

This is the final draft of the contribution published as:

Schröter, I., Paasche, H., Dietrich, P., Wollschläger, U. (2015):
Estimation of catchment-scale soil moisture patterns based on terrain data and sparse TDR
measurements using a fuzzy c-means clustering approach
Vadose Zone J. **14** (11), 10.2136/vzj2015.01.0008

The publisher's version is available at:

<http://dx.doi.org/10.2136/vzj2015.01.0008>

1 **Estimation of catchment-scale soil moisture patterns**
2 **based on terrain data and sparse TDR measurements**
3 **using a fuzzy c-means clustering approach**

4
5 Ingmar Schröter^{*}, Hendrik Paasche, Peter Dietrich, Ute Wollschläger

6
7 I. Schröter, H. Paasche, P. Dietrich, U. Wollschläger, Dept. Monitoring and Exploration
8 Technologies, Helmholtz Centre for Environmental Research - UFZ, Permoserstraße 15, 04318
9 Leipzig, Germany; P. Dietrich, U. Wollschläger, WESS – Water and Earth System Sciences
10 Competence Cluster, Keplerstraße 17, 72074 Tübingen, Germany; P. Dietrich, Centre for
11 Applied Geoscience, University of Tübingen, Hölderlinstraße 12, 72074 Tübingen, Germany

12 ^{*}Corresponding author (ingmar.schroeter@ufz.de)

13
14
15 Contact details (^{*}corresponding author):

16 Ingmar Schröter

17 Department Monitoring and Exploration Technologies

18 Helmholtz Centre for Environmental Research – UFZ

19 Permoserstraße 15, 04318 Leipzig – Germany

20 ingmar.schroeter@ufz.de

21 Tel. +49 341 235 1754

22 Fax +49 341 234 51939

23 **Impact Statement**

24 We present an efficient method for sampling and spatial estimation of soil moisture at the small
25 catchment scale which is based on terrain data and sparse soil moisture measurements.

26 **Abstract**

27 Accurate characterization of spatial soil moisture patterns and their temporal dynamics is
28 important to infer hydrological fluxes and flow pathways and to improve the description and
29 prediction of hydrological models. Recent advances in ground-based and remote sensing
30 technologies provide new opportunities for temporal information on soil moisture patterns.
31 However, spatial monitoring of soil moisture at the small catchment scale (0.1-1 km²) remains
32 challenging and traditional *in situ* soil moisture measurements are still indispensable. This paper
33 presents a strategic soil moisture sampling framework for a low-mountain catchment. The
34 objectives were to: (i) find *a priori* a representative number of measurement locations, (ii)
35 estimate the soil moisture pattern on the measurement date and (iii) to assess the relative
36 importance of topography for explaining soil moisture pattern dynamics. The fuzzy c-means
37 sampling and estimation approach (FCM SEA) was used to identify representative measurement
38 locations for in-situ soil moisture measurements. The sampling was based on terrain attributes
39 derived from a DEM. Five TDR measurement campaigns were conducted from April to October
40 2013. The TDR measurements were used to calibrate the FCM SEA to estimate the soil moisture
41 pattern. For wet conditions the FCM SEA performed better than under intermediate conditions
42 and was able to reproduce a substantial part of the soil moisture pattern. A temporal stability
43 analysis shows a transition between states characterized by a re-organization of the soil moisture
44 pattern. This indicates that, at the investigated site, under wet conditions topography is a major

45 control that drives water redistribution whereas for the intermediate state other factors become
46 increasingly important.

47 **Keywords:** soil moisture, sampling design, cluster analysis, switching of states,
48 pattern analysis

49 **Introduction**

50 Despite the importance of soil moisture patterns to derive information for hydrological,
51 ecological, and pedological studies, spatial monitoring of soil moisture at the small catchment
52 scale (0.1-1 km²) remains a challenge. The interplay between static properties (e.g., topography,
53 soil, geology) and dynamic processes (e.g., vegetation growth, evapotranspiration) is the reason
54 why soil moisture is highly variable in space and time and the characterization of this variability
55 is one of the major challenges within the hydrological sciences (Vereecken et al., 2014).

56 To observe the spatio-temporal dynamics of soil moisture at the small catchment scale, numerous
57 measurement techniques are available including traditional *in situ* field measurements with
58 various types of soil moisture sensors, geophysical measurement techniques, and passive and
59 active microwave remote sensing (see e.g., Robinson et al., 2008; Wagner et al., 2007 for
60 reviews). A drawback of the application of geophysical techniques is that the temporal resolution
61 is often low and restricted to a few snapshots during the year and that the effectiveness of the
62 different methods to spatially map soil moisture often depends on site conditions (e.g., soil
63 texture, soil moisture state). Passive and active remote sensing data on airborne and space borne
64 platforms have been widely used for detecting soil moisture at different wavelengths. However,
65 at the small catchment scale remote sensing data are not yet operationally available to depict soil
66 moisture patterns at adequate temporal and spatial resolution (Bronstert et al., 2012). For this
67 reason, traditional *in situ* soil moisture measurements are still indispensable because they are

68 straightforward in application and provide most accurate data. However, there is a trade-off
69 between sampling density and spatial scale to observe soil moisture pattern dynamics by point
70 measurements. Recent advances in sensing technology, in particular through wireless sensor
71 networks (Cardell-Oliver et al., 2005; Bogaen et al., 2010), allow for automated soil moisture
72 monitoring in real time for the hillslope to the small catchment scale (Martini et al., 2015;
73 Rosenbaum et al., 2012; (Penna et al., 2009; Bogaen et al., 2010; Qu et al., 2015). However,
74 these measurements are costly and hard to maintain, especially at agriculturally used sites. This is
75 the reason why conventional soil moisture data collection is often done using portable sensors.
76 Several studies collected *in situ* soil moisture on uniform grids or densely distributed and
77 comprise hundreds of points for multiple sampling campaigns (Western and Grayson, 1998a;
78 Wilson et al., 2003; Takagi and Lin, 2012; Hu and Si, 2014). Grayson et al. (1997) and Western
79 et al. (1999a) demonstrated the use of “lots of points” (LOP) measurements (~500) to study the
80 dynamics of spatial soil moisture patterns and their controls at the 10 ha Tarrawarra catchment in
81 a temperate region in Australia. In their studies, they recognized that soil moisture patterns tend
82 to switch between two preferred states depending on the seasonal evapotranspiration and
83 precipitation ratio and the moisture state of the catchment. The observed patterns indicated a
84 different degree of spatial organization (i.e., more organized for wet than dry state) related to the
85 dominant hydrological process controlling the soil moisture distribution. The dry state is
86 dominated by vertical fluxes, with local controls including soil properties and local terrain (areas
87 of high convergence) shaping the spatial pattern. The wet state is dominated by lateral water
88 movement through both surface and subsurface paths, with nonlocal controls which represent the
89 dominant influence of catchment terrain on the distribution of soil moisture. Merz and Plate
90 (1997) showed that organization in spatial patterns of soil moisture at particular times and soil

91 properties may have a dominant influence on catchment runoff. Moreover, Western and Grayson
92 (2000) demonstrated the use of observed patterns to improve the development, calibration and
93 testing of a distributed model.

94 Systematic comprehensive observation campaigns like the ones of the Tarrawarra study (Western
95 et al., 1999b) or the Mahurangi catchment (Woods et al., 2001) are rare because they are labor
96 intensive and time consuming. Therefore, there is a need for new measurement designs which are
97 technically and economically feasible, while maximizing information content (Soulsby et al.,
98 2008). The challenging questions to design such strategic sampling schemes for *in situ* soil
99 moisture measurements include: (i) how to define *a priori* an appropriate number of
100 representative measurement locations?, and (ii) how to regionally estimate soil moisture patterns
101 (from point observations) for a target area by using sparse *in situ* measurements and readily
102 available ancillary data so that the derived spatial information matches the model resolution or
103 the footprint of high-resolution remote sensing data to which it is being compared?

104 Vachaud et al. (1985) were the first who introduced the concept of temporal stability (TS) in soil
105 moisture monitoring to reduce the measurement effort to characterize the spatial soil moisture
106 pattern of large fields. TS is described as the temporal persistence of a spatial pattern which
107 implies that particular locations exist in the field that always display mean behavior while others
108 are persistently wet or dry (Kachanoski and de Jong, 1988). Currently, the number of
109 publications on TS of soil moisture is growing quickly mainly in applying the TS concept to
110 select one or more locations out of a larger sampling volume to estimate average soil moisture for
111 the field to catchment scale (Grayson and Western, 1998; Martínez-Fernández and Ceballos,
112 2005; Martinez et al., 2008; Schneider et al., 2008; Zucco et al., 2014). However, such a
113 catchment average soil moisture monitoring (CASMM), introduced by Grayson and Western

114 (1998), does not aim to describe the spatial patterns (e.g., zones of saturation), relevant for the
115 movement of water, and to shed light on the reason for TS and its controls. Therefore, other
116 sampling techniques are needed that use available knowledge about structure forming processes
117 allocated to disciplines of pedology, biology, hydrology, and geomorphology, to expand sparse
118 measurements to a continuous representation of the soil moisture pattern (Schulz et al., 2006).
119 Recently, this approach has shaped the field of hydropedology (Lin, 2003) that uses pedometrics
120 to study the spatial patterns of soil properties and the related temporal dynamics of water in the
121 vadose zone. In this framework, various sampling techniques have been proposed and tested to
122 efficiently capture catchment/environmental conditions. These include stratified random sampling (SRS)
123 (McKenzie and Ryan, 1999; De Gruijter et al., 2006), response surface sampling (RSS) (Lesch, 2005) or
124 conditioned Latin hypercube sampling (cLHS) (Minasny and McBratney, 2006; Schmidt et al., 2014), to
125 name a few, to find sampling locations assisted and guided by the presence of ancillary data, such as
126 terrain attributes, geophysical measurements, remote sensing images or vegetation maps. The sampling
127 points are then chosen to optimize the soil property/ancillary data relationship. Furthermore, Werbylo and
128 Niemann (2014) tested the effectiveness of SRS and cLHS to select a limited number of points for soil
129 moisture monitoring based on topographic data and to calibrate two different models to estimate soil
130 moisture patterns at three catchments.

