in two fundamentally different states. On the one
10 hand, we found rainfall-driven data clusters, usually characterized by strong relationships between dispersion and distance. Their spatial extend roughly matches the hillslope scale. On the other hand, we found clusters covering the vegetation period. In drying and then dry soil conditions there is no particular spatial dependence in soil moisture patterns, but the values are highly similar beyond hillslope scale.

By combining uncertainty propagation with information theory, we were able to calculate the information content of spatial
15 similarity with respect to measurement uncertainty (when are patterns different outside of uncertainty margins?). We were able to prove that the spatial information contained in soil moisture observations is highly redundant and can be compressed to only a fragment of the original data volume without significant information loss.

Our most interesting finding is that even a few soil moisture time series bear a considerable amount of information about dynamic changes of soil moisture. We argue that distributed soil moisture sampling reflects an organized catchment state,
20 where soil moisture variability is not random. Thus, only a small amount of observation points is necessary to capture soil moisture dynamics.

1 Introduction

Although soil water is by far the smallest fresh water stock on earth, it plays a key role in the functioning of terrestrial ecosystems. Soil moisture controls (preferential) infiltration and runoff generation and is a limiting factor for vegetation growth. Plant-available soil water affects the Bowen ratio i.e. the partitioning of net radiation energy in latent and sensible heat, and last

5 but not least it is an important control for soil respiration and related trace gas emissions. Technologies and experimental strategies to observe soil water dynamics across scales have been at the core of hydrological research agenda for more than 20 years (Topp et al., 1982, 1984). Since the early days of Topp spatially and temporally distributed Time Domain Reflectometry (TDR) and Frequency Domain Reflectometry (FDR) measurements have been widely used to characterize soil moisture dynamics at the transect (eg. Blume et al., 2009) , hillslope (eg. Starr et al., 2002; Brocca et al., 2007) and catchment scales (eg. Western

10 et al., 2004; Bronstert et al., 2012). A common conclusion of this type of study is that "Soil moisture at the headwater scale exhibits huge spatial variability and single or even distributed TDR measurements yield non-representative data." as literally stated in Zehe et al. (2010, p.874 l.11-14) (Vereecken et al., 2008; Vanderlinden et al., 2012; Brocca et al., 2012).

Though large spatial variability seems to be a generic feature of soil moisture, there is also evidence that ranks of distributed soil moisture observations are largely stable in time as observed at the plot (Blume et al., 2009; Zehe et al., 2010), hillslope

15 (Brocca et al., 2007, 2009), and even catchment scale (Grayson et al., 1997). Given the partly contradicting nature of these findings, one may wonder whether the huge spatial variability of soil moisture dynamics is indeed the dead end for spatially distributed point sampling?

Soil moisture dynamically responds mainly to two forcing regimes namely rainfall driven wetting or radiation driven drying. The related controlling factors and processes differ largely and operate at different spatial and temporal scales. The soil moisture

20 pattern reflects thus the multitude of these influences and the strengths of the respective forcing vary over time (Bárdossy and Lehmann, 1998). This manifests in changes in the spatial covariance structure (Lark, 2012; Schume et al., 2003), either in form of changing nugget to sill ratios (spatially explained variance) (Zehe et al., 2010) or state dependent variogram ranges (spatial extent of correlation) (Western et al., 2004). In a homogeneous, flat and non-vegetated landscape the soil moisture pattern shortly after the event would be the imprint of the precipitation pattern and provide predictive information about its spatial

25 covariance. On the long term the spatial pattern and the related predictive information will gradually dissipate due to depletion of the gradients in soil water potentials by lateral soil water flows. On the other hand, in a heterogeneous landscape driven by spatially uniform block rain events, the spatial pattern of soil moisture would be a largely stable imprint of different landscape properties controlling through-fall, infiltration as well as vertical and lateral soil water redistribution.

Landscape heterogeneity is thus a perquisite for a temporally persistent spatial variability of soil moisture time series.

30 While most catchments are strongly heterogeneous, it is striking how spatially organized they are (Dooge, 1986; Sivapalan, 2003; McDonnell et al., 2007; Zehe et al., 2014; Bras, 2015). Spatial organization manifests for instance through systematic and structured patterns of catchment properties, such as a catena. This might naturally lead to a systematic variability of those processes controlling wetting and drying of the soil. One approach to diagnose and model systematic variability is based on

the covariance between observations in relation to their separating distance (Burgess and Webster, 1980) and geo-statistical interpolation or simulations methods (Kitanidis and Vomvoris, 1983; Ly et al., 2011; Pool et al., 2015).

A spatial co-variance function describes how linear statistical dependence of observations declines with increasing separating distance up to the distance of statistical independence. This is often expressed as experimental variogram. Geostatistics

5 relies on several assumptions such as second order stationarity (see e.g. Lark (2012) or Burgess and Webster (1980)), which are ultimately important for interpolation. Due the above-mentioned dynamic nature of soil moisture observations, the most promising avenue for interpolation would be a spatio-temporal geostatistical modeling of our data (Ma, 2002; De Cesare et al., 2002; Ma, 2003; Snepvangers et al., 2003; Jost et al., 2005).

However, here we take a different avenue, as we do not intent to interpolate. One of our goals is detecting dynamic changes

10 in the spatial soil moisture pattern. Following Sampson and Guttorp (1992) we relate the statistical dispersion of soil moisture observations to their separating distance to characterize how their similarity and predictive information declines with separating distance (see section 2). More specifically, we analyze temporal changes in the spatial dispersion of distributed soil moisture data. We want to find out whether typical patterns emerge in time, how those relate to the different forcing regimes and whether those patterns are recurrent in time. The latter is an indicator for predictability and (self) - organization in dynamic systems

15 (Wendi and Marwan, 2018; Wendi et al., 2018).

Zehe et al. (2014) argued that spatial organization manifests through a similar hydrological functioning. This is in line with the idea of Wagener et al. (2007) on catchment classification, or the early idea of a geomorphological unit hydrograph (Rodríguez-Iturbe et al., 1979; Sivapalan et al., 2011; Patil and Stieglitz, 2012). Recently, Loritz et al. (2018) corroborated the idea of Zehe et al. (2014) and showed that hydrological similarity of discharge time series implies that they are redundant.

20 Thus, they can be compressed without information loss (Weijs et al., 2013). More specifically, Loritz et al. (2018) showed that a set of 105 hillslope models yielded, despite their strong differences in topography, a strongly redundant runoff response. Using Shannon entropy (Shannon, 1948) Loritz et al. (2018) showed that the ensemble could be compressed to a set of 6 to 8 typical hillslopes without performance loss. Here we adopt this idea and investigate the redundancy of spatially distributed soil moisture data, which is a measure of how representative they are.

25 The core objective of this study is to provide evidence that distributed soil moisture time series provide, despite their strong spatial and persistent variability, representative information on soil moisture dynamics. More specifically, we test the following hypotheses:

– **H1:** Radiation-driven drying and rainfall-driven wetting leave different fingerprints in the soil moisture pattern, which manifests in temporal changes in the dispersion functions.

30 – **H2:** Both forcing regimes and their seasonal variability may be identified through temporal clustering of dispersion functions.

– **H3:** Spatial dispersion is more pronounced during and shortly after rainfall driven wetting conditions, which implies a stronger spatial similarity of soil moisture.
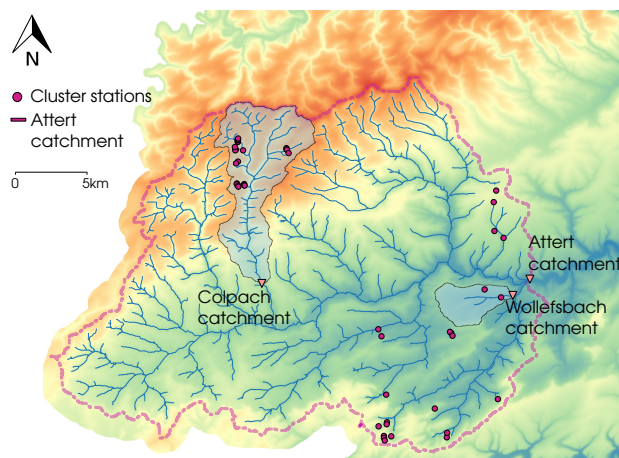
Figure 1. Attert experimental catchment in Luxembourg and Belgium. The purple dots show the sensor cluster stations installed during the CAOS project. Here we focus on those cluster stations within the Colpach catchment. Figure taken from Loritz et al. (2017).

– **H4:** Soil moisture time series are redundant, which implies they are compressible without information loss. However, the degree of compressibility is changing over time.

We test these hypotheses using a distributed soil moisture data set collected in the Colpach catchment as explained in the next section. In section 2 we give an overview of our method, including a general description of the clustering algorithm used. The result consists of three parts: spatial dispersion functions, temporal patterns in their emergence and some insights on generalization (or compressibility) of these functions. A discussion and summary of our method finish this work.

2 Methods

2.1 Study area and soil moisture data set

We base our analyses on the CAOS data set, which was collected in the Attert experimental watershed between 2012 and 2017 and is explained in Zehe et al. (2014). The Attert catchment is situated in western Luxembourg and Belgium (Figure 1). Mean monthly temperatures range from $18°$C in July to $0°$C in January. Mean annual precipitation is approximately 850 mm (Pfister et al., 2000). The catchment covers three geological formations, Devonian schists of the Ardennes massif in the northwest, a mixture of Triassic sandy marls in the center and a small area on Luxembourg Sandstone on the southern catchment border (Martinez-Carreras et al., 2012). The respective soils in the three areas are haplic Cambisols in the schist, different types of Stagnosols in the marls area and Arenosols in the sandstone (IUSS Working Group, 2006; Sprenger et al., 2016). The distinct differences in geology are also reflected in topography and land use. In the schist area, land use is mainly forest on steep slopes of the valleys, which intersect plateaus that are used for agriculture and pastures. The marls area has very gentle slopes and is mainly used for pastures and agriculture, while the sandstone area is forested on steep topography.
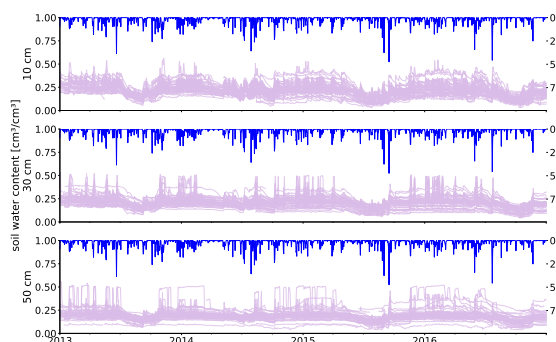
Figure 2. Soil moisture data overview. Soil moisture observations in 10 cm (top), 30 cm (middle) and 50 cm (bottom).

The experimental design is based on spatially distributed, clustered point measurements within replicated hillslopes to infer on their typical controls on hydrological dynamics. To this end 4-6 sensor clusters were placed in hillslopes at the top, midslope and hill foot sector along the anticipated flow paths. Within each of those clusters, soil moisture was recorded in three replicate profiles in 10, 30 and 50 cm depth using Decagon 5TE sensors. While the design was stratified to sample different geological

5 settings (schist, marls, sandstone), different aspects and land use (deciduous forest and pasture), we focus here on those sensors installed in the schist area. In total we used 19 cluster locations and thus 57 profiles of 171 time series.

Soil moisture in the Colpach catchment exhibits high but temporally persistent spatial variability (Fig. 2). For each point in time a wide range of water content values can be observed across the catchment. The range of soil moisture observations is generally wider in winter than in summer. From visual inspection it appears, that the heterogeneity in observations is not

10 purely random but systematic as the measurements are rank stable throughout long periods. One has to note that the different cluster locations differ in is aspect, slope and landuse. From the data shown in figure 2, two sensors have been removed. Both measured in 50 cm and can be seen in the figure at the very bottom. Both recorded values close to or even below $0.1 \text{ cm}^3\text{cm}^{-3}$ for the whole period of four years. Additionally, the plateaus lasting for a couple of days at constant $0.5 \text{ cm}^3\text{cm}^{-3}$ in 50 cm and 30 cm were removed.

15 ## 2.2 Dispersion of soil moisture observations as function of their distance

We seek for the temporal emergence and relaxation of spatial dependencies/similarities between the observation points. For our analysis the selected data was aggregated to mean daily soil moisture values θ. This data set is further aggregated within a moving window of one month as described by equation (1).

$$z_x(t) = \frac{\sum_t^{t+b} \theta_x}{b} \tag{1}$$

For each time step $t = 1, 2, \ldots, (L - b)$ and a time series length of L in days. For our analysis, b is set to 30 (days), x is the observation location. We tested different window sizes, as we expect that different processes control the emergence of spatial dependence at different temporal scales. The chosen window size was most suitable to detect seasonal effects.

To estimate the spatial dependence structure between observations, we relate their pairwise separation distance to a measure of pairwise similarity. Here, we further define the statistical spatial dispersion to be a measure for spatial similarity. We compare the empirical distribution of pairwise value differences at different distances. Statistically, a more dispersed empirical distribution is less well described by its mean value. Thus, observations taken at a specific distance are more similar in value, if they are less dispersed.

To estimate the dispersion, we use the Cressie-Hawkins estimator (Cressie and Hawkins, 1980). This estimator is more robust to extreme values and the contained power transformation handles skewed data better than estimators based on the arithmetic mean (Bárdossy and Kundzewicz, 1990; Cressie and Hawkins, 1980). The estimator is given by equation (2):

$$a_t(h) = \frac{1}{2} \left(\frac{1}{N(h)} \sum_{i,j} \sqrt{|z(x_i) - z(x_j)|} \right)^4 \left(0.457 + \frac{0.494}{N(h)} + \frac{0.045}{N^2(h)} \right)^{-1} \tag{2}$$

with $z(x_{i,j})$ given by equation (1) for each pair of observation locations x_i, x_j. h is the separating distance lag between these points and $N(h)$ the number of points formed at the given lag h. To calculate the dispersion at one distance, it is necessary to form distance lag classes, as there are hardly two observation point pairs x_i, x_j at the exact same separating distance. 10 classes were formed with a maximum separation distance of 1200 m. Observation point pairs further apart than 1200 m are most likely located on different hillslopes. Here, the distance is not a good depended variable for similarity anymore, as observation locations might share land use or terrain aspect. This would result is smaller dispersion, but not in dependence of distance. The lag classes are not equidistant, but with a fixed $N(h)$ for all classes. This is further discussed in section 2.3.

2.3 Clustering of dispersion functions

We analyzed how and if meaningful spatial dispersion functions emerge and whether those develop into stable configurations. To tackle the hypotheses formulated in the introduction a clustering is applied to the dispersion functions derived for each window. The clustering algorithm should form groups of functions that are more similar to each other than to members of other clusters. The similarity between two dispersion functions is calculated by the Euclidean vector distance between the dispersion values forming the function. This distance is defined by equation (3):

$$d(\boldsymbol{u}, \boldsymbol{v}) = \sqrt{(\boldsymbol{u} - \boldsymbol{v})^2} \tag{3}$$

with $\boldsymbol{u}, \boldsymbol{v}$ being two dispersion function vectors. This is the Euclidean distance of two points in the (higher dimensional) value space of the dispersion function's distance lags. Two identical dispersion functions are represented by the same point in this value space and hence their distance is zero. Thus, distance lags are not equidistant, as this could lead to empty lag classes. Empty lag classes result in an undefined position in the value space, which has to be avoided. The clustering algorithm cannot use the number of clusters as a parameter, as this can hardly be determined a priori. One clustering algorithm meeting
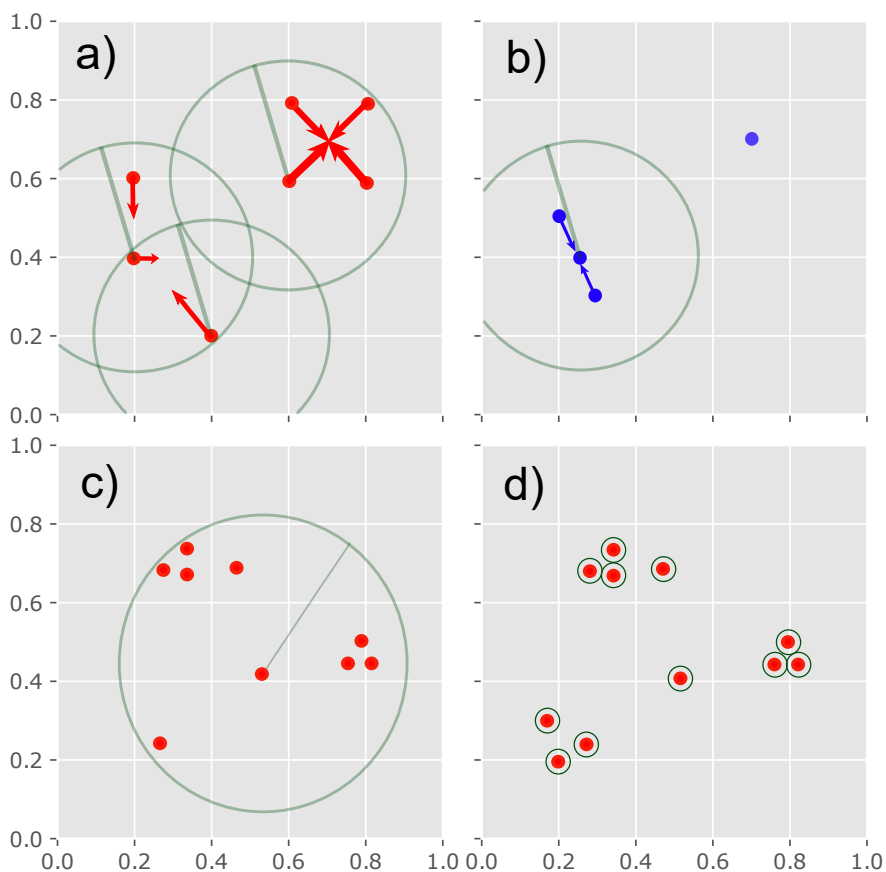
Figure 3. Schematic procedure of the Mean shift algorithm in $\mathbb{R}^2$. **a):** Red dots indicate hypothetical samples to be clustered. The circles are illustrating the flat kernel of the centered sample at first iteration. The bandwidth parameter is illustrated by the radius. The red arrows indicate the shift of the respective sample onto the geometric mean of all samples inside the current kernel. Note that three points on the left side are shifted differently, as the upper and lower point do not lie in each others kernel. **b):** Second iteration step after a). The blue dots are the shifted means from a) and will be used as the input sample for the next iteration. The procedure finishes when no points can be 'shifted' anymore. **c):** Example of a large bandwidth (radius), which will result in only one cluster at convergence. **d):** Example of a too small bandwidth, where no point will be shifted at all.

these requirements is the *Mean shift* algorithm (Fukunaga and Hostetler, 1975). The actual code implementation is taken from Pedregosa et al. (2011), which follows the Comaniciu and Meer (2002) variant of Mean shift.

2.3.1 Mean shift algorithm

Mean shift starts by forming a cluster for each sample on its own. Here, a *sample* corresponds to one dispersion function. We

5   will illustrate the fundamental mechanism of the algorithm for the two-dimensional case, as the samples can easily be plotted

in $\mathbb{R}^2$ (see figure 3 a, b). Mean shift works iteratively. In each iteration, a window is shifted over all samples, which can be thought of as coordinate points in the two-dimensional case (see fig. 3 a). This window is called a *kernel* that is controlled by a size parameter called *bandwidth*. In the two-dimensional case, this can be thought of a circle with a radius set to the given bandwidth as shown in figure 3 a. In each kernel position, the center of sample density is calculated and the current sample is

5   *shifted* onto this point, which is the new cluster mean, called cluster centroid. On the next iteration, the newly created cluster centroids are used as the new (input) samples, as shown in figure 3 b. The iterations stop when the shifting means converge (centroids do not change their position anymore). We substitute the centroids calculated on the last iteration by the original sample closest to this point. Thus, we choose the most representative dispersion function for the cluster.

Mean shift is sensitive to the selected bandwidth. Two clusters whose centroids are within one bandwidth length will be

10   shifted into a combined cluster before convergence is met. As a result a bandwidth parameter chosen too big might classify all samples as a single cluster as indicated in figure 3 c. In case the bandwidth is chosen too small many tiny clusters with just a few members will be the result. Figure 3 d shows an extreme example, where no sample will shift anywhere. We tested different bandwidth parameters at a few examples and set the bandwidth to the 30% percentile of all pairwise vector distances between the dispersion functions of one year and depth. We chose the so-called flat kernel as a kernel, which would result in a

15   circle in the two-dimensional case and a N-dimensional sphere in the $\mathbb{R}^N$, where N is the number of lag classes used for the dispersion function.