131 To find representative sampling locations for soil moisture measurements *a priori* knowledge of
132 key drivers of soil moisture patterns is essential. The general idea behind this is based on the soil-
133 landscape paradigm proposed by Jenny (1941) which was further generalized and formulated by
134 McBratney et al. (2003). In short, the spatial variability of a soil attribute is the result of spatially
135 referenced soil forming factors (environmental covariates) which can be used to establish soil
136 spatial prediction functions (McBratney et al., 2003). Considering soil moisture as a dynamic soil
137 attribute, a general prediction model can be described as

$$138 \quad \theta(x, y, z, t) = f(Q) \quad [1]$$

139 where Q is a set of p environmental variables (i.e., surrogate patterns) that provide information
140 about the underlying catchment characteristics. They can, for example, be derived from GIS,
141 proximal soil sensing, and remote sensing. $\theta(x, y, z, t)$ stands for the soil moisture at some spatial
142 location $x, y, (z)$ at time t . The general problem consists in the definition of zones (clusters) with
143 similar environmental characteristics which are assumed to show similar hydrological response
144 (i.e., soil moisture dynamics). In each zone, a soil moisture monitoring location and a set of
145 collocated environmental variables exist to build a function f which is flexible enough to
146 describe a nonlinear relationship. Based on this empirical quantitative function, spatial estimates
147 are made from observation data to infer soil moisture at unsampled locations. To follow this
148 approach a main task is the identification of representative zones and to derive the soil-landscape
149 relationships. With the advent of pattern recognition in the late 1960s the usage of the fuzzy
150 clustering technique has continuously found its way into geosciences. Since then, fuzzy
151 classification has been applied in many fields to extract knowledge based on ancillary data for
152 automated landform classification (Burrough et al., 2000; MacMillan et al., 2000) and soil
153 classification (Odeh et al., 1992; Triantafilis et al., 2013). However, the performance of the fuzzy
154 clustering also depends on the considered set of factors and is sensitive to the selected number of
155 clusters (Stevenson et al., 2015; Sun et al., 2012). Fuzzy classes have also been used to identify
156 sampling locations for mapping soil types (Odeh et al., 1990) and soil moisture (Van Arkel and
157 Kaleita, 2014). Furthermore, the study of Schmidt et al., (2014) where a fuzzy sampling scheme
158 was compared to other sampling schemes supports the usefulness of this method in combination
159 with nonlinear multiple regression (i.e. random forest) to predict soil properties at the field-scale.

160 The general advantage of the fuzzy classification method is that it allows for class overlap to
161 account for gradual transitions which often occur in the environment (Burrough et al., 2000).

162 The objectives of this study are to apply a strategic sampling design based on a fuzzy c-means
163 clustering technique (Paasche et al., 2006) to identify *a priori* a limited number of representative
164 sampling locations to monitor near-surface soil moisture dynamics at the small catchment scale,
165 (ii) to apply the sampling framework to spatially estimate soil moisture pattern on the
166 measurement date, and (iii) to assess the relative importance of topography for explaining soil
167 moisture pattern dynamics at the small catchment scale for the catchment under investigation.

168 **Theoretical background of the FCM sampling and estimation** 169 **approach (FCM SEA)**

170 *Fuzzy c-means clustering technique*

171 The fuzzy c-means (FCM) clustering technique can be applied to stratify a catchment, based on a
172 set of environmental covariates (i.e., proxies for soil moisture controlling factors), into a specified
173 number of clusters. In this paper, we call these clusters soil landscape descriptors (SLDs) where it
174 is expected that samples belonging to each SLD show similar soil moisture response. The
175 geospatial data of the catchment are arranged in a $n \times p$ matrix \mathbf{X} , with element x_{ik} , n is the
176 number of data points (pixels) and p is the number of predictor variables (e.g., environmental
177 variables such as terrain attributes) with $i = 1 \dots n$ and $k = 1 \dots p$. The optimal classification for a
178 selected number of c clusters is iteratively found so that the multivariate within-cluster variance
179 is as small as possible (Burrough et al., 2000). This is obtained by minimizing the following
180 objective function (Bezdek, 1974):

181 $J_{FCM} = \sum_{i=1}^n \sum_{j=1}^c m_{ij}^{\phi} (D_{ij})^2$, where [2]

182 m_{ij} denotes the degree of membership of a data point to a distinct cluster and D_{ij} is a selected

183 distance measure, defined here as Euclidean: $D_{ij}^2 = \|\mathbf{x}_i - \mathbf{v}_j\|^2$.

184 After selecting the number of clusters and minimizing the objective function (Eq. 2), the final

185 outcome of the FCM clustering provides a fuzzy membership matrix \mathbf{M} with elements m_{ij} and a

186 matrix of cluster centers \mathbf{V} with elements v_{jk} . In fuzzy logic, each element is assigned a partial

187 membership to all clusters based on their distance to the respective cluster center. The

188 membership values vary between zero and one; the the value the closer the element is to the

189 corresponding cluster center. The exponent ϕ determines the degree of fuzziness. For ϕ

190 approaching one, the algorithm resembles a crisp classification algorithm that only allows an

191 individual pixel/data point to lie in one mutually exclusive cluster, while for larger values it

192 allows an individual to be partial member of all clusters (i.e., the membership to a specific cluster

193 is more fuzzy). In our study, ϕ was set to 1.6, which is widely accepted as suitable choice in

194 literature (Bezdek et al., 1984). The membership m_{ij} of the i th object to the j th cluster is

195 determined by

196
$$m_{ij} = \frac{D_{ij}^{-2/(\phi-1)}}{\sum_{l=1}^c D_{il}^{-2/(\phi-1)}},$$
 [3]

197 where the sum of all membership values of an element across all clusters is unity. The cluster

198 center v_{jk} of the j th cluster for the k th attribute is calculated as

199
$$v_{jk} = \frac{\sum_{i=1}^n (m_{ij})^{\phi} x_{ik}}{\sum_{i=1}^n (m_{ij})^{\phi}}.$$
 [4]

200 One important issue in all clustering techniques is the choice of an appropriate number of clusters

201 which is inherently subjective. This is known as “cluster validity problem” which strives to offer

202 a quantitative (statistical) measure which indicates how well the algorithm has identified the
 203 structure that is present in the underlying ancillary data. Therefore, many authors have proposed
 204 validity functionals in order to solve the validity problem. The fuzzy performance index (FPI)
 205 and the normalized classification entropy (NCE) are two of these functionals which were found
 206 to be most useful (Roubens, 1982) and are widely used in literature (Burrough et al., 2000;
 207 Triantafilis et al., 2003). The FPI is a measure of the degree of fuzziness and the NCE indicates
 208 the degree of disorganization in the classification. The least fuzzy and disorganized number of
 209 clusters is considered optimal (Odeh et al., 1990). After finding the optimal number of clusters,
 210 defuzzification techniques (Leekwijck and Kerre, 1999) are used to define crisp classes in order
 211 to perform a geoscientific interpretation of every class. The general workflow of this approach
 212 has been used by others (e.g., Paasche and Eberle, 2009) and is displayed in Figure 1.

213 Our proposed method is based on the work of Paasche et al. (2006) but works in a modified way
 214 (Fig. 2) because the purpose of our sampling strategy is to identify locations that represent the
 215 range of values of each ancillary variable so that the sampling is less likely to be redundant.
 216 Therefore, we consider the FCM clustering as imaging technique using the membership matrix \mathbf{M}
 217 and the center matrix \mathbf{V} to store the spatial heterogeneity and average cluster-specific
 218 information, respectively. These elements are of central importance because they describe the
 219 complex landscape structure in a reduced form. Using a simple mixing law, the ancillary data can
 220 be reconstructed from the membership and center matrices (Fig. 2).

221 The reconstructed maps are stored in a matrix \mathbf{B} and the reconstructed value of the i th data point
 222 for the k th attribute is calculated as a weighted sum over all clusters:

$$223 \quad b_{ik} = \sum_{j=1}^c m_{ij} v_{jk}. \quad [5]$$

224 Comparing the reconstructed and the attribute data allows the quantification of informational loss
225 during the clustering procedure. To account for this the total absolute difference (TAD) was
226 calculated:

$$227 \quad TAD = \sum_{i=1}^n |\mathbf{b}_i - \mathbf{x}_i| \quad [6]$$

228 where \mathbf{b}_i and \mathbf{x}_i are reconstructed and attribute data points.

229 A simple L-curve analysis (e.g., Twarakavi et al., 2010) of informational loss over different
230 numbers of clusters allows for the identification of an optimal number of clusters. L-curve
231 analysis is a well-established technique to identify the optimal regularization strength for ill-
232 posed discrete optimization techniques (e.g., Lassonde, 2001) which constrains the
233 spatial/structural complexity of the final solution. In our case, the number of clusters constrains
234 the complexity of the solution of the cluster analysis and thus the ability to capture the spatial
235 heterogeneity of all considered ancillary data in the fuzzy membership matrix. The selection of an
236 optimal number of clusters ensures that sufficient information is stored in the membership matrix
237 \mathbf{M} to reconstruct the complex patterns/structures of the ancillary data. In case the “elbow” point
238 cannot be detected uniquely by visual inspection it is advisable to better increase the number of
239 clusters towards a very conservative optimal number of clusters to avoid the risk of losing
240 structural information and we have followed this strategy in this study. Figure 3 conceptually
241 shows the TAD relationship for different cluster solutions and the “elbow” point where the
242 information loss saturates. The fuzzy membership matrix of this optimal solution (c_{opt}) is then
243 further analyzed for the identification of suitable sampling locations (e.g., Hachmöller and
244 Paasche, 2013). Note, for the case of noise-free ancillary data, one could rank the sampling
245 locations according to the values of maximal membership to a cluster. For example, the point,
246 where maximal membership to a cluster is achieved would be the most appropriate one for soil

247 moisture sampling. However, the ancillary data are not free of noise components which are not
248 specifically known for every data point. According to the principles of error propagation, this
249 noise will propagate into the membership values not allowing for a strict ranking of optimal
250 sampling locations. Thus, we are only searching for a point with high membership value to a
251 cluster but not for the point with highest membership. Practically, this makes our approach very
252 applicable to regions of rough and hardly or inaccessible terrain, since the condition to sample a
253 very specific point in the study region is not restricted to a single sampling location but rather to a
254 subset of points. Therefore, the memberships of c_{opt} are used to identify areas within the zones
255 which have at least a membership of 0.8 to a corresponding class. For each class a minimum of
256 one location is selected to form a sparse set of sampling points.

257 *Estimation of soil moisture patterns*

258 Following Paasche et al. (2006) and (Hachmöller and Paasche, 2013) we use sparse
259 measurements and the fuzzy membership matrix describing the spatial heterogeneity of the
260 attributes to estimate soil moisture spatial patterns . This is done by calibrating each cluster with
261 a point soil moisture measurement thus extending the center matrix \mathbf{V} achieved by clustering the
262 ancillary/physiographic data sets. This extension is done by adding a column to the matrix
263 containing the measured soil moisture value for each cluster obtained from the sampling locations
264 derived with the FCM clustering. Taking the fuzzy membership matrix as spatial weighting
265 information (equation 5) for the soil moisture information stored in the extended center matrix \mathbf{A}
266 we achieve a spatially continuous prediction of soil moisture distribution (Fig.2). This approach
267 holds if the variable of interest is related to at least one of the underlying physiographic attributes
268 influencing the pattern in the fuzzy membership matrix. Prediction is done without knowing or
269 specifying the exact relation of the target value (i.e., soil moisture) to any of the underlying

270 ancillary data. However, since only structural heterogeneity described by the membership values
271 is used to guide the prediction of the spatial distribution of the target value, this approach will
272 likely fail, if the ancillary database does not reflect the various dependencies of the target value.

273 **Site Description**

274 The Schäfertal research site (11°03'E, 51°39'N) is a small low-mountain catchment and is part of
275 the long-term Earth observation network TERENO (Zacharias et al., 2011). It is located on a
276 plain in the Lower Harz Mountains in central Germany, approximately 150 km southwest of
277 Berlin (Fig. 4). The Schäfertal catchment has a humid continental climate and the mean annual
278 air temperature is 6.8 °C ranging from -1.8 °C in January to 15.5 °C in July (Ollesch et al., 2006).
279 Mean annual precipitation is about 630 mm (time series 1968-2006) which is low compared to
280 other low mountain-areas in Germany due to the leeward position of the region on the eastern
281 slope of the Harz Mountains (Reinstorf, 2010). From April to September, the climatic water
282 balance yield according to the closest weather station, 9 km away, in Harzgerode, is negative (-64
283 mm) whereas the long-term mean of the annual climatic water balance is +126 mm (Abdank et
284 al., 1995).