2.4   Clusters compression based on the cluster centroids

The next step is to generate a representative dispersion function for each cluster. The straightforward representative is the cluster centroid. All dispersion functions are calculated with the same parameters, including the maximum separating distance

20   of 1200 m. At larger lags we found indeed instances of declining dispersion values, because we related point pairs at different hillslopes which were located in similar hillslope positions (e.g. up-slope or down-slope), land use and aspects. To facilitate the comparison of the dispersion functions we decided to monotonize them. In geostatistics this is usually done through fitting of a theoretical variogram model to the experimental variogram, which assures monotony and positive definiteness. Here we do not force a specific shape by a fitting a model function. Instead we use the technique of monotonizing the cluster centroid

25   as suggested by Hinterding (2003) using the PAVA-Algorithm (Barlow et al., 1972). The implementation is from Pedregosa et al. (2011). This way, the final compressed dispersion functions are monotonically increasing, while still reflecting the shape properties of the cluster members. We suggest that the number of representative functions needed to cover all observed spatial dispersion functions of one calendar year is used as a measure for the strength of spatial organization. Additionally, it is insightful to judge the information loss that goes along with this compression, as a high compression with little information

30   loss is understood as a manifestation of spatial and temporal organization of soil moisture dynamics.

In line with Loritz et al. (2018) we use the Shannon Entropy as measure for the compression without information loss. It requires treatment of the clusters as discrete probability density functions, which in turn implies a careful selection of an appropriate binning. Motivated by Loritz et al. (2018), we use as minimum meaningful separating distance the uncertainty in the dispersion function, calculated via uncertainty propagation (see section 2.5.1).

2.5 Uncertainty propagation and compression quality

2.5.1 Uncertainty propagation

Soil moisture measurements have considerable measurement uncertainty. The complex medium to observe leads to a typical measurement precision of 1 - 3 $cm^3 cm^{-3}$ reported by manufacturers. For our uncertainty propagation we assume an absolute

5   uncertainty/measurement error $\Delta \theta$ of 0.02 $cm^3 cm^{-3}$.

Next we propagate these uncertainties into the dispersion functions and the distances among those. As we assume the measurement uncertainties to be statistically independent, we use the Gaussian uncertainty propagation to calculate error bands / margins . In a general form, for any function $f(z)$ and an absolute error Δz the propagated error Δf can be calculated. In our case z is itself a function of x, the observation location, and the general form is given by equation (4).

$$\Delta f = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial f}{\partial z(x_i)} \Delta z(x_i) \right)^2} \tag{4}$$

To apply equation (4) for our method, the measurement uncertainty $\Delta \theta$ is propagated into the dispersion estimator given by equation (2). The dispersion estimator is derived with respect to $z(x)$ and following equation (1), the uncertainty in $z(x)$, Δz, is denoted as $\Delta z = \Delta \theta = 0.02 cm^3 cm^{-3}$. Then, with given Δz, we can propagate the uncertainty into the dispersion function. As the dispersion function is a function of the spatial lag h, we need to propagate the uncertainty Δa (uncertainty of dispersion

15   estimator) for each value of h. At the same time, following equation (2), for each h, $z(x_i) - z(x_i + h)$ is a fixed set of point pairs. Instead of propagating uncertainty through equation (2), we can substitute $z(x_i) - z(x_i + h)$ by δ, the pairwise differences, for each value of h. The uncertainty $\Delta \delta$ is given by equation (5)

$$\Delta \delta = \sqrt{\Delta z_i^2 + \Delta z_{i+h}^2} = \sqrt{2} \Delta z \tag{5}$$

The uncertainty of dispersion Δa is then defined by equation (6):

$$\delta a = \frac{\partial a}{\partial \delta} \Delta \delta \tag{6}$$

$$= 2c \left(\frac{1}{N} \sum_{i=1}^{N} (|\delta_i|)^{\frac{1}{2}} \right)^3 \cdot \frac{1}{N} \left(\sum_{i=1}^{N} |\delta_i|^{-1} \right)^{\frac{1}{2}} \cdot \Delta \delta \tag{7}$$

Where the factors from equation (2) that stay constant in the derivative are denoted as c and defined in equation (8). In line with equation (2) N is the number of observation pairs available for a given lag class h and therefore constant for a single calculation. $\Delta \delta$ and δ are the substitutes for z, as described above (see equation (5)).

$$c = \frac{1}{2} * (0.457 + \frac{1}{N} + \frac{0.045}{N^2})^{-1} \tag{8}$$

The last step is to propagate the uncertainty into the distance function as defined in equation (3). This distance is the foundation of Mean shift on the one hand and the compression quality calculations presented in section 2.5.2 below on the other hand.

Following equation (4) the propagated uncertainty ΔD can be calculated by the derivative of equation 3 with respect to each of the vectors multiplied by the corresponding value of Δa, which results in equation (9):

$$\Delta d_{\boldsymbol{u},\boldsymbol{v}} = \sqrt{\left(\frac{\delta d}{\delta \boldsymbol{u}}\boldsymbol{\Delta u}\right)^2 + \left(\frac{\delta d}{\delta \boldsymbol{u}}\boldsymbol{\Delta u}\right)^2} \tag{9}$$

$$= \sqrt{\frac{1}{2}\sum_{i=1}^{n}(|\boldsymbol{u}-\boldsymbol{v}|^{\frac{1}{2}})\sum_{i=1}^{n}((2(|\boldsymbol{u}_i-\boldsymbol{v}_i|)\Delta\boldsymbol{u}_i)^2 + (2(|\boldsymbol{u}_i-\boldsymbol{v}_i|)\Delta\boldsymbol{v}_i)^2)} \tag{10}$$

Where $\boldsymbol{u},\boldsymbol{v}$ are two spatial dispersion function vectors as defined and used in equation (3). $\boldsymbol{\Delta u}, \boldsymbol{\Delta v}$ are the vectors of uncertainties for $\boldsymbol{u},\boldsymbol{v}$, where Δv_i is the uncertainty propagated into the ith lag class as shown in equation (6). n is the number of lag classes and, thus the length of each vector $\boldsymbol{u},\boldsymbol{v},\boldsymbol{\Delta u},\boldsymbol{\Delta v}$.

Equation (9) is applied to all possible combinations of dispersion functions $\boldsymbol{u},\boldsymbol{v}$ to get all possible uncertainties in dispersion function distances.

2.5.2 Compression quality

The Shannon entropy (Shannon, 1948) of all pairwise dispersion function distances is used as measure for information content. The Shannon entropy of a discrete probability density function of states (patterns in this case) is maximized for the uniform distribution. It corresponds to the number of yes/no questions one has to ask to determine the state of a system. The minimum entropy is zero, which corresponds to the deterministic case that the system state is always known. A common way to define spatial organisation of a physical system is through its distance from the maximum entropy state (Kondepudi and Prigogine, 1998; Kleidon, 2012). The deviation of the entropy of the dispersion functions in a cluster from its maximum value is thus a measure for their redundancy and thus similarity.

For a discrete frequency distribution of n bins, the information entropy H is defined as:

$$H = -\sum_{n} p_n log_2(p_n) \tag{11}$$

Where p_n is the relative probability of the nth bin. H is calculated for each depth in each year individually to compare the information content across years and depths.

To assure comparability we use one binning for all calculations of H (across years and depths). To achieve this, all pairwise distances between all spatial dispersion functions of all four years in all three depths are calculated. The discrete frequency distribution is formed from 0 up to the global maximum distance (between two dispersion functions) calculated using equation (3). The bins are formed equidistant using a width of the maximum function distance that still lies within the error margins calculated using equation (9). Thus, the information content of the spatial heterogeneity is calculated with respect to the expected uncertainties. This way we can be sure to distinguish exclusively those spatial dispersion functions that lie outside the error margins.

The Kullback-Leibler divergence measures the information loss due to compression. To compress the series of dispersion functions, each cluster member is expressed by its centroid function. Now, we need to calculate the amount of information lost
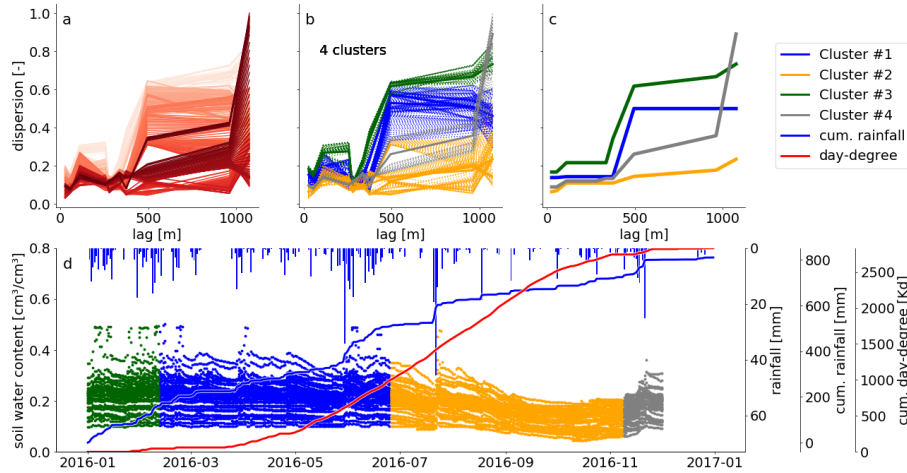
Figure 4. Spatial dispersion functions in 30 cm of 2016 based on on a window size of 30 days.

a): Spatial dispersion function for each position of the moving window. The red color saturation is indicating the window position. The darker the red the higher the Julian day.

b): The same dispersion functions as presented in a). Here the color indicates cluster membership as identified by the Mean shift algorithm.

c): Compressed spatial dispersion information represented by corrected cluster centroids. The colors match the clusters as presented in b).

d): Soil moisture time series of 2016 in 30 cm depth. The colors identify the cluster membership of the spatial dispersion function of the current window location and is matching the colors in b) and c). The bars on the top show the daily precipitation sums. The solid blue line is the cumulative daily precipitation sum and the red line the cumulative sum of all mean daily temperatures $> 5°$C.

in this process. To calculate the mean information content of the compressed series each cluster member is substituted by the respective cluster centroid. This substitution is obviously not a compression in a technical sense, but necessary to calculate the Kullback-Leibler divergence. Then a frequency distribution for compressed series X and the uncompressed series Y can be calculated. The Kullback-Leibler divergence D_{KL} of X, Y is given in equation (12):

$$D_{KL}(X,Y) = H(X||Y) - H(Y) \tag{12}$$

Where $H(X||Y)$ is the cross entropy of X and Y and defined by equation (13):

$$H(X||Y) = \sum_{x \in X} p(x) * log_2 p(y) \tag{13}$$

Where $p(x)$ is the empirical non-exceedance probability of the frequency distribution X and $p(y)$ of Y, respectively.