285 The size of the agriculturally used catchment is 144 ha. Its elevation ranges from 393 m a.s.l.
286 from the outlet of the catchment to 445 m a.s.l. on the highest ridge. The catchment is V-shaped
287 with a first order stream in the valley and with gentle to moderate slopes (up to 20%) on both
288 sides of the stream. The site is characterized by four distinct landforms: 1) north-facing slope, 2)
289 south-facing slope, both intensively used for agriculture, 3) valley bottom with pasture or
290 meadow and 4) topographic depressional areas (swales) disrupting the slopes on both sides of the

291 stream. In autumn 2012, parts of the grassland area in the western part of the catchment were
292 transformed into arable land.

293 The catchment is underlain by Devonian greywacke and shale which are covered by a complex of
294 periglacial layers with different fractions of silt and rock fragments (Altermann, 1985). Different
295 soils evolved according to the sequence of the cover layer and landscape position. The dominant
296 soils comprise Luvisols and Cambisols on the hillslopes and peaty Gleysols in the valley bottom
297 (Borchardt, 1982).

298 **Material and Methods**

299 *Selection of ancillary data*

300 To capture the spatio-temporal variability of soil moisture in the Schäfertal catchment the
301 sampling needs to be driven based on the digital implementation of external drivers of soil
302 formation that control processes of water redistribution. Therefore, the selection of appropriate
303 ancillary data requires understanding of the physical significance of patterns of specific
304 environmental variables to provide information which is in some way related to soil moisture
305 (Grayson et al., 2002). We follow a very simple approach and use topographic information for
306 this purpose. Recalling the soil-landscape paradigm (Hudson, 1992), topography at the small
307 catchment scale is an integral factor that has a lasting influence on the effects of gravitation,
308 water, biota, microclimate and soil formation and is therefore one of the most widely used static
309 factors that affects runoff processes (Western et al., 1999a; Beaudette et al., 2013). Of course
310 there is no reason per se to expect that topography solely explains soil moisture variability.
311 Previous studies show that in some settings the use of terrain indices perform well, while in other
312 settings they have been shown to perform poorly (Western et al., 2004). So we are well aware

313 that topography is not the only factor controlling soil moisture pattern dynamics but we can test
314 how much of the soil moisture pattern can be described in our research catchment by following
315 this rather simple approach. To this end, we used a digital elevation model (DEM) to derive
316 terrain attributes that may be related to key hydrological processes controlling the spatial
317 distribution of soil moisture in the Schäfertal catchment.

318 *Digital Terrain Modelling*

319 Terrain information was obtained from a high-resolution $1 \times 1 \text{ m}^2$ digital elevation model
320 (DEM1) derived from an airborne laser scanning (LIDAR) of the Schäfertal catchment
321 (GeoBasis-DE / LVermGeo LSA, 2009). Such a high resolution DEM was used to describe the
322 spatial arrangement of topographic structures (e.g., swales) in detail as they are relevant for soil
323 moisture redistribution and to approximate the sampling support for soil moisture measurements.
324 To reduce the amount of noise in the LIDAR data and to better represent primary topographic
325 attributes within the catchment, a $10 \times 10 \text{ m}^2$ filter window was applied to the original DEM
326 using ArcGIS 10.1 (ESRI, Redland, CA). By choosing this filter window, we calculated the mean
327 elevation of all cells within the window and applied the mean value to the corresponding center
328 cell. For practical reasons, the smoothed DEM was additionally resampled into a $2 \times 2 \text{ m}^2$ DEM
329 (DEM2) with bilinear interpolation. For hydro-pedological applications, a wide range of terrain
330 attributes are available that describe relevant features in the catchment with respect to soil
331 moisture redistribution (Moore et al., 1991; Behrens et al., 2010) and we assume that they are
332 useful for the identification of representative monitoring locations. We derived four topographic
333 attributes: slope, elevation, total annual incoming solar radiation (TIR) and SAGA wetness index
334 (SWI). Each attribute independently contributes information about local and contextual landscape
335 conditions and is commonly used in the literature (Western et al., 1999a; Wilson et al., 2005;

336 Takagi and Lin, 2012). By this means, the elevation was used to describe landscape position and
337 the gravitational potential energy that drives water flow. The slope is indicative to represent the
338 hydraulic gradient that drives surface and near-subsurface fluxes (Western et al., 1999a). The TIR
339 was used as an index for evapotranspiration and microclimate and the SWI to represent zones of
340 surface saturation. Note that we used SWI instead of topographic wetness index (TWI) as
341 introduced by Beven and Kirkby (1979). The SWI is a modified version of the TWI to account
342 for a more realistic prediction for cells situated in valley floors with small vertical distance to a
343 channel (Böhner and Selige, 2006). All terrain attributes were calculated with SAGA-GIS
344 (System for Automated Geoscientific Analyses) (see Wilson and Gallant (2000) for algorithms).
345 These computations yielded gridded data sets of 4 attributes (Fig. 5a-d) with a spatial resolution
346 of 2 m and form the input data for the fuzzy c-means cluster algorithm.

347 *Selection of Sampling Locations for Estimation of Soil Moisture Patterns*

348 The FCM clustering was conducted separately for the arable land and for the grassland. We
349 derived 20 clusters for arable land and 10 clusters for grassland, respectively. Figure 6 shows the
350 distribution of clusters with the clusters for both land uses being merged in one map. Taken
351 together, they form the 30 SLDs for the Schäferfetal catchment. Then, each cluster was assigned
352 with at least one “sparse” measurement point. In this study, we determined altogether 50
353 measurement points for TDR by placing two measurement points in each cluster on the arable
354 land (hillslope areas) and one in each grassland cluster (riparian zone). The two measurement
355 locations for the arable land were chosen as the total area is much larger than that of grassland
356 and this way to account for within cluster variation and to avoid local extrema which would
357 influence the interpolation. For a catchment area of $\sim 1.5 \text{ km}^2$ a number of 50 measurement points
358 is “sparse” in the sense that this number would be way too small for an interpolation of soil

359 moisture maps based on geostatistical approaches such as kriging. Nevertheless, conducting the
360 analysis with 30 measurement points (one per cluster) would also be feasible.

361 For validation of the soil moisture maps, we used an independent set of 44 TDR measurement
362 points distributed all over the catchment. Thirty of these locations were obtained from simple
363 random sampling across the whole catchment and 14 points were available from a Latin
364 hypercube sampling which was available from an already running time series of measurement
365 campaigns and was designed to provide local ground truth (soil moisture and vegetation
366 parameters) for remote sensing measurements. However, all these measurement points were
367 independent from the locations determined for the FCM sampling.

368 In combination with the calibration points a total of 94 monitoring locations were determined for
369 the catchment (Fig. 7), which lead to a measurement effort of a half day work. Each sampling
370 point was georeferenced with a Leica GPS1200 (Leica, Heerbrugg, Switzerland) with a lateral
371 and vertical resolution of 0.1 m, and the locations were marked. The near-surface (0-10 cm)
372 volumetric soil moisture was sampled using a portable TDR100 time domain reflectometer
373 (TDR) (Campbell Scientific, Logan, UT) with a custom-made three-rod probe, with a basic
374 accuracy of $\pm 0.02 \text{ m}^3/\text{m}^3$. Before measuring, TDR probes were calibrated for soil moisture
375 estimation with measurements in water and air. Volumetric soil moisture $\theta \text{ [m}^3/\text{m}^3]$ was
376 calculated using the CRIM (Complex Refractive Index Model) formula according to Roth et al.
377 (1990). Furthermore, soil temperature was measured with a DT-300 handheld LCD-thermometer
378 (Votcraft, Hirschau, Germany), with a basic accuracy of $\pm 1 \text{ }^\circ\text{C}$, in 3-6 cm depth. The
379 temperature data were used for temperature correction of the dielectric permittivity of water ϵ_w .
380 The dielectric permittivity of the soil matrix ϵ_s was set to an estimated value of 4.6 [-]. The
381 porosity of the soil was estimated based on soil information data (Borchardt, 1982) and was set to

382 0.4 for Cambisols and Luvisols on the slopes, and to 0.6 and 0.75, respectively, for the Gleysols
 383 in the lower relief positions near the channel and for the peat soils. At each point, campaign-
 384 based TDR measurements were conducted to acquire local soil moisture contents based on the
 385 average of three replicate TDR measurements.

386 *Soil Moisture Measurement Campaigns and Timing*

387 From spring to autumn 2013 five measurements campaigns (T1 to T5) were conducted in the
 388 Schäfertal catchment. The timing of the five field campaigns was designed to capture a range of
 389 soil moisture states. The moisture contents covered a wet (April-May 2013) moisture state due to
 390 the late snow melt in 2013 and an intermediate (September-October 2013) soil moisture state
 391 typical for fall conditions in the Lower Harz Mountains. A dry state could not be captured for the
 392 entire catchment due to limited access to the cropped fields during the summer.

393 All points were measured within a few hours (< 6h) to minimize the effect of evapotranspiration
 394 and drainage processes on the soil moisture measurements. The measurement campaigns were
 395 further optimized by calculating the shortest route for visiting each point with the Concorde
 396 Travelling Salesman (TSP) Solver (Applegate et al., 2001, see
 397 www.math.uwaterloo.ca/tsp/concorde/).

398 The temporal persistence of soil moisture patterns was tested by calculating the non-parametric
 399 Spearman rank correlation coefficient r_s for the various sampling dates (Vachaud et al., 1985).

400 The Spearman rank correlation coefficient for soil moisture values measured at observation times
 401 u_1 and u_2 is computed as

$$402 \quad r_s = 1 - \frac{6 \sum_{i=1}^n (R_{i,u_1} - R_{i,u_2})^2}{n(n^2 - 1)} \quad [7]$$

403 where n is the number of point soil moisture observations in the catchment, R_{ij} is the rank of θ_{ij}
 404 at location i at observation time u . The closer r_s is to unity, the more temporally stable are the
 405 patterns. In this respect, our intention to apply this coefficient is to check for temporal changes in
 406 soil moisture spatial organization which would not be explainable by the static topographic data
 407 used in our FCM clustering procedure.

408 ***FCM validation and estimation error***

409 To assess the prediction accuracy of this nonlinear estimation technique we calculated the Nash–
 410 Sutcliffe coefficient of efficiency (NSE) (Nash and Sutcliffe, 1970) and the root mean square
 411 error (RMSE). Therefore, the independent validation data set composed of 44 soil moisture
 412 observation points over the whole catchment was used. The NSE was used to explain how well
 413 the model matches the observed soil moisture pattern. The RMSE indicates the accuracy of the
 414 model to match the observed soil moisture. The NSE is defined as

$$415 \quad NSE = 1 - \frac{\sum_{t=1}^a (\theta_t - \hat{\theta}_t)^2}{\sum_{t=1}^a (\theta_t - \bar{\theta})^2} \quad [8]$$

416 where a is the number of data points in the validation data set, θ_t is the t th observed soil moisture,
 417 $\hat{\theta}_t$ is the predicted soil moisture of the t th observation and $\bar{\theta}$ is the average of the observations.