3 Results

3.1 Dispersion functions over time

Figure 4 a) shows the spatial dispersion functions for all moving window positions in 2016 for the 30 cm sensors. The position of the moving window in time can be retraced by the line color, more reddish means a later Julian day. Each of the spatial
5 dispersion functions relates the dispersion for all pairwise observations to their separating distance in the corresponding lag class. Dispersion declines with separating distance, as small values correspond to observations which have similar values while large values suggest the opposite. As expected, the dispersion is a suitable metric for similarity/dependency of observations.

The spatial dispersion functions occur in several distinct shapes with each of these shapes corresponding to a certain period in time. More specifically from Figure 4 a) one can identify groups of functions of similar color plotting close to each other.
10 Dispersion functions of similar red saturation, which reflects proximity in time, are also similar in shape, and this in turn reflects similar spatial patterns. Similar dispersion functions were grouped using the Mean shift clustering algorithm (Fig. 4 b); here, the color indicates the cluster membership.

To provide further insight on the temporal occurrence of cluster members, we colored the soil moisture time series according to the color codes of the identified clusters (Fig. 4 d). The blue parts of the soil moisture time series were classified into Cluster
15 #1, while the the orange was classified into Cluster #2. Note that cluster memberships are constant for long periods of time, which means that also the soil moisture patterns are persistent over these periods. As new observations did not change the patterns during these periods, they were redundant in time.

As the spatial dispersion functions in the presented example are redundant in time, we compressed the information by replacing the dispersion function within one cluster by the cluster centroid. All four representative functions shown in Figure 4
20 c) exhibit increasing dispersion with separating distance. For the blue and green cluster this happens step-wise at a characteristic distance of 500 m. That reminds us of a Gaussian variogram, which can also show a step-wise characteristic. The small grey cluster shows an increase at 500 and another one at 1000 m separating distance. In contrast the orange cluster, however, shows a only a gentle increase with distance.

In the vegetation period observations are similar even at large separating distances. As the orange cluster (Figure 4 c and
25 d) covers significant parts of the vegetation period, the influence of vegetation on spatial soil water dynamics is considered to be dominant. Interestingly, dispersion functions in this cluster start with small values that only gently increase with separating distance. That means soil moisture becomes more homogeneous. Outside of the vegetation period, different spatial patterns can be observed, with increasing dissimilarity with separating distances. The part of the blue cluster overlapping with vegetation period shows still higher soil moisture values. The transition to the orange cluster sets in as the soil moisture drops (Fig. 4 d).
30 This suggests that vegetation influences, such as root water uptake, smooth out variability in soil water content, leading to a more homogeneous pattern in space, as further discussed in section 4.3.
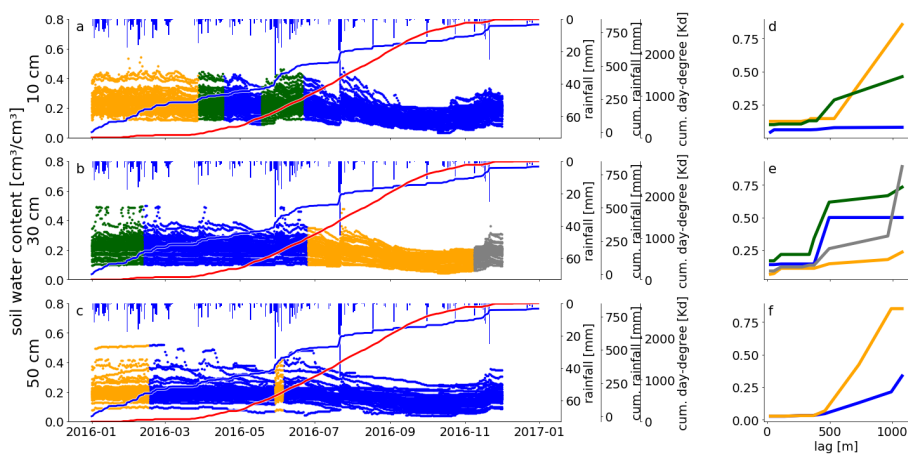
Figure 5. Soil moisture time-series of 2016 in all three depths with respective cluster centroids. The three rows show the data from 10 cm (a), 30 cm (b) and 50 cm (c). The colors indicate the cluster membership of the corresponding dispersion function of the respective window position. The cluster centroids for the respective depth is shown in d, e and f.

3.2 Dispersion time series as a function of depth

Figure 5 shows the time series of the dispersion functions in all depths in the year 2016. Note that the coloring between the sub-figure is arbitrary, due to Mean shift, that means there is no connection between the orange cluster between the three figures.

5   In comparison to the dispersion functions in 30 cm (Fig. 5 b) the soil moisture signal in 10 cm (Fig. 5 a) is more variable in time. A look at the centroid of the orange winter cluster (Fig. 5 d) reveals a higher spatial heterogeneity in winter and spring at large separating distances. At the same time the observations get spatially more homogeneous in summer, particularly when the blue cluster emerges, because the dispersion at large lags decreases significantly. We can still find a summer-recession cluster in 10 cm, but compared to the depth of 30 cm we also find this spatial footprint of continuous drying earlier in the year around

10   May. This is likely due to a higher sensitivity to rising temperatures. Note that during May there was only little rainfalls and the soil moisture is already declining. This blue cluster shows very small dispersion values for all separating distance classes (Fig. 5 d). Thus, for both the 10 and the 30 cm depth the values in summer are similar, even at large distances.

    The green clusters emerge with strong rainfall events after longer previous dry spells (Fig. 5). We would have expected a third occurrence at the beginning of August, but the soil may already be too dry to bear a detectable dependency on separating

15   distance (Remember that the blue cluster does not show increasing dispersion with distance).

    Observations at 50 cm depth show a clear spatial dependency throughout the whole year. We cannot identify a summer cluster, Mean shift yielded two clusters and rainfall forcing does not have a clear influence on their occurrence or transition. The 50 cm dispersion functions (Fig. 5 f) show a clear spatial dependence and are similar. Not only the soil moisture observations have become much more homogeneous with depth, also the dispersion functions are more similar in shape. The magnitudes of
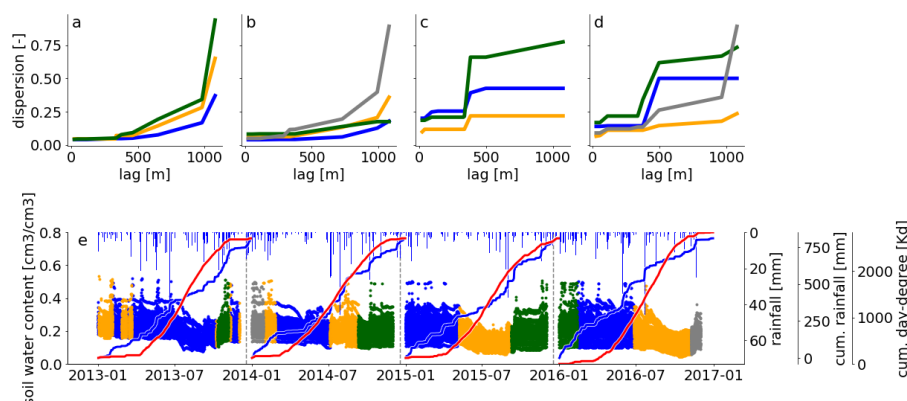
Figure 6. Soil moisture timeseries of all years in 30 cm depth (e) and the respective cluster centroids (a-d). The colors of the soil moisture data indicate the cluster membership of the corresponding dispersion function of the respective window position and correspond with the color of the cluster centroid (in a-d). The cumulative rainfalls (blue) and cumulative temperature sums (red) are shown for each year individually.

the dispersion are larger for the orange cluster than for the blue one. We find dispersion functions with characteristic length of 500 m and the blue cluster persists throughout most of the year. The orange cluster occurs in the cool and wet start of the year, showing a larger dispersion and thus stronger dissimilarity at larger lags (Fig. 5 f). Interestingly this cluster occurs again in early June after an intense rainfall period. However, a similar rainfall period in August does not trigger the emergence of this

5 orange cluster as the top soil above 50 cm is so dry, that even this strong wetting signal does not reach the depth of 50 cm (Fig. 5 c). This behavior reveals the low pass behavior of the top soil, which causes a strong decoupling of the soil moisture pattern in 50 cm depth from event scale changes.

3.3 Recurring spatial dispersion over the years

Table 1 summarizes the most important features of the clustering for all observation depths. Soil moisture patterns and their
10 clustering appear generally to be clearer for 2015 and 2016. The vegetation period is more often characterized by a typical cluster and dispersion functions more often reveal a clear spatial dependency. In some cases, mostly 2014, no spatial dependency of dispersion functions could be observed throughout the whole year. Space-time patterns became more stable in 2015 and 2016. Note that annual rainfall sums were higher in 2013 and 2014, while 2015 and 2016 had significantly more precipitation in the first half of the year followed by a dry summer (compared to 2013 and 2014).