418 The RMSE is determined by

$$419 \quad RMSE = \left(\frac{1}{a} \sum_{t=1}^a (\theta_t - \hat{\theta}_t)^2 \right)^{1/2}. \quad [9]$$

420 The statistic software R (R Development Core Team, 2012) and the R package e1071 version
 421 1.6-1 (Meyer et al., 2014) was used to carry out the FCM clustering and to perform all statistical
 422 analyses.

423 **Results**

424 *Soil moisture measurements*

425 The main statistics of the five selected TDR sampling campaigns are provided in Table 1. The
426 first TDR campaign was conducted on April 17, 2013 right after the snowmelt, and was followed
427 by a second campaign on April 23, 2013 after one week of drying, and a third one on May 8,
428 2013 three weeks later with a small amount of rain between the second and the third campaign.
429 The fourth campaign was conducted on September 25, 2013 followed by the fifth measurement
430 on October 2, 2013 after one week of drying. T1 to T3 represent a wet moisture state and show
431 high spatial mean soil moisture values (0.34, 0.26, and 0.27 m³/m³, respectively) throughout the
432 catchment whereas T4 and T5 represent an intermediate state with moderate spatial mean soil
433 moisture values (0.19 and 0.18 m³/m³, respectively). The spatial patterns of the five TDR
434 measurement campaigns are displayed in Fig. 8.

435 All five measurement dates show a valley-dependent pattern of soil moisture with higher values
436 occurring in the valley bottom. However, the pattern for the wet moisture states (T1 to T3) is
437 more pronounced with higher ranges between 0.49 to 0.55 m³/m³ and higher standard deviations
438 from 0.09 to 0.11 m³/m³ for soil moisture contents (Tab. 1). The very high moisture contents in
439 the valley bottom (> 0.5 m³/m³) highlight locations with a peaty soil layer and high porosities at
440 the top of the soil profile that are in strong contrast to the soil moisture of the mineral soils with
441 lower porosities which, however are also at or close to saturation right after the snowmelt (0.3 to
442 0.4 m³/m³) (Fig. 8a-7c). Moreover, soil moisture values are higher on the north-exposed slope
443 than on the south-exposed slope which cannot be explained by topography only and which is also
444 contradictory to what would be expected from atmospheric forcing. The pattern for the

445 intermediate state (8d-8e) is less prominent with smaller soil moisture ranges between 0.36 and
446 $0.37 \text{ m}^3/\text{m}^3$ (Tab. 1) than the one observed during the wet state that displays strong alignment
447 along the valley and converging hillslopes.

448 *Estimated soil moisture patterns*

449 Figure 9 shows the soil moisture patterns for all five sampling campaigns estimated for the entire
450 catchment area with the FCM clustering approach using 50 TDR measurement points for
451 calibration based on the topography-based sampling scheme (Fig. 2).

452 All estimated patterns show the topographic dependence of soil moisture with increasingly wet
453 areas in the depression lines (swales) and in the valley bottom. For the measurement dates T1 to
454 T3 (Fig. 9a–9c) they also indicate the difference in soil moisture between the northern and the
455 southern hillslope with higher soil moisture contents predicted for the northern slope. To test how
456 well topography is suited to reproduce the observed soil moisture patterns, the prediction
457 accuracy was estimated using the Nash-Sutcliffe Coefficient of Efficiency (NSE) based on the
458 measurements at the 44 validation points. The predicted and observed soil moisture values for the
459 five measurement dates are displayed in Figure 10. In addition to that, Table 3 shows the
460 performance when the validation is separately conducted for arable land and grassland (riparian
461 zone) and for the whole catchment. For the measurement dates T1 to T3 during the wet state, the
462 FCM SEA performs well for the whole catchment and shows high NSE values (T1: 0.78; T2:
463 0.73; T3: 0.59) (Fig. 10). In contrast, the FCM performance is weaker during intermediate soil
464 moisture states on dates T4 and T5 (T4: 0.34; T5: 0.41). The RMSE for the whole catchment
465 yielded for all five sampling dates similar results with moderate accuracy ($0.05\text{--}0.06 \text{ m}^3/\text{m}^3$).
466 However, there is a strong contrast in the performance between the two land use types. For arable
467 land the estimates are almost always more accurate than for grassland with much smaller RMSE

468 (Table 3). For the wet state, T1 to T3, arable land outperforms grassland and shows more
 469 consistent NSE values than grassland. For the intermediate state, T4 to T5, the performance for
 470 both land use types decreases and shows negative NSE values for grassland and a negative and a
 471 positive NSE value for arable land. Note that negative NSE values indicate that the observed
 472 average soil moisture is a better estimate of the observed pattern than the estimated pattern
 473 obtained with the FCM SEA.

474 *Temporal stability of soil moisture patterns*

475 Table 2 provides an overview of the Spearman rank correlation coefficients calculated for all
 476 measurement dates. For the wet soil moisture state (T1 to T3) the spatial patterns show a high
 477 rank correlation ($r_s \geq 0.87$) indicating that the patterns are very similar during this time.
 478 Comparison of the moisture patterns of the wet state (T1 to T3) with the intermediate state (T4
 479 and T5) results in significantly lower rank correlations ($r_s \leq 0.66$), while the moisture patterns
 480 during the intermediate states exhibit again a higher rank correlation ($r_s = 0.79$). This analysis
 481 shows that the soil moisture patterns in the Schäfertal reorganize from the wet to the intermediate
 482 state.

483 **Discussion**

484 *Controls of spatio-temporal organization of soil moisture for the wet and intermediate states*

485 For the Schäfertal catchment, the performance of the FCM SEA decreases as the catchment
 486 becomes drier (Fig. 10). Moreover, there is significant shift in NSE values between the wet and
 487 the intermediate soil moisture conditions, indicating that the relative importance of topography
 488 diminishes and the relative importance of factors such as soil heterogeneity (i.e., texture,
 489 structure) and vegetation (i.e., crop type, land use, density) become increasingly important and

490 drive soil moisture variation. Additionally, we attribute the shift to the fact that
491 evapotranspiration is the controlling factor fostering vertical flow processes within the soil profile
492 during the summer and fall months. This would also support the work of Albertson and Montaldo
493 (2003) who provided a theoretical framework that demonstrated that vegetation can reduce soil
494 moisture spatial variability as soon as a positive covariance between transpiration and the soil
495 moisture field emerges. Since the estimated spatial pattern described from the fuzzy membership
496 information relies on topographical attributes only we are not able to reproduce this pattern re-
497 organization.

498 For the wet state, the FCM SEA based on a combination of single terrain attributes already
499 explained between 59% and 78% of soil moisture variability. The predicted patterns of soil
500 moisture (Fig. 9) reflect the terrain features well showing high soil moisture in the riparian zone
501 (valley) and converging areas (swales/hollows). Our findings correspond well to the hydrotope
502 map of the Schäfertal (Borchardt, 1982) that indicates a shallow groundwater table close to the
503 surface in the central part of riparian zone and in the converging areas. Additionally, field
504 observation shows failed sprouting due to waterlogging with the beginning of the growing season
505 in the converging areas. For this reason, there is evidence that the estimated maps produce a
506 realistic pattern under wet conditions and thus topography is an important control for soil
507 moisture patterns during these times in the Schäfertal catchment.

508 For dryer states our proposed terrain-driven sampling and estimation approach performs less
509 accurate but is still capable of explaining between 34% and 41% of total variance. This is still
510 relatively good compared to most other studies which rarely explained more than 50% of soil
511 moisture variability using topographic data (Western et al., 1999a; Takagi and Lin, 2012; Wilson
512 et al., 2005; Beaudette et al., 2013). However, for both moisture states the accuracy of the

513 estimated soil moisture is moderate with an average RMSE of $0.06 \text{ m}^3/\text{m}^3$. Our proposed FCM
514 SEA showed seasonality in the prediction accuracy that is in line with the studies that used terrain
515 indices to explain soil moisture variability. However, it should be noted that at our site, the FCM
516 SEA performance is better and more accurate (small RMSE) for arable land than for the
517 grassland areas in the valley bottom (Fig. 10, Table 3). We presume that this is due to the fact
518 that topography is less prominent in the riparian zone and expect other factors such as
519 groundwater influence and soil properties to be more important. Moreover, these areas show also
520 small-scale variability in soil moisture due to microtopography which is caused by hummocks
521 and depressions of the grassland patches which is not captured by the DEM.

522 Our temporal stability analysis on the observed soil moisture patterns provides an additional
523 explanation for the seasonality in the prediction accuracy of the FCM SEA. The analysis shows
524 that temporal re-organization of the soil moisture spatial pattern occurs during the year in the
525 Schäfertal catchment. There is at least a transition from a highly organized pattern during wet
526 conditions towards a more uniform distribution under intermediate conditions. Previous studies
527 that examined seasonal changes in near-surface soil moisture spatial organization attributed the
528 reorganization of soil moisture patterns to variable rates of evapotranspiration and root water
529 uptake by plants (Hupet and Vanclooster, 2002; Baroni et al., 2013), a change from
530 predominantly lateral soil water movement to predominantly vertical soil water movement
531 (Grayson et al., 1997; Western et al., 1999a), or to simply soil textural differences (Famiglietti et
532 al., 1998).

533 The results have implications for our FCM SEA and demonstrate that the actual SLDs (Fig. 6), as
534 expected, are not suitable to explain temporal dynamics for the entire range of soil moisture. This
535 agrees with the findings of Takagi and Lin (2012) who stated that a soil-landform unit (SLU) is

536 not a reasonable indicator of soil moisture spatial organization under dry conditions while
537 different SLUs can show the same moisture content. It is clear from our analysis and others
538 (Wilson et al., 2004) that to make a good state-space prediction of soil moisture, we need a better
539 *a priori* stratification of the catchment from SLUs towards meaningful hydrological response
540 units (HRUs). There is no doubt that other proxy data related to processes that control the spatial
541 distribution of soil moisture should be included in the FCM SEA. In particular the integration of
542 (static) soil and (dynamic) vegetation properties is an important step while they strongly affect
543 soil moisture variation (Baggaley et al., 2009; Hu and Si, 2014). In addition, dynamic soil
544 moisture patterns obtained from a time-lapse sequence of SAR data and cosmic-ray probes
545 (Zreda et al., 2008) as an emerging technology may also provide useful information to further
546 constrain the locations of HRUs in our research catchment.

547 Nevertheless, our proposed FCM SEA is promising for collecting data in small catchments in an
548 efficient way to characterize and analyze temporal dynamics of soil moisture. It provides a
549 synergistic integration of various proxy data which can be related to soil moisture. Based on a set
550 of surrogated patterns, hydrologically relevant structures were explored with the FCM clustering
551 technique which is a common technique in pattern recognition science. The explored patterns
552 represent SLDs, and in a best case HRUs, which are described by fuzzy membership maps. In
553 each SLD at least one representative sampling point is selected, and taken together they allow in
554 combination with the fuzzy membership maps a realistic estimation of the actual soil moisture
555 pattern. The main advantage of the FCM SEA is the small number of measurement points that
556 form the basis to characterize and predict soil moisture patterns. Earlier studies that made
557 measurements on regular grids like in the Tarrawarra catchment (10.5 ha, ~500 points) or that
558 used common geostatistical techniques (such as kriging) like in the Shale Hills catchment (7.7 ha,

559 ~189 points) need hundreds of points to obtain estimates of the spatial distribution of soil
560 moisture (Western et al., 1998b; Lin et al., 2006). For the FCM SEA a smaller number of
561 observation points is needed (144 ha, ~50 points) to estimate soil moisture maps. Thus, FCM
562 SEA is a promising approach providing at least as accurate results for wet moisture states than
563 traditional techniques but with considerably less effort. However, for the intermediate state and
564 especially for the grassland areas the FCM SEA performance is poor and demonstrates the lack of
565 topographic data to explain soil moisture variability. On the other hand to take more than one
566 sample out of each cluster might be of benefit to further improve the prediction accuracy. At
567 present only one sampling point was selected per cluster in the grassland areas and two sampling
568 points per cluster for arable land to calibrate the FCM interpolation method. Our results show that
569 the small scale variability of soil moisture is much higher and especially in clusters that show
570 large within-cluster variability it is advisable to increase the number of sampling points. To test
571 the performance of FCM SEA in a broader sense and for other sites the use of comprehensive
572 datasets from existing networks can motivate further research to verify the estimated maps for
573 different soil moisture states. In addition, with the FCM SEA we could suggest how to reduce the
574 number of existing measurement points in the networks and make the observations more efficient
575 with respect to measurement costs and maintenance effort. Nevertheless, this requires initially a
576 very good prediction of the soil moisture maps and for the moment we first need to improve our
577 method by adding additional variables (e.g. soil texture, land use) to reach this goal, which is an
578 ongoing work in the Schäfertal catchment. At the moment there is still some work to be done and
579 one should also not forget that the performance of the method might be site specific with respect
580 to the information/attributes required to estimate good soil moisture maps. At one site, the

581 clustering might work well with topographic data only while at other sites other properties (e.g.,
582 texture) might be responsible for driving soil moisture dynamics.

583 When further improved with additional variables such as soil texture and vegetation and carefully
584 validated, such predicted maps might become useful for validation and calibration of remote
585 sensing data (see e.g. Crow et al., 2012) and distributed models, while in future remote sensing
586 might replace time consuming measurements. For the moment, ground based measurements are
587 still indispensable and our proposed framework might become valuable to accompany remote
588 sensing campaigns and modeling approaches to provide insights into the hydrological behavior of
589 small catchments.

590 *A priori* knowledge in terms of a careful and hypothesis-driven selection of soil-moisture
591 controlling variables is a first step to increase the sampling efficiency for soil moisture
592 monitoring and to test the ability of the integrated data to describe the temporal dynamics of soil
593 moisture. In a second step and based on the previously obtained results, *a posteriori* knowledge is
594 gained and can be included in the next step by adding further proxy data that account for
595 processes which were missing in the first step. Therefore, we see our proposed FCM SEA as a
596 new learning framework for understanding the function of hydrological systems.

597 **Summary and Conclusions**

598 In this paper we applied a terrain-based FCM sampling and estimation approach (FCM SEA) to
599 identify and characterize temporal dynamics of soil moisture in a small-scale catchment. A set of
600 topographic attributes was selected (i.e., elevation, slope, SWI, TIR) to represent lateral flow and
601 topographically modulated evaporative forcing. Based on this data set the FCM SEA identifies *a*
602 *priori* an appropriate number of representative monitoring locations by stratifying the landscape

603 in SLDs (unique combinations of topographic attribute values). At these points, near surface soil
604 moisture (0-10 cm) was measured at five different sampling dates and the FCM SEA method is
605 able to predict reasonable soil moisture patterns.

606 For the Schäferfetal catchment results indicate that there is a transition between states characterized
607 by a re-organization of the soil moisture pattern. During wet conditions, there is high degree of
608 spatial organization which decreases as the soil gradually dries (intermediate conditions).

609 The independent validation revealed that the FCM SEA performed well and was able to explain
610 0.59% to 0.78% of the spatial variability of soil moisture under wet conditions, whereas under
611 intermediate conditions its explanatory power decreased. However, the terrain-based FCM SEA
612 was still able to account for more than 34% of the variability. Therefore, for the Schäferfetal
613 catchment, the FCM SEA is promising and superior to most studies that generally explained
614 <50% variance.

615 We attribute the formation of the two distinct soil moisture patterns to a combination of several
616 factors: (i) under wet conditions topography is the major control and drives water redistribution
617 due to surface and subsurface lateral flow; (ii) at intermediate states the relative importance of
618 other factors such as soil texture and vegetation become increasingly important. However, a
619 detailed investigation of the relative contribution of these factors has not been done so far and
620 will be part of future studies.

621 **Acknowledgements**

622 We would like to thank Stefan Gerlach (Agrargenossenschaft Siptenfelde) who has provided
623 access to his fields and for his cooperation throughout this project. Assistance in field data
624 collection was provided by Thomas Grau, Edoardo Martini, Andreas Schmidt, Martin Schrön,

625 and Simon Kögler. We also thank Frido Reinstorf and Hermann John (University of Applied
626 Sciences Magdeburg-Stendal) for collaboration and support, Steffen Zacharias for discussions
627 and Hans-Jörg Vogel for comments on the manuscript. The DEM was provided with kind support
628 of the LVermGeo LSA. We would also like to thank the two anonymous reviewers for
629 constructive comments and suggestions which very much helped to improve the paper. This work
630 was funded by the Helmholtz Alliance EDA – Remote Sensing and Earth System Dynamics,
631 through the Initiative and Networking Fund of the Helmholtz Association, Germany. The
632 research was supported by TERENO (TERrestrial ENvironmental Observatories) and by
633 Helmholtz Impulse and Networking Fund through Helmholtz Interdisciplinary Graduate School
634 for Environmental Research (HIGRADE) (Bissinger and Kolditz, 2008).

635

636 **Figure Captions**

637 **Figure 1** Flow chart illustrating the use of the FCM cluster analysis for data/ model integration
 638 which is traditionally used for classification purposes.

639 **Figure 2** Flow chart illustrating the FCM clustering technique used to find the adequate number
 640 of clusters for a strategic sampling and to regionally predict the target value (e.g., soil moisture).

641 **Figure 3** Total absolute difference (TAD) for different number of clusters. Red line indicates
 642 appropriate number of clusters according to elbow method.

643 **Figure 4** Location, topography and initial land use of the Schäfertal catchment. Meanwhile the
 644 grassland between the two roads has been transformed into arable land. Topographical data:
 645 DGM1 © GeoBasis-DE / LVermGeo LSA [2012, A13-6001119-2012]

646 **Figure 5** Terrain attributes (a) elevation, (b) SAGA wetness index , (c) slope, and (d) annual
 647 potential incoming solar radiation derived form a 2-m DEM for the Schäfertal catchment. White
 648 lines represent creek and roads that were masked and not used in the analysis. Topographical
 649 data: DGM1 © GeoBasis-DE / LVermGeo LSA [2012, A13-6001119-2012]

650 **Figure 6** Map of the 30 SLDs obtained with the FCM SEA.

651 **Figure 7** Distribution of the soil moisture measurement locations. Black: Locations obtained from
 652 the FCM clustering technique (see Fig. 2) used for calibration (50 points). Red: Independent
 653 locations used for validation (44 points). White lines represent creek and roads.

654 **Figure 8** Observed soil moisture patterns in the Schäfertal catchment for five occasions. Each dot
 655 represents the average volumetric soil moisture of three replicate TDR measurements in the top
 656 10 cm of the soil profile.

657 **Figure 9** Predicted maps of volumetric soil moisture content using the FCM interpolation method
 658 (Fig. 2) for different moisture states.

659 **Figure 10** Comparison between the predicted volumetric soil moisture using the FCM
 660 interpolation method and the observed soil moisture for the five sampling dates (a-d). Horizontal
 661 bars indicate ± 1 standard deviation of the observed soil moisture. (NSE: Nash-Sutcliffe
 662 coefficient of efficiency; RMSE: root mean square error)

663 **Table 1** Main characteristics of the selected measurement campaigns in the Schäferfetal

| Campaign | Date | Sample Size | Soil Moisture [m^3/m^3] | | | | | | Antecedent |
|----------|------------|-------------|---|--------|--------------------|------|------|-------|----------------------------------|
| | | | Mean | Median | Standard Deviation | Min | Max | Range | Precipitation 5 Days, [mm] |
| T1 | 17/04/2013 | 94 | 0.34 | 0.29 | 0.11 | 0.22 | 0.71 | 0.49 | 15.0 |
| T2 | 23/04/2013 | 94 | 0.26 | 0.23 | 0.12 | 0.14 | 0.69 | 0.55 | 0 |
| T3 | 08/05/2013 | 94 | 0.27 | 0.24 | 0.09 | 0.16 | 0.69 | 0.53 | 8.8 |
| T4 | 25/09/2013 | 94 | 0.19 | 0.17 | 0.07 | 0.12 | 0.48 | 0.36 | 0 |
| T5 | 02/10/2013 | 94 | 0.18 | 0.16 | 0.07 | 0.10 | 0.47 | 0.37 | 0.6 |

664

665 **Table 2** Spearman rank correlation coefficients between dates for the entire study period

| Date | 17/04/2013 | 23/04/2013 | 08/05/2013 | 25/09/2013 | 02/10/2013 |
|----------|------------|------------|------------|------------|------------|
| Campaign | T1 | T2 | T3 | T4 | T5 |
| T1 | 1 | | | | |
| T2 | 0.95 | 1 | | | |
| T3 | 0.87 | 0.92 | 1 | | |
| T4 | 0.57 | 0.49 | 0.48 | 1 | |
| T5 | 0.66 | 0.62 | 0.59 | 0.79 | 1 |

666

667 **Table 3** Performance of the FCM SEA for the validation of arable land, grassland and the whole
 668 catchment. (NSE: Nash-Sutcliffe coefficient of efficiency; RMSE: root mean square error)

| Campaign | Date | validation for arable land | | validation for grassland | | validation for whole catchment (see Fig. 10) | |
|----------|------------|-------------------------------|------|-----------------------------|------|---|------|
| | | NSE | RMSE | NSE | RMSE | NSE | RMSE |
| T1 | 17/04/2013 | 0.54 | 0.05 | 0.50 | 0.08 | 0.78 | 0.06 |
| T2 | 23/04/2013 | 0.52 | 0.04 | 0.37 | 0.10 | 0.73 | 0.06 |
| T3 | 08/05/2013 | 0.56 | 0.03 | 0.06 | 0.11 | 0.59 | 0.06 |
| T4 | 25/09/2013 | -0.35 | 0.04 | -0.28 | 0.09 | 0.34 | 0.06 |
| T5 | 02/10/2013 | 0.28 | 0.03 | -0.31 | 0.09 | 0.41 | 0.05 |