15 To further illuminate inter annual changes in soil moisture patterns we present the time series of cluster memberships for the sensors in 30 cm for the entire measuring campaign in Figure 6. From this example it becomes obvious that patterns are recurring. 2013 and 2014 cluster centroids look different from the following two years. Dispersion values increase with distance in all centroids in 2013 and 2014, while 2015 and 2016 show a sudden increase at 400 - 500 m (Fig. 6 a-d). 2015 and 2016 are segmented by Mean shift in a similar way, and cluster centroids reveal that the green clusters in both years are
20 actually the same. This cluster emerges with the the occurrence of the largest rainfall event in the observation period and lasts

Table 1. Qualitative description of method success in all years and depths. The results from years other than 2016 and all depths were inspected visually and are summarized here for sake of completeness. The first two columns identify the year and sensor depth. N gives the number of clusters found by Mean shift. The remaining three columns state if specific features existed in the given result. *Vegetation period* marks whether or not the vegetation period was roughly covered by a single, or two, clusters. *Spatial structure*: Does a dependency of dispersion on distance exist **outside** the vegetation period? *rainfall transition*: Were cluster transitions accompanied by a rainfall event in close (temporal) proximity? This feature is marked 'yes' if it was more often the case than it was not.

Year	depth	N	vegetation period	spatial structure	rainfall transition
2013	10	4	yes	no	no
2013	30	3	no	yes	yes
2013	50	6	no	yes	yes
2014	10	3	yes	no	yes
2014	30	4	no	yes	no
2014	50	5	no	yes	yes
2015	10	3	yes	yes	no
2015	30	3	yes	yes	yes
2015	50	3	no	yes	no
2016	10	3	no	yes	yes
2016	30	4	yes	yes	yes
2016	50	2	no	yes	no

for around 5 months. All dispersion functions within this cluster look more or less the same (see Appendix, Fig. A2, b). Similar observations can be made between 2014 and 2015. Here, the green and blue cluster seem to be an inter-annual cluster. However, in contrast to to 2015/2016 the dispersion functions here are of different shape (see Appendix, Fig. A1, b). Hence, the cluster transition indicated between 2014 and 2015 is indeed a real transition. When looking at cluster memberships throughout the whole period, the division into calendar years is rather meaningless, while the division in hydrological years is much more appropriate, as it is reflected by the cluster membership and its changes.

Distinct summer recessions in soil moisture are only identified in 2015 and 2016. Evapo-transpiration (indicated by the temperature-sum curves in Fig. 6 e) is dominating over rainfall input (blue sum curve) in the soil moisture signal. Mean shift could identify a significantly distinct spatial dependency in dispersion, as shown by the two orange centroids in figure 6 c) and d). They are both distinct from the other centroids in the same period by showing only a gentle increase in dispersion. A likely reason for the absence of a distinct summer recessions in 2014 is the rather wet and cold spring and summer, as can be seen from the flat rainfall-sum curve during that period (Fig. 6 e). In 2013 this identification did not work. Possible reasons are provided in the discussion section 4.5.

Table 2. Information content and information loss due to compression. N gives the number of clusters identified for the given year and depth. The information content is given as Shannon entropy H, which is the expectation value of information in information theory. The information loss after compression is given by the Kullback-Leibler divergence between the compressed and uncompressed series of dispersion functions. The last column relates D_{KL} to H

Year	Depth	N	H [bit]	2^H	D_{KL} [bit]	$D_{KL} * (H + D_{KL})^{-1}$
2013	10 cm	4	0.97	0.93	0.44	0.31
2013	30 cm	3	1.49	2.22	0.06	0.04
2013	50 cm	6	2.0	3.98	0.13	0.06
2014	10 cm	3	1.35	1.83	0.22	0.14
2014	30 cm	4	1.57	2.47	0.3	0.16
2014	50 cm	5	2.44	5.95	0.28	0.1
2015	10 cm	3	1.87	3.52	0.18	0.09
2015	30 cm	3	1.18	1.39	0.09	0.07
2015	50 cm	3	2.39	5.71	0.9	0.27
2016	10 cm	3	2.49	6.2	0.76	0.23
2016	30 cm	4	1.44	2.07	0.02	0.02
2016	50 cm	2	3.21	10.32	2.5	0.44

3.4   Redundant spatial dispersion functions

We calculated the Shannon entropy for all soil moisture time series for all years and depths (Table 2). As explained in section 2.5.2 this reflects the intrinsic uncertainty of the clusters. Most entropy values are within a range of $1 < H < 2.5$. The maximum possible entropy for a uniform distribution of the used binning is 3.55. The Kullback-Leibler divergence D_{KL} is a measure for

5 the information loss due to the compression of the cluster onto the centroid dispersion function. In the overwhelming majority of the cases, the information loss is one magnitude smaller than the intrinsic uncertainty and ranges between $0.01 < D_{KL} < 0.4$. Hence, the information loss due to compression is negligible. There is one exception in 2016 (50 cm).

The clusters obtained in 30 cm for the year 2016 (compare 3.1) showed an entropy of 1.44. Compared to this value, the Kullback-Leibler divergence caused by compression of only 0.02 is small, if not negligible. The last column of table 2 relates

10 D_{KL} to the overall uncertainty. It contributes less than one third in almost all cases (2016, 50 cm is the only exception). In the majority of the cases it does not contribute more than 20%.

According to equation (11) the Shannon entropy is derived from an discrete, empirical probability distribution. As it is calculated using the binary logarithm, 2^H gives the amount of discriminable states in this discrete distribution. This number of states is deemed to be a reasonable upper limit for the number of clusters for Mean shift. A higher number of clusters than 2^H

15 appears meaningless, and this assures that only those clusters are separated, which are separated by a distance large than the margin of uncertainty.

4 Discussion

In line with our central hypothesis **H1** - that radiation-driven drying and rainfall-driven wetting leave different fingerprints in the soil moisture pattern which manifests in temporal changes in the dispersion functions - we found strong evidence that soil water water dynamics is organized in space and time. And more importantly our findings reveal that this organization is not

5   static but exhibits dynamic changes which are closely related to seasonal changes in forcing regimes. A direct consequence is that soil moisture observations are despite of their strong spatial heterogeneity, quite predictable in time. This manifests through a temporal persistence of characteristic spatial patterns, which are expressed using the statistical dispersion of soil moisture observations in dependence of their separating distance. These characteristic patterns can be captured through clustering of the dispersion functions using their vector distance to express their similarity. As measure for the degree of organization we

10 used the information loss that goes along with the compression of the entire cluster on the most representative cluster member. Here we found that this compression adds negligible uncertainty compared to the intrinsic uncertainty, caused by propagation of measurements uncertainties. We thus conclude that soil moisture observations are heterogeneous, but for long periods in an spatially organized manner, that persists in many cases for months.

In the following we will discuss our main findings that similarity in space leads to dynamic similarity in time, the way

15 we utilized the measurement uncertainty to determine the information content and how two different processes forcing soil moisture dynamics induce two fundamentally different spatial pattern.

4.1 Spatial similarity persists in time

We related the dispersion of pairwise point observations to their separating distance. For brevity and due to their shape we called these relationships *dispersion functions*. We emphasize that this term is not meant in a strict sense and no mathematical

20 functional relationship, analogous to a theoretical model, has been fitted to the experimental dispersion functions. Despite the fact that the presented functions are empirical, they show clear, recurrent shapes in many occasions.

We found spatial similarity to persist in time. We calculated the dispersion functions using a moving window approach resulting in one dispersion function for each day as the starting position of one window. These functions describe the spatial dependence of dispersion and thus the average similarity of observations as function of their separating distance. These were

25 grouped into clusters based on their vector distance. In line with **H2** - that both forcing regimes and their seasonal variability may be identified through temporal clustering of dispersion functions - the results (Fig. 4, 6, 5) provided evidence that similar dispersion functions emerge in fact very closely in time. Generally they appeared in continuous periods or blocks in time and their changes coincided with changes or a switch in the forcing regimes. In case we can relate the emergence of such a cluster more quantitatively to the nature and strength of a specific forcing event/process, we can analyze for how long this

30 event/process imprints the spatial pattern of soil moisture observations. Or in other words: we can analyze how long a catchment state *remembers* a disturbance.

Although cluster memberships occur in temporally continuous blocks in all depths throughout all years, for a few cases we could not relate their emergence to distinct changes in the analyzed forcing. This implies that **H2** needs to be partly revised. At

the same time especially Figure 6 showed that soil moisture observations in 30 cm are spatially heterogeneous, as can be seen by the wide variability of soil moisture observations at each time step. But the dispersion functions reflect spatial organization which persists and even recurs in time.

Dispersion functions in 50 cm show a clear spatial dependency throughout the year, with distinct differences within and
5   outside the vegetation period. In 50 cm of 2016 this is different. We find essentially two clusters that do not separate the data series by vegetation period. The shape of the two centroids (Fig. 5, f) is similar, only at large distances they differ in value. That means, from orange to blue cluster observations became more similar at large separating distances. Heavy rainfall disturbs this pattern leading to stronger dissimilarity at larger distances and that pattern lasted for a couple of weeks. Then, evapotranspiration driven drying smooths out soil moisture variability and during a similarly strong rainfall event in summer,
10 the cluster can not emerge again as the soil is already too dry. The soil acts as a low pass filter here, which filters out any change in state above a specific frequency. This happens mainly due to dispersion of the infiltrating and percolating water through, or storage in the soil matrix. By the time it reaches the deep layers, spatial differences are eliminated. This kind of behaviour is well known and was already reported in the early 90s (Entekhabi et al., 1992; Wu et al., 2002). More recently (Rosenbaum et al., 2012) "found large variations in spatial soil moisture patterns in the topsoil, mostly related to meteorological forcing. In
15 the subsoil, temporal dynamics were diminished due to soil water redistribution processes and root water uptake". In the same year, Takagi and Lin (2012) analyzed a dataset of 106 locations in a forested catchment in the US for spatial organization in soil moisture patterns. They found a seasonal change in more shallow depths (30 cm), controlled by rainfall and evapotranspiration. In deeper depths patterns became more temporally persistent. All these findings are in line with our results and conclusions.

4.2   Uncertainty analysis

20 We related the evaluation of compression quality directly to the measurement uncertainty. This was achieved by Gaussian error propagation of measurement uncertainty into the dispersion functions and their distances. The latter allowed definition of a minimum separable vector distance between two dispersion functions that are different with respect to the error margin. We based the bin width for calculating the Shannon entropy on this minimum distance, because this assured that the Shannon entropy gives the information content of each cluster *with respect to the uncertainty*. On this basis it was possible to assess
25 compression quality not only by the number of meaningful clusters found, but also on the information lost due to compression with respect to uncertainty.