669

670 References

- 671 Abdank, H., Steininger, M., Altermann, M., 1995. Untersuchungen zur Landnutzung und zum
672 Gewässerschutz im Unterharz. *Mitteilungen Dtsch. Bodenkd. Ges.* 77 166–173.
- 673 Albertson, J.D., Montaldo, N., 2003. Temporal dynamics of soil moisture variability: 1.
674 Theoretical basis. *Water Resour. Res.* 39, 1274. doi:10.1029/2002WR001616
- 675 Altermann, M., 1985. Standortkennzeichnung landwirtschaftlich genutzter Gebiete des östlichen
676 Harzes, Habilitationsschrift. Universität Rostock.
- 677 Applegate, D., Bixby, R., Chvátal, V., Cook, W., 2001. TSP Cuts Which Do Not Conform to the
678 Template Paradigm, in: Jünger, M., Naddef, D. (Eds.), *Computational Combinatorial
679 Optimization, Lecture Notes in Computer Science.* Springer Berlin Heidelberg, pp. 261–
680 303.
- 681 Baggaley, N., Mayr, T., Bellamy, P., 2009. Identification of key soil and terrain properties that
682 influence the spatial variability of soil moisture throughout the growing season. *Soil Use
683 Manag.* 25, 262–273. doi:10.1111/j.1475-2743.2009.00222.x
- 684 Baroni, G., Ortuani, B., Facchi, A., Gandolfi, C., 2013. The role of vegetation and soil properties
685 on the spatio-temporal variability of the surface soil moisture in a maize-cropped field. *J.
686 Hydrol.* 489, 148–159. doi:10.1016/j.jhydrol.2013.03.007
- 687 Beaudette, D.E., Dahlgren, R.A., O’Geen, A.T., 2013. Terrain-Shape Indices for Modeling Soil
688 Moisture Dynamics. *Soil Sci. Soc. Am. J.* 77, 1696–1710. doi:10.2136/sssaj2013.02.0048
- 689 Behrens, T., Zhu, A.-X., Schmidt, K., Scholten, T., 2010. Multi-scale digital terrain analysis and
690 feature selection for digital soil mapping. *Geoderma* 155, 175–185.
691 doi:10.1016/j.geoderma.2009.07.010
- 692 Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin
693 hydrology / Un modèle à base physique de zone d’appel variable de l’hydrologie du
694 bassin versant. *Hydrol. Sci. Bull.* 24, 43–69. doi:10.1080/02626667909491834

- 695 Bezdek, D.J.C., 1974. Numerical taxonomy with fuzzy sets. *J. Math. Biol.* 1, 57–71.
696 doi:10.1007/BF02339490
- 697 Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: The fuzzy c-means clustering algorithm.
698 *Comput. Geosci.* 10, 191–203. doi:10.1016/0098-3004(84)90020-7
- 699 Bissinger, V., Kolditz, O., 2008. Helmholtz Interdisciplinary Graduate School for Environmental
700 Research (HIGRADE) GAIA 1(2008):71–73.
- 701 Bogena, H.R., Herbst, M., Huisman, J.A., Rosenbaum, U., Weuthen, A., Vereecken, H., 2010.
702 Potential of Wireless Sensor Networks for Measuring Soil Water Content Variability.
703 *Vadose Zone J.* 9, 1002–1013. doi:10.2136/vzj2009.0173
- 704 Böhner, J., Selige, T., 2006. Spatial prediction of soil attributes using terrain analysis and climate
705 regionalisation, in: Böhner, J., McCloy, K.R., Strobl, J. (Eds.), *SAGA—Analyses and*
706 *Modelling Applications*. Göttinger Geographische Abhandlungen, Göttingen, pp. 13–28.
- 707 Borchardt, D., 1982. Geoökologische Erkundung und hydrologische Analyse von
708 Kleinzugsgebieten des unteren Mittelgebirgsbereiches, dargestellt am Beispiel von
709 Experimentalgebieten der oberen Selke/Harz. *Petermanns Geogr. Mitteilungen* 482 251–
710 262.
- 711 Bronstert, A., Creutzfeldt, B., Graeff, T., Hajnsek, I., Heistermann, M., Itzerott, S., Jagdhuber, T.,
712 Kneis, D., Lück, E., Reusser, D., Zehe, E., 2012. Potentials and constraints of different
713 types of soil moisture observations for flood simulations in headwater catchments. *Nat.*
714 *Hazards* 60, 879–914. doi:10.1007/s11069-011-9874-9
- 715 Burrough, P.A., van Gaans, P.F.M., MacMillan, R.A., 2000. High-resolution landform
716 classification using fuzzy k-means. *Fuzzy Sets Syst.* 113, 37–52. doi:10.1016/S0165-
717 0114(99)00011-1
- 718 Cardell-Oliver, R., Kranz, M., Smettem, K., Mayer, K., 2005. A Reactive Soil Moisture Sensor
719 Network: Design and Field Evaluation. *Int. J. Distrib. Sens. Netw.* 1, 149–162.
720 doi:10.1080/15501320590966422
- 721 Crow, W.T., Berg, A.A., Cosh, M.H., Loew, A., Mohanty, B.P., Panciera, R., de Rosnay, P.,
722 Ryu, D., Walker, J.P., 2012. Upscaling sparse ground-based soil moisture observations
723 for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.* 50,
724 RG2002. doi:10.1029/2011RG000372
- 725 De Gruijter, J.J., Brus, D., Bierkens, M., Knotters, M., 2006. *Sampling for natural resource*
726 *monitoring*. Springer, Berlin.
- 727 Famiglietti, J.S., Rudnicki, J.W., Rodell, M., 1998. Variability in surface moisture content along
728 a hillslope transect: Rattlesnake Hill, Texas. *J. Hydrol.* 210, 259–281. doi:10.1016/S0022-
729 1694(98)00187-5
- 730 Grayson, R.B., Blöschl, G., Western, A.W., McMahon, T.A., 2002. Advances in the use of
731 observed spatial patterns of catchment hydrological response. *Adv. Water Resour.* 25,
732 1313–1334. doi:10.1016/S0309-1708(02)00060-X
- 733 Grayson, R.B., Western, A.W., 1998. Towards areal estimation of soil water content from point
734 measurements: time and space stability of mean response. *J. Hydrol.* 207, 68–82.
735 doi:10.1016/S0022-1694(98)00096-1
- 736 Grayson, R.B., Western, A.W., Chiew, F.H.S., Blöschl, G., 1997. Preferred states in spatial soil
737 moisture patterns: Local and nonlocal controls. *Water Resour. Res.* 33, 2897–2908.
738 doi:10.1029/97WR02174

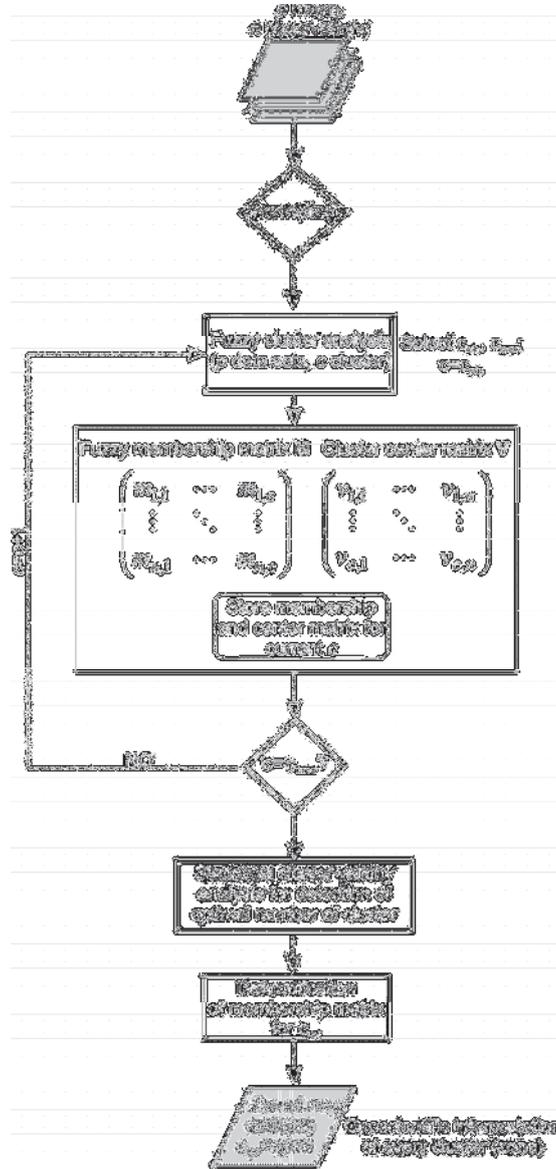
- 739 Hachmöller, B., Paasche, H., 2013. Integration of surface-based tomographic models for zonation
740 and multimodel guided extrapolation of sparsely known petrophysical parameters.
741 *GEOPHYSICS* 78, EN43–EN53. doi:10.1190/geo2012-0417.1
- 742 Hudson, B.D., 1992. The Soil Survey as Paradigm-based Science. *Soil Sci. Soc. Am. J.* 56, 836.
743 doi:10.2136/sssaj1992.03615995005600030027x
- 744 Hupet, F., Vanclooster, M., 2002. Intraseasonal dynamics of soil moisture variability within a
745 small agricultural maize cropped field. *J. Hydrol.* 261, 86–101. doi:10.1016/S0022-
746 1694(02)00016-1
- 747 Hu, W., Si, B.C., 2014a. Revealing the relative influence of soil and topographic properties on
748 soil water content distribution at the watershed scale in two sites. *J. Hydrol.*,
749 Determination of soil moisture: Measurements and theoretical approaches 516, 107–118.
750 doi:10.1016/j.jhydrol.2013.10.002
- 751 Hu, W., Si, B.C., 2014b. Revealing the relative influence of soil and topographic properties on
752 soil water content distribution at the watershed scale in two sites. *J. Hydrol.*,
753 Determination of soil moisture: Measurements and theoretical approaches 516, 107–118.
754 doi:10.1016/j.jhydrol.2013.10.002
- 755 Jenny, H., 1941. *Factors of Soil Formation: A System of Quantitative Pedology*. McGraw-Hill
756 Book Co., New York.
- 757 Kachanoski, R.G., de Jong, E., 1988. Scale dependence and the temporal persistence of spatial
758 patterns of soil water storage. *Water Resour. Res.* 24, 85–91.
759 doi:10.1029/WR024i001p00085
- 760 Lassonde, M., 2001. *Approximation, Optimization and Mathematical Economics*, Softcover
761 reprint of the original 1st ed. 2001. ed. Physica.
- 762 Leekwijck, W.V., Kerre, E.E., 1999. Defuzzification: criteria and classification. *Fuzzy Sets Syst.*
763 108, 159–178. doi:10.1016/S0165-0114(97)00337-0
- 764 Lesch, S.M., 2005. Sensor-directed response surface sampling designs for characterizing spatial
765 variation in soil properties. *Comput. Electron. Agric., Applications of Apparent Soil*
766 *Electrical Conductivity in Precision Agriculture* 46, 153–179.
767 doi:10.1016/j.compag.2004.11.004
- 768 Lin, H., 2003. *Hydropedology Bridging Disciplines, Scales, and Data*. *Vadose Zone J.* 2, 1–11.
769 doi:10.2113/2.1.1
- 770 Lin, H.S., Kogelmann, W., Walker, C., Bruns, M.A., 2006. Soil moisture patterns in a forested
771 catchment: A hydropedological perspective. *Geoderma* 131, 345–368.
772 doi:10.1016/j.geoderma.2005.03.013
- 773 MacMillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W., 2000. A generic procedure for
774 automatically segmenting landforms into landform elements using DEMs, heuristic rules
775 and fuzzy logic. *Fuzzy Sets Syst.* 113, 81–109. doi:10.1016/S0165-0114(99)00014-7
- 776 Martinez, C., Hancock, G.R., Kalma, J.D., Wells, T., 2008. Spatio-temporal distribution of near-
777 surface and root zone soil moisture at the catchment scale. *Hydrol. Process.* 22, 2699–
778 2714. doi:10.1002/hyp.6869
- 779 Martínez-Fernández, J., Ceballos, A., 2005. Mean soil moisture estimation using temporal
780 stability analysis. *J. Hydrol.* 312, 28–38. doi:10.1016/j.jhydrol.2005.02.007
- 781 Martini, E., Wollschläger, U., Kögler, S., Behrens, T., Dietrich, P., Reinstorf, F., Schmidt, K.,
782 Weiler, M., Werban, U., Zacharias, S., 2015. Spatial and Temporal Dynamics of