In line with **H4** spatial patterns of soil moisture were found to be persistent over longer time periods. In many instances we found onlyfew clusters within one year and compression was possible with small if not negligible information loss. That means, while one cluster persists in time, small daily changes in the dispersion functions do not add significant new information
30 on spatial similarity. Hence, the whole cluster can be represented by only the centroid function. We conclude that this is a manifestation of strongly organized state which persist for a considerable time, as most observations were redundant during these periods.

From equation (11) it can be seen, that the Shannon entropy changes substantially with the binning. Therefore, it is of crucial importance to define a meaningful binning based on objective criteria. We suggest that only a discrimination into bins larger

than of error margins makes sense, because smaller differences cannot be resolved based on the precision of the sensors. For the application presented in this work, this is important because otherwise one could not compare the compression quality between depths or years, as different binnings lead to different Shannon entropy values, even for the same data. Hence, it would be difficult to analyze effects or differences of spatial dispersion in depth or over the years. We thus conclude that that

5 the Shannon entropy should only be used if the measurement uncertainties of the data are properly propagated.

We provided an example of how the quality of a compression can be assessed. Instead of considering the number of clusters (compression rate) only, we linked the compression rate to the resulting information loss. We could show that in the majority of the cases substantial compression rates could be achieved, which are accompanied by negligible information losses. We thus suggest that the trade off between compression rate and information loss should be used as compression quality measure.

10 **4.3 Different dominant processes lead to different patterns**

Outside of the vegetation period, we found a recurring picture of spatial dispersion functions with characteristic lengths clearly smaller than the typical extent of hillslopes. Dispersion functions were calculated in three depths for every day throughout four years. In the overwhelming majority there is an characteristic length at which the dispersion function shows a sudden rise in dispersion. For spatial lags smaller than this distance the dispersion is usually very small. Higher lags show much higher and

15 more variable dispersion values. This characteristic length is approx. 500 m. This corresponds to a common hillslope length for the Colpach catchment. During the vegetation period variability at large separating distance was smoothed out. Dispersion functions were flattened at large distances suggesting similarity even at distances larger than typical slope length. We thus conclude that there is dependence of the dispersion on the rainfall pattern, which is reflected in their shape and characteristic lengths. In the vegetation period, this heterogeneity is reduced by the vegetation and soil moisture values are similar at all

20 distances. This confirms **H2** and suggest that vegetation is the dominant factor in smoothing out soil moisture variability. A similar conclusion is drawn by Meyles et al. (2003), who identified 'preferred states in soil moisture' (Grayson et al., 1997; Western and Grayson, 1998; Western et al., 1999) and could relate the state transition to a significant change in the characteristic length of their geostatistical analysis.

In winter, late autumn and early spring more than one cluster of dispersion functions is dominating. Outside the vegetation

25 period we observed multiple cluster transitions. In many cases these cluster transitions coincided with a shift in rainfall regimes. Either the first stronger rainfall event after a longer period without rainfall sets in, or one of the heaviest rainfall events of that year occurs. There are also instances with recurring clusters that develop more than once (eg. Fig. 4, 5 a, 5 c, 6 e). As these periods are controlled by rainfall either different rainfall patterns or different hydrological processes are dominating. Depending on antecedent wetness, rainfall amounts and rainfall intensity, infiltration and subsurface flow processes can change and thus

30 also alter the soil moisture pattern. Although this may only be a coincidence, we found the green cluster in 2016 (Fig. 4) to form with strong rainfall input setting in after a period of little rainfall. Similar observations can be made for other years, unfortunately not in all cases. Consequently, we can neither confirm nor reject **H3** - that spatial dispersion is more pronounced during and shortly after rainfall driven wetting conditions, which implies a stronger spatial similarity of soil moisture.

During the vegetation period no spatial dependence is detectable. For the vegetation period, we found usually only one or at maximum two clusters (Table 2). These clusters are characterized by showing no dependence of dispersion on separating distance. That means, evapotranspiration forcing the system to drier states is doing this in a (spatially) homogeneous manner. Dispersion functions are not only flattened when the catchment is dry, they are already flat while the system is drying. Similar

5 observations have been reported for the Tarrawarra catchment in Australia (Grayson et al., 1997; Western and Grayson, 1998; Western et al., 1999). Although these works focused on the relation of spatial organization to topographic indices, no spatial correlation of soil moisture observations could be found for the dry period. This is comparable to our to our findings about dispersion functions during the vegetation period. It has to be noted that the lowest soil moisture values, i.e. residual moisture, are only observed for very short periods in time. At residual soil moisture all sensors show essentially the same absolute value

10 (which leads to small dispersion as well).

We conclude that cluster transitions were often triggered by rainfall events. But as not each strong rainfall event caused a cluster transition, further investigations on the rainfall events, their spatial characteristics and relation to the moisture state are needed.

4.4 Mean shift as a diagnostic tool

15 We used Mean shift mainly as a diagnostic tool to cluster dispersion functions based on their similarity. Similarity is measured by the Euclidean distance between two dispersion function vectors. This Euclidean distance does, however, not provide information on the underlying cause of dissimilarity and thus a simple shift along the y-axis can result in the same level of dissimilarity as a change in the shape of the dispersion function. We observed some cluster separations that were caused by minor differences in mean dispersion, but essentially describing the same spatial dependency.

20 It is possible to train better Mean shift algorithm instances. As described in the methods, we selected the bandwidth parameter for Mean shift, to yield meaningful results for the entire data set. The same parameter was used for all subsets to cluster dispersion functions on the same basis. This makes the clustering procedure itself comparable and thus, the number of identified clusters can support result interpretation. Nevertheless, it is likely that better bandwidth parameters can be found for each data subset individually, that overcomes misclassifications as described and discussed. Our objective, however, was to find clustering

25 results that can directly be compared to each other (instead of comparing hyper-parameters).

Dispersion functions operate in a higher dimensional space and might be affected by the curse of dimensionality. Mean shift clusters data points based on their distance to each other. Following the theory of the curse of dimensionality, with each added dimension (of these points), the difference of maximum and minimum distance between points become less significant (Beyer et al., 1999). On the one hand, we wish to resolve dispersion functions on as many distance lag classes as possible to gain more

30 insight on spatial dependencies. On the other hand, each additional lag class possibly decreases the performance of Mean shift (or any other clustering algorithm) and turns the results less meaningful. We calculated dispersion functions using a 30 days aggregation window and therefore end up with 335 points for Mean shift. However, despite the high number of points and the resulting uncertainty of cluster identification the clusters identified here seem plausible.

Mean shift is sensitive for the bandwidth parameter. As described in the methods (section 2.3), the bandwidth parameter has to be specified and has direct influence on the amount of clusters formed by the algorithm. We found a suitable parameter through trial-and-error. It would be more satisfactory to infer this crucial parameter from the data or supplementary information gathered at field campaigns. However, to our knowledge there is no such method or procedure to infer bandwidth parameters
5   for Mean shift from a data sample.

4.5   Limitations of the proposed method

Successful clustering does not point out spatial dependency. Mean shift can cluster functions without spatial dependency, as it uses their distance and no actual covariance between the functions. In this case the clustering is based on differences in mean, which may not even be statistically significant. The Mean shift algorithm is not meant to test clusters for statistical
10   independence. If two groups of points are separated or not depends only on the bandwidth parameter. Therefore the centroid functions of each cluster have to be checked for their shape and the information on spatial dependency that follows from that shape.

Our approach to find suitable bins to calculate the Shannon entropy is sensitive to outliers. We decided to rather define the width of a bin instead of their number. The reasons and necessity to do so were discussed in detail in section 4.2. As a width
15   we used the uncertainty propagated into dispersion function distances. From all distances that still were within uncertainty margins, we used the maximum value. In cases where this maximum distance is an outlier, it will influence the whole entropy calculations. This is a limitation to our method, but an acceptable one as it is still superior to other approaches from our point of view. Choosing the maximum distance within each year or depth (or both) will yield more bins for entropy calculations and therefore a wider range of values, but it would be very hard to compare these values.
20   From the point of view of the monitoring network, it has to be mentioned that the analysis of the 2013 data is likely to be less reliable, as during this period of installation the number of sensors was still lower than in the following years.

Due to the sampling design and the amount of observation points, we did not systematically test for differences of forest vs pasture plots, but ran our analyses across the two land covers. The fundamentally different shapes of cluster centroids in the summer clusters and, thus, the strong effect of vegetation altering soil moisture patterns might be partly more pronounced due
25   to the sampling design and not easy transferable to other sites. In our opinion, we would have made the same observations with a more stratified sampling design, as this is systematic catchment behavior, but we can neither confirm nor reject this.

5   Conclusions

We presented a new method to identify periods of similar spatial dispersion present in a data set. While soil moisture observations might be spatially heterogeneous, spatial patterns are much more persistent in time. We found two fundamentally different
30   states: On the one hand rainfall-driven cluster formations, usually characterized by strong relationships between dispersion and separating distance and a characteristic length, roughly matching the hillslope scale. On the other hand the vegetation period

cluster. A drying and then dry soil exhibits dispersion functions without spatial structure. Interestingly, these functions flatten out by minimizing the dispersion on large distance lags and we can thus see how the soil acts as a low pass filter.

While these long lasting periods of similar spatial patterns help us to understand how and when the soil is wetting or dying in an organized manner, there are possible applications beyond this. One could use the identification of clusters to stratify

5 data based on spatial dispersion for combined modeling. Then, for example, a set of spatio-temporal geostatistical models or hydrological models applied on each period separately, might in combination return reasonable catchment responses.

Our most interesting finding is that even a few soil moisture time series bear a considerable amount of predictive information about dynamic changes of soil moisture. We argue that distributed soil moisture sampling reflects an organized catchment state, where soil moisture variability is not random. Persistence spatial variability is hence not the dead end of distributed soil

10 moisture sampling because only a small amount of observation points is necessary to capture soil moisture dynamics.

Code and data availability. Major parts of the analysis are based on the scipy (Jones et al., 2001–), scikit-learn (Pedregosa et al., 2011) and scikit-gstat package (Mälicke and Schneider, 2019). All plots were generated using the matplotlib package (Hunter, 2007). The full analysis scripts are published on Github (https://github.com/mmaelicke/soil-moisture-dynamics-companion-code) (Mälicke, 2019).

The data is available upon request.