- 783 Hillslope-Scale Soil Moisture Patterns: Characteristic States and Transition Mechanisms.
784 Vadose Zone J. 0, 0. doi:10.2136/vzj2014.10.0150
- 785 McBratney, A.B., Santos, M.L.M., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117,
786 3–52. doi:10.1016/S0016-7061(03)00223-4
- 787 McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental
788 correlation. *Geoderma* 89, 67–94. doi:10.1016/S0016-7061(98)00137-2
- 789 Merz, B., Plate, E.J., 1997. An analysis of the effects of spatial variability of soil and soil
790 moisture on runoff. *Water Resour. Res.* 33, 2909–2922. doi:10.1029/97WR02204
- 791 Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., [C++-code - Chih-Chung
792 Chang (libsvm)], [C++-code - Chih-Chen Lin (libsvm)], 2014. e1071: Misc Functions of
793 the Department of Statistics (e1071), TU Wien, R package version 1.6-1 <[http://cran.r-](http://cran.r-project.org/package=e1071)
794 [project.org/package=e1071](http://cran.r-project.org/package=e1071) (Accessed February 10, 2013).
- 795 Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the
796 presence of ancillary information. *Comput. Geosci.* 32, 1378–1388.
797 doi:10.1016/j.cageo.2005.12.009
- 798 Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: A review of
799 hydrological, geomorphological, and biological applications. *Hydrol. Process.* 5, 3–30.
800 doi:10.1002/hyp.3360050103
- 801 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A
802 discussion of principles. *J. Hydrol.* 10, 282–290. doi:10.1016/0022-1694(70)90255-6
- 803 Odeh, I.O.A., Chittleborough, D.J., McBratney, A.B., 1992. Soil Pattern Recognition with Fuzzy-
804 c-means: Application to Classification and Soil-Landform Interrelationships. *Soil Sci.*
805 *Soc. Am. J.* 56, 505. doi:10.2136/sssaj1992.03615995005600020027x
- 806 Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1990. Design of optimal sample spacings
807 for mapping soil using fuzzy-k-means and regionalized variable theory. *Geoderma* 47,
808 93–122. doi:10.1016/0016-7061(90)90049-F
- 809 Ollesch, G., Kistner, I., Meissner, R., Lindenschmidt, K.-E., 2006. Modelling of snowmelt
810 erosion and sediment yield in a small low-mountain catchment in Germany. *CATENA* 68,
811 161–176. doi:10.1016/j.catena.2006.04.005
- 812 Paasche, H., Eberle, D.G., 2009. Rapid integration of large airborne geophysical data suites using
813 a fuzzy partitioning cluster algorithm: a tool for geological mapping and mineral
814 exploration targeting. *Explor. Geophys.* 40, 277–287.
- 815 Paasche, H., Tronicke, J., Holliger, K., Green, A.G., Maurer, H., 2006. Integration of diverse
816 physical-property models: Subsurface zonation and petrophysical parameter estimation
817 based on fuzzy c-means cluster analyses. *Geophysics* 71, H33–H44.
818 doi:10.1190/1.2192927
- 819 Penna, D., Borga, M., Norbiato, D., Dalla Fontana, G., 2009. Hillslope scale soil moisture
820 variability in a steep alpine terrain. *J. Hydrol.* 364, 311–327.
821 doi:10.1016/j.jhydrol.2008.11.009
- 822 Qu, W., Bogena, H.R., Huisman, J.A., Vanderborght, J., Schuh, M., Priesack, E., Vereecken, H.,
823 2015. Predicting subgrid variability of soil water content from basic soil information.
824 *Geophys. Res. Lett.* 42, 2014GL062496. doi:10.1002/2014GL062496
- 825 R Development Core Team, 2012. R: A language and environment for statistical computing. R
826 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0
827 <<http://www.R-project.org>> (Accessed August 15, 2012).

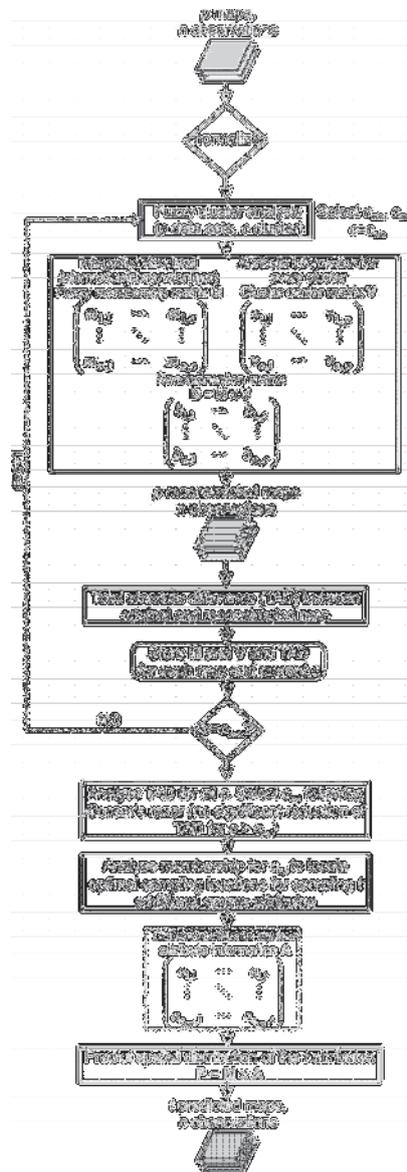
- 828 Reinstorf, F., 2010. Schäfertal, Harz Mountains, Germany. Poster. In: Schumann, S., B.
829 Schmalz, H. Meesenburg, and U. Schröder (Eds.), “Status and Perspectives of
830 Hydrology in Small Basins”, Results and recommendations of the International
831 Workshop in Goslar-Hahnenklee, Germany 2009, and Inventory of Small Hydrological
832 Research Basin. 30 March – 2 April 2009, Goslar-Hahnenklee, Germany.
- 833 Robinson, D.A., Campbell, C.S., Hopmans, J.W., Hornbuckle, B.K., Jones, S.B., Knight, R.,
834 Ogden, F., Selker, J., Wendroth, O., 2008. Soil Moisture Measurement for Ecological and
835 Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone J.* 7, 358.
836 doi:10.2136/vzj2007.0143
- 837 Rosenbaum, U., Bogena, H.R., Herbst, M., Huisman, J.A., Peterson, T.J., Weuthen, A., Western,
838 A.W., Vereecken, H., 2012. Seasonal and event dynamics of spatial soil moisture patterns
839 at the small catchment scale. *Water Resour. Res.* 48, n/a–n/a.
840 doi:10.1029/2011WR011518
- 841 Roth, K., Schulin, R., Flühler, H., Attinger, W., 1990. Calibration of time domain reflectometry
842 for water content measurement using a composite dielectric approach. *Water Resour. Res.*
843 26, 2267–2273. doi:10.1029/WR026i010p02267
- 844 Roubens, M., 1982. Fuzzy clustering algorithms and their cluster validity. *Eur. J. Oper. Res.* 10,
845 294–301. doi:10.1016/0377-2217(82)90228-4
- 846 Schmidt, K., Behrens, T., Daumann, J., Ramirez-Lopez, L., Werban, U., Dietrich, P., Scholten,
847 T., 2014. A comparison of calibration sampling schemes at the field scale. *Geoderma*
848 232–234, 243–256. doi:10.1016/j.geoderma.2014.05.013
- 849 Schneider, K., Huisman, J.A., Breuer, L., Zhao, Y., Frede, H.-G., 2008. Temporal stability of soil
850 moisture in various semi-arid steppe ecosystems and its application in remote sensing. *J.*
851 *Hydrol.* 359, 16–29. doi:10.1016/j.jhydrol.2008.06.016
- 852 Schulz, K., Seppelt, R., Zehe, E., Vogel, H.J., Attinger, S., 2006. Importance of spatial structures
853 in advancing hydrological sciences. *Water Resour. Res.* 42, W03S03.
854 doi:10.1029/2005WR004301
- 855 Soulsby, C., Neal, C., Laudon, H., Burns, D.A., Merot, P., Bonell, M., Dunn, S.M., Tetzlaff, D.,
856 2008. Catchment data for process conceptualization: simply not enough? *Hydrol. Process.*
857 22, 2057–2061. doi:10.1002/hyp.7068
- 858 Stevenson, B.A., McNeill, S., Hewitt, A.E., 2015. Characterising soil quality clusters in relation
859 to land use and soil order in New Zealand: An application of the phenoform concept.
860 *Geoderma* 239–240, 135–142. doi:10.1016/j.geoderma.2014.10.003
- 861 Sun, X.-L., Zhao, Y.-G., Wang, H.-L., Yang, L., Qin, C.-Z., Zhu, A.-X., Zhang, G.-L., Pei, T., Li,
862 B.-L., 2012. Sensitivity of digital soil maps based on FCM to the fuzzy exponent and the
863 number of clusters. *Geoderma, Entering the Digital Era: Special Issue of Pedometrics*
864 2009, Beijing 171–172, 24–34. doi:10.1016/j.geoderma.2011.03.016
- 865 Takagi, K., Lin, H.S., 2012. Changing controls of soil moisture spatial organization in the Shale
866 Hills Catchment. *Geoderma* 173–174, 289–302. doi:10.1016/j.geoderma.2011.11.003
- 867 Triantafilis, J., Gibbs, I., Earl, N., 2013. Digital soil pattern recognition in the lower Namoi
868 valley using numerical clustering of gamma-ray spectrometry data. *Geoderma* 192, 407–
869 421. doi:10.1016/j.geoderma.2012.08.021
- 870 Triantafilis, J., Odeh, I.O.A., Minasny, B., McBratney, A.B., 2003. Elucidation of physiographic
871 and hydrogeological features of the lower Namoi valley using fuzzy k-means