15 **Appendix A: Detailed result plots of 30 cm in 2014 and 2015**

Author contributions. The methodology was developed by MM, supervised by EZ and discussed with SH. The data was provided by TB and MW. All code was developed by MM. The manuscript was written by MM, with contributions of EZ in the introduction, discussion and the formulas. SH supplied the field and data descriptions. Structure, narrative and language of the manuscript was revised and significantly improved by TB.

20 *Competing interests.* The authors declare to have no competing interests.

Acknowledgements. We thank the German Ministerium für Wissenschaft, Forschung und Kunst, Baden-Württemberg for funding the V-FOR-WaTer project. We thank the German Research Foundation (DFG) for funding of the CAOS research unit FOR 1598. We especially acknowledge Britta Kattenstroth and Tobias Vetter, the technicians in charge of the maintenance of the monitoring network. The authors also acknowledge support by Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of Karlsruhe Institute of Technology

25 (KIT). The service charges for this open access publication have been covered by a Research Centre of the Helmholtz Association.
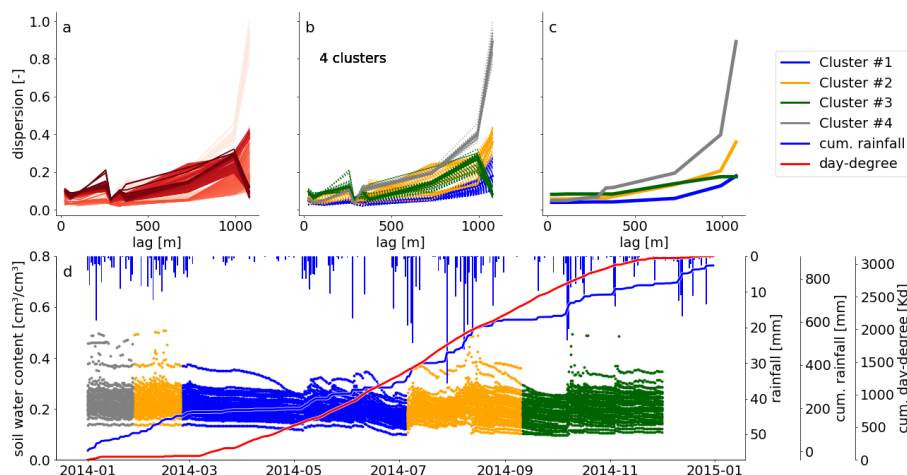
Figure A1. Spatial dispersion functions in 30 cm of 2014 based on on a window size of 30 days.

a): Spatial dispersion function for each position of the moving window. The red color saturation is indicating the window position. The darker the red the higher the Julian day.

b): The same dispersion functions as presented in a). Here the color indicates cluster membership as identified by the Mean shift algorithm.

c): Compressed spatial dispersion information represented by corrected cluster centroids. The colors match the clusters as presented in b).

d): Soil moisture time series of 2014 in 30 cm depth. The colors identify the cluster membership of the spatial dispersion function of the current window location and is matching the colors in b) and c). The bars on the top show the daily precipitation sums. The solid blue line is the cumulative daily precipitation sum and the red line the cumulative sum of all mean daily temperatures $> 5°$C.

References

Bárdossy, A. and Kundzewicz, Z. W.: Geostatistical methods for detection of outliers in groundwater quality spatial fields, Journal of Hydrology, 115, 343–359, https://doi.org/10.1016/0022-1694(90)90213-H, https://www.sciencedirect.com/science/article/pii/002216949090213H, 1990.

5   Bárdossy, A. and Lehmann, W.: Spatial distribution of soil moisture in a small catchment. Part 1: Geostatistical analysis, Journal of Hydrology, 206, 1–15, https://doi.org/10.1016/S0022-1694(97)00152-2, 1998.

Barlow, R. E., Bartholomew, D., Bremner, J., and Brunk, H.: Statistical Inference Under Order Restrictions: Theory and Application of Isotonic Regression, Tech. rep., MISSOURI UNIV COLUMBIA DEPT OF STATISTICS, 1972.

Beyer, K., Goldstein, J., Ramakrishnan, R., and Shaft, U.: When Is "Nearest Neighbor" Meaningful?, in: International conference on
10   database theory, pp. 217–235, Springer, Berlin, Heidelberg, https://doi.org/10.1007/3-540-49257-7_15, http://link.springer.com/10.1007/3-540-49257-7{_}15, 1999.

Blume, T., Zehe, E., and Bronstert, A.: Use of soil moisture dynamics and patterns at different spatio-temporal scales for the investigation of subsurface flow processes, Hydrology and Earth System Sciences, 13, 1215–1233, https://doi.org/10.5194/hess-13-1215-2009, 2009.

Bras, R. L.: Complexity and organization in hydrology: A personal view, https://doi.org/10.1002/2015WR016958, http://doi.wiley.com/10.
15   1002/2015WR016958, 2015.
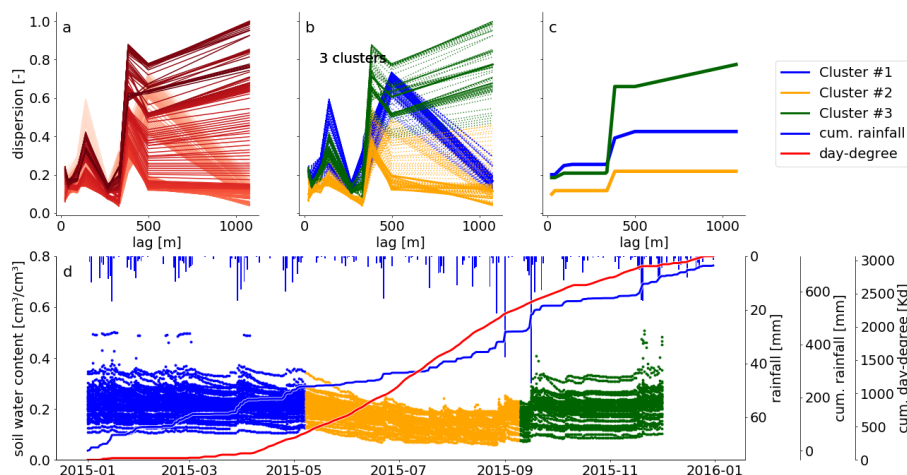
Figure A2. Spatial dispersion functions in 30 cm of 2015 based on on a window size of 30 days.

a): Spatial dispersion function for each position of the moving window. The red color saturation is indicating the window position. The darker the red the higher the Julian day.

b): The same dispersion functions as presented in a). Here the color indicates cluster membership as identified by the Mean shift algorithm.

c): Compressed spatial dispersion information represented by corrected cluster centroids. The colors match the clusters as presented in b).

d): Soil moisture time series of 2015 in 30 cm depth. The colors identify the cluster membership of the spatial dispersion function of the current window location and is matching the colors in b) and c). The bars on the top show the daily precipitation sums. The solid blue line is the cumulative daily precipitation sum and the red line the cumulative sum of all mean daily temperatures $> 5°$C.

Brocca, L., Morbidelli, R., Melone, F., and Moramarco, T.: Soil moisture spatial variability in experimental areas of central Italy, Journal of Hydrology, 333, 356–373, https://doi.org/10.1016/j.jhydrol.2006.09.004, https://www.sciencedirect.com/science/article/pii/S0022169406004781, 2007.

Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R.: Soil moisture temporal stability over experimental areas in Central Italy, Geo-
5   derma, 148, 364–374, https://doi.org/10.1016/j.geoderma.2008.11.004, 2009.

Brocca, L., Tullo, T., Melone, F., Moramarco, T., and Morbidelli, R.: Catchment scale soil moisture spatial-temporal variability, Journal of Hydrology, 422-423, 63–75, https://doi.org/10.1016/j.jhydrol.2011.12.039, https://www.sciencedirect.com/science/article/pii/S0022169411009437, 2012.

Bronstert, A., Creutzfeldt, B., Graeff, T., Hajnsek, I., Heistermann, M., Itzerott, S., Jagdhuber, T., Kneis, D., Lück, E., Reusser, D., and
10   Zehe, E.: Potentials and constraints of different types of soil moisture observations for flood simulations in headwater catchments, Natural Hazards, 60, 879–914, https://doi.org/10.1007/s11069-011-9874-9, http://link.springer.com/10.1007/s11069-011-9874-9, 2012.

Burgess, T. M. and Webster, R.: Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging, Journal of Soil and Science, 31, 315–331, 7 figs, 1 table, 27 refs, https://doi.org/10.1111/j.1365-2389.1980.tb02084.x, 1980.

Comaniciu, D. and Meer, P.: Mean shift: A robust approach toward feature space analysis, IEEE Transactions on pattern analysis and machine
15   intelligence, 24, 603–619, 2002.

Cressie, N. and Hawkins, D. M.: Robust estimation of the variogram: I, Journal of the International Association for Mathematical Geology, 12, 115–125, https://doi.org/10.1007/BF01035243, 1980.

De Cesare, L., Myers, D., and Posa, D.: FORTRAN programs for space-time modeling, Computers & Geosciences, 28, 205–212, 2002.

Dooge, J. C.: Looking for hydrologic laws, Water Resources Research, 22, 1986.

5 Entekhabi, D., Rodriguez-Iturbe, I., and Bras, R. L.: Variability in large-scale water balance with land surface-atmosphere interaction, Journal of Climate, 5, 798–813, 1992.

Fukunaga, K. and Hostetler, L.: The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Transactions on information theory, 21, 32–40, 1975.

Grayson, R. B., Western, A. W., Chiew, F. H. S., and Blöschl, G.: Preferred states in spatial soil moisture patterns: Local and nonlo-
10 cal controls, Water Resources Research, 33, 2897–2908, https://doi.org/10.1029/97WR02174, http://onlinelibrary.wiley.com/doi/10.1029/97WR02174/full, 1997.

Hinterding, A.: Entwicklung hybrider Interpolationsverfahren für den automatisierten Betrieb am Beispiel meteorologischer Größen, Ph.D. thesis, Institut für Geoinformatik, Universität Münster, Münster, Germany, 2003.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90–95,
15 https://doi.org/10.1109/MCSE.2007.55, 2007.

IUSS Working Group: WRB, 2006, World reference base for soil resources, 2006.

Jones, E., Oliphant, T., Peterson, P., et al.: SciPy: Open source scientific tools for Python, http://www.scipy.org/, [Online; accessed <today>], 2001–.

Jost, G., Heuvelink, G., and Papritz, A.: Analysing the space–time distribution of soil water storage of a forest ecosystem using spatio-
20 temporal kriging, Geoderma, 128, 258–273, 2005.

Kitanidis, P. K. and Vomvoris, E. G.: A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, Water Resources Research, 19, 677–690, https://doi.org/10.1029/WR019i003p00677, 1983.

Kleidon, A.: How does the Earth system generate and maintain thermodynamic disequilibrium and what does it imply for the future of the planet?, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370, 1012–1040, 2012.

25 Kondepudi, D. and Prigogine, I.: From heat engines to dissipative structures, in: Modern Thermodynamics, John Wiley & Sons Chichester, 1998.

Lark, R. M.: Towards soil geostatistics, Spatial Statistics, 1, 92–99, https://doi.org/10.1016/j.spasta.2012.02.001, 2012.

Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., and Zehe, E.: Picturing and modeling catchments by representative hillslopes, Hydrology and Earth System Sciences, 21, 1225–1249, https://doi.org/10.5194/hess-21-1225-2017, https:
30 //www.hydrol-earth-syst-sci.net/21/1225/2017/, 2017.

Loritz, R., Gupta, H., Jackisch, C., Westhoff, M., Kleidon, A., Ehret, U., and Zehe, E.: On the dynamic nature of hydrological similarity, Hydrology and Earth System Sciences Discussions, 2018, 1–38, 2018.

Ly, S., Charles, C., and Degré, A.: Geostatistical interpolation of daily rainfall at catchment scale: The use of several variogram models in the Ourthe and Amblève catchments, Belgium, Hydrology and Earth System Sciences, 15, 2259–2274, https://doi.org/10.5194/hess-15-
35 2259-2011, 2011.

Ma, C.: Spatio-temporal covariance functions generated by mixtures, Mathematical geology, 34, 965–975, 2002.

Ma, C.: Spatio-temporal stationary covariance models, Journal of Multivariate Analysis, 86, 97–107, 2003.

Martinez-Carreras, N., Krein, A., Gallart, F., Iffly, J.-F., Hissler, C., Pfister, L., Hoffmann, L., and Owens, P. N.: The influence of sediment sources and hydrologic events on the nutrient and metal content of fine-grained sediments (attert river basin, Luxembourg), Water, Air, & Soil Pollution, 223, 5685–5705, 2012.

McDonnell, J. J., Roderick, M. L., Selker, J., Vaché, K., Hinz, C., Hooper, R., Grant, G., Sivapalan, M., Kirchner, J., Weiler, M., Dunn, S., and Haggerty, R.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resources Research, 43, https://doi.org/10.1029/2006wr005467, http://doi.wiley.com/10.1029/2006WR005467, 2007.

Meyles, E., Williams, A., Ternan, L., and Dowd, J.: Runoff generation in relation to soil moisture patterns in a small Dartmoor catchment, Southwest England, Hydrological Processes, 17, 251–264, https://doi.org/10.1002/hyp.1122, http://doi.wiley.com/10.1002/hyp.1122, 2003.

Mälicke, M.: Companion Code for: Soil moisture: variable in space but redundant in time. (10.5194/hess-2019-574), https://doi.org/10.5281/zenodo.3532729, 2019.

Mälicke, M. and Schneider, H. D.: Scikit-GStat 0.2.6: A scipy flavoured geostatistical analysis toolbox written in Python., https://doi.org/10.5281/zenodo.3531816, 2019.

Patil, S. and Stieglitz, M.: Controls on hydrologic similarity: Role of nearby gauged catchments for prediction at an ungauged catchment, Hydrology and Earth System Sciences, 16, 551–562, https://doi.org/10.5194/hess-16-551-2012, www.hydrol-earth-syst-sci.net/16/551/2012/, 2012.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Pfister, L., Humbert, J., and Hoffmann, L.: Recent trends in rainfall-runoff characteristics in the Alzette river basin, Luxembourg, Climatic Change, 45, 323–337, 2000.

Pool, M., Carrera, J., Alcolea, A., and Bocanegra, E. M.: A comparison of deterministic and stochastic approaches for regional scale inverse modeling on the Mar del Plata aquifer, Journal of Hydrology, 531, 214–229, https://doi.org/10.1016/j.jhydrol.2015.09.064, 2015.

Rodríguez-Iturbe, I., Devoto, G., and Valdés, J. B.: Discharge response analysis and hydrologic similarity: The interrelation between the geomorphologic IUH and the storm characteristics, Water Resources Research, 15, 1435–1444, https://doi.org/10.1029/WR015i006p01435, http://doi.wiley.com/10.1029/WR015i006p01435, 1979.

Rosenbaum, U., Bogena, H. R., Herbst, M., Huisman, J. A., Peterson, T. J., Weuthen, A., Western, A. W., and Vereecken, H.: Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale, Water Resources Research, 48, https://doi.org/10.1029/2011WR011518, http://doi.wiley.com/10.1029/2011WR011518, 2012.

Sampson, P. D. and Guttorp, P.: Nonparametric Estimation of Nonstationary Spatial Covariance Structure, Journal of the American Statistical Association, 87, 108–119, https://doi.org/10.1080/01621459.1992.10475181, https://www.tandfonline.com/doi/abs/10.1080/01621459.1992.10475181, 1992.

Schume, H., Jost, G., and Katzensteiner, K.: Spatio-temporal analysis of the soil water content in a mixed Norway spruce (Picea abies (L.) Karst.)-European beech (Fagus sylvatica L.) stand, Geoderma, 112, 273–287, https://doi.org/10.1016/S0016-7061(02)00311-7, 2003.

Shannon, C. E.: A mathematical theory of communication, Bell system technical journal, 27, 379–423, 1948.

Sivapalan, M.: Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection?, Hydrological Processes, 17, 1037–1041, https://doi.org/10.1002/hyp.5109, http://doi.wiley.com/10.1002/hyp.5109, 2003.

Sivapalan, M., Yaeger, M. A., Harman, C. J., Xu, X., and Troch, P. A.: Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry, Water Resources Research, 47, https://doi.org/10.1029/2010WR009568, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010WR009568, 2011.

Snepvangers, J., Heuvelink, G., and Huisman, J.: Soil water content interpolation using spatio-temporal kriging with external drift, Geoderma, 112, 253–271, 2003.

Sprenger, M., Seeger, S., Blume, T., and Weiler, M.: Travel times in the vadose zone: Variability in space and time, Water Resources Research, 52, 5727–5754, 2016.

Starr, J., Timlin, D., Cady, R., and Nicholson, T.: Evaluating Uncertainties in Ground-Water Recharge Estimates Through Advanced Monitoring, in: AGU Spring Meeting Abstracts, 2002.

Takagi, K. and Lin, H. S.: Changing controls of soil moisture spatial organization in the Shale Hills Catchment, Geoderma, 173-174, 289–302, https://doi.org/10.1016/j.geoderma.2011.11.003, https://www.sciencedirect.com/science/article/pii/S0016706111003132{#}f0015, 2012.

Topp, G., Davis, J., and Annan, A.: Electromagnetic Determination of Soil Water Content Using TDR: I. Applications to Wetting Fronts and Steep Gradients 1, Soil Science Society of America Journal, 46, 672–678, 1982.

Topp, G., Zebchuk, W., Davis, J., and Bailey, W.: The measurement of soil water content using a portable TDR hand probe, Canadian Journal of Soil Science, 64, 313–321, 1984.

Vanderlinden, K., Vereecken, H., Hardelauf, H., Herbst, M., Martínez, G., Cosh, M. H., and Pachepsky, Y. A.: Temporal stability of soil water contents: A review of data and analyses, https://doi.org/10.2136/vzj2011.0178, https://www.soils.org/publications/vzj/abstracts/11/4/vzj2011.0178, 2012.

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, https://doi.org/10.1029/2008WR006829, http://doi.wiley.com/10.1029/2008WR006829, 2008.

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, Geography Compass, 1, 901–931, https://doi.org/10.1111/j.1749-8198.2007.00039.x, http://doi.wiley.com/10.1111/j.1749-8198.2007.00039.x, 2007.

Weijs, S. V., Van De Giesen, N., and Parlange, M. B.: Data compression to define information content of hydrological time series, Hydrology and Earth System Sciences, 17, 3171–3187, https://doi.org/10.5194/hess-17-3171-2013, www.hydrol-earth-syst-sci.net/17/3171/2013/, 2013.

Wendi, D. and Marwan, N.: Extended recurrence plot and quantification for noisy continuous dynamical systems, Chaos, 28, 085 722, https://doi.org/10.1063/1.5025485, http://aip.scitation.org/doi/10.1063/1.5025485, 2018.

Wendi, D., Marwan, N., and Merz, B.: In Search of Determinism-Sensitive Region to Avoid Artefacts in Recurrence Plots, International Journal of Bifurcation and Chaos, 28, 1850 007, https://doi.org/10.1142/s0218127418500074, https://www.worldscientific.com/doi/abs/10.1142/S0218127418500074, 2018.

Western, A. W. and Grayson, R. B.: The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements, Water Resources Research, 34, 2765–2768, 1998.

Western, A. W., Grayson, R. B., Blöschl, G., Willgoose, G. R., and McMahon, T. A.: Observed spatial organization of soil moisture and its relation to terrain indices, Water Resources Research, 35, 797–810, https://doi.org/10.1029/1998WR900065, http://doi.wiley.com/10.1029/1998WR900065, 1999.

Western, A. W., Zhou, S. L., Grayson, R. B., McMahon, T. A., Blöschl, G., and Wilson, D. J.: Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, Journal of Hydrology, 286, 113–134, https://doi.org/10.1016/j.jhydrol.2003.09.014, 2004.

Wu, W., Geller, M. A., and Dickinson, R. E.: The response of soil moisture to long-term variability of precipitation, Journal of Hydrometeo-
5   rology, 3, 604–613, 2002.

Zehe, E., Graeff, T., Morgner, M., Bauer, A., and Bronstert, A.: Plot and field scale soil moisture dynamics and subsurface wetness control on runoff generation in a headwater in the Ore Mountains, Hydrology and Earth System Sciences, 14, 873–889, https://doi.org/10.5194/hess-14-873-2010, 2010.

Zehe, E., Ehret, U., Pfister, L., Blume, T., Schroeder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K., and Others:
10   HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments, Hydrology and Earth System Sciences, 18, 4635–4655, 2014.