- 872 classification of EM34 data. *Environ. Model. Softw.* 18, 667–680. doi:10.1016/S1364-
873 8152(03)00053-7
- 874 Twarakavi, N.K.C., Šimůnek, J., Schaap, M.G., 2010. Can texture-based classification optimally
875 classify soils with respect to soil hydraulics? *Water Resour. Res.* 46, W01501.
876 doi:10.1029/2009WR007939
- 877 Vachaud, G., Passerat De Silans, A., Balabanis, P., Vauclin, M., 1985. Temporal Stability of
878 Spatially Measured Soil Water Probability Density Function. *Soil Sci. Soc. Am. J.* 49,
879 822. doi:10.2136/sssaj1985.03615995004900040006x
- 880 Van Arkel, Z., Kaleita, A.L., 2014. Identifying sampling locations for field-scale soil moisture
881 estimation using K-means clustering. *Water Resour. Res.* 50, 7050–7057.
882 doi:10.1002/2013WR015015
- 883 Vereecken, H., Huisman, J.A., Pachepsky, Y., Montzka, C., van der Kruk, J., Bogen, H.,
884 Weihermüller, L., Herbst, M., Martinez, G., Vanderborght, J., 2014. On the spatio-
885 temporal dynamics of soil moisture at the field scale. *J. Hydrol., Determination of soil*
886 *moisture: Measurements and theoretical approaches* 516, 76–96.
887 doi:10.1016/j.jhydrol.2013.11.061
- 888 Wagner, W., Bloeschl, G., Pampaloni, P., Calvet, J.-C., Bizzarri, B., Wigneron, J.-P., Kerr, Y.,
889 2007. Operational readiness of microwave remote sensing of soil moisture for hydrologic
890 applications. *Nord. Hydrol.* 38, 1–20. doi:10.2166/nh.2007.029
- 891 Werbylo, K.L., Niemann, J.D., 2014. Evaluation of sampling techniques to characterize
892 topographically-dependent variability for soil moisture downscaling. *J. Hydrol.,*
893 *Determination of soil moisture: Measurements and theoretical approaches* 516, 304–316.
894 doi:10.1016/j.jhydrol.2014.01.030
- 895 Western, A., Grayson, R., 2000. Soil moisture and runoff processes at Tarrawarra. In: Grayson,
896 R.; Blöschl G. (Eds.), *Spatial Patterns in Catchment Hydrology Observations and*
897 *Modelling*. Cambridge: Cambridge University Press; 2000. p. 209–246 [chapter 9].
- 898 Western, A.W., Blöschl, G., Grayson, R.B., 1998. Geostatistical characterisation of soil moisture
899 patterns in the Tarrawarra catchment. *J. Hydrol.* 205, 20–37. doi:10.1016/S0022-
900 1694(97)00142-X
- 901 Western, A.W., Grayson, R.B., 1998. The Tarrawarra Data Set: Soil moisture patterns, soil
902 characteristics, and hydrological flux measurements. *Water Resour. Res.* 34, 2765–2768.
903 doi:10.1029/98WR01833
- 904 Western, A.W., Grayson, R.B., Blöschl, G., Willgoose, G.R., McMahon, T.A., 1999a. Observed
905 spatial organization of soil moisture and its relation to terrain indices. *Water Resour. Res.*
906 35, 797–810. doi:10.1029/1998WR900065
- 907 Western, A.W., Grayson, R.B., Blöschl, G., Willgoose, G.R., McMahon, T.A., 1999b. Observed
908 spatial organization of soil moisture and its relation to terrain indices. *Water Resour. Res.*
909 35, 797–810. doi:10.1029/1998WR900065
- 910 Western, A.W., Grayson, R.B., Green, T.R., 1999c. The Tarrawarra project: high resolution
911 spatial measurement, modelling and analysis of soil moisture and hydrological response.
912 *Hydrol. Process.* 13, 633–652. doi:10.1002/(SICI)1099-1085(19990415)13:5<633::AID-
913 HYP770>3.0.CO;2-8
- 914 Western, A.W., Zhou, S.-L., Grayson, R.B., McMahon, T.A., Blöschl, G., Wilson, D.J., 2004.
915 Spatial correlation of soil moisture in small catchments and its relationship to dominant

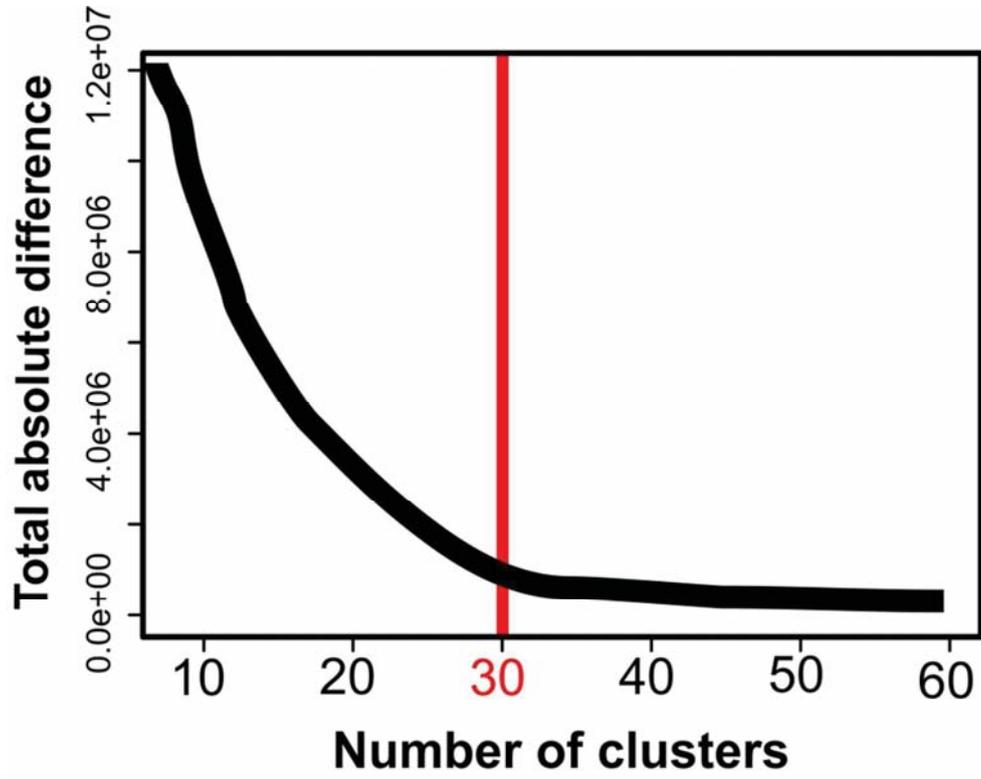
- 916 spatial hydrological processes. *J. Hydrol.* 286, 113–134.
917 doi:10.1016/j.jhydrol.2003.09.014
- 918 Wilson, D.J., Western, A.W., Grayson, R.B., 2005. A terrain and data-based method for
919 generating the spatial distribution of soil moisture. *Adv. Water Resour.* 28, 43–54.
920 doi:10.1016/j.advwatres.2004.09.007
- 921 Wilson, D.J., Western, A.W., Grayson, R.B., 2004. Identifying and quantifying sources of
922 variability in temporal and spatial soil moisture observations. *Water Resour. Res.* 40,
923 W02507. doi:10.1029/2003WR002306
- 924 Wilson, D.J., Western, A.W., Grayson, R.B., Berg, A.A., Lear, M.S., Rodell, M., Famiglietti,
925 J.S., Woods, R.A., McMahon, T.A., 2003. Spatial distribution of soil moisture over 6 and
926 30 cm depth, Mahurangi river catchment, New Zealand. *J. Hydrol.* 276, 254–274.
927 doi:10.1016/S0022-1694(03)00060-X
- 928 Wilson, J.P., Gallant, J.C., 2000. *Terrain Analysis: Principles and Applications*. John Wiley &
929 Sons, New York [etc.].
- 930 Woods, R., Grayson, R., Western, A., Duncan, M., Wilson, D., Young, R., Ibbitt, R., Henderson,
931 R., McMahon, T., 2001. Experimental Design and Initial Results from the Mahurangi
932 River Variability Experiment: MARVEX, in: Lakshmi, V., Albertson, J., Schaake, J.
933 (Eds.), *Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling*.
934 American Geophysical Union, pp. 201–213.
- 935 Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank,
936 M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P.,
937 Hajsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J.C., Papen, H.,
938 Priesack, E., Schmid, H.P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., Vereecken, H.,
939 2011. A Network of Terrestrial Environmental Observatories in Germany. *Vadose Zone J.*
940 10, 955–973. doi:10.2136/vzj2010.0139
- 941 Zreda, M., Desilets, D., Ferré, T.P.A., Scott, R.L., 2008. Measuring soil moisture content non-
942 invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophys. Res. Lett.* 35,
943 L21402. doi:10.1029/2008GL035655
- 944 Zucco, G., Brocca, L., Moramarco, T., Morbidelli, R., 2014. Influence of land use on soil
945 moisture spatial–temporal variability and monitoring. *J. Hydrol.*, Determination of soil
946 moisture: Measurements and theoretical approaches 516, 193–199.
947 doi:10.1016/j.jhydrol.2014.01.043
948



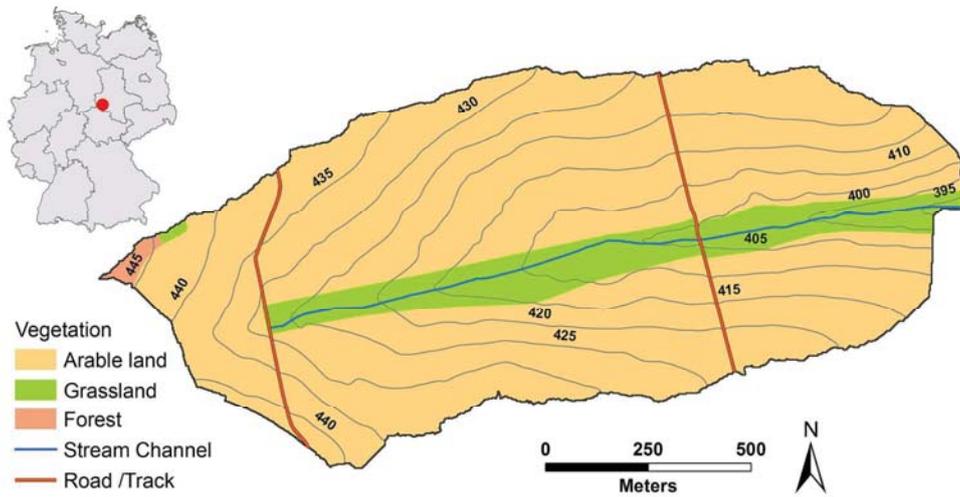
Flow chart illustrating the use of the FCM cluster analysis for data/ model integration which is traditionally used for classification purposes.
84x181mm (300 x 300 DPI)



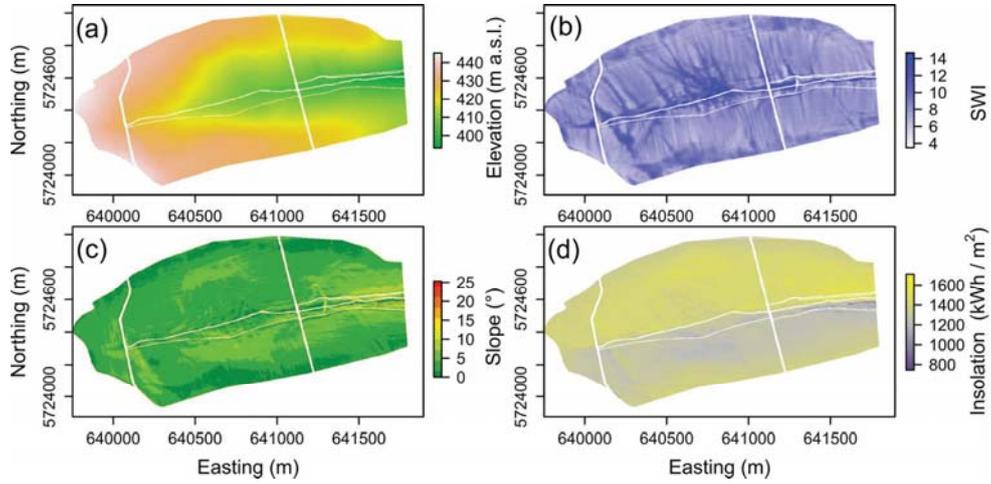
Flow chart illustrating the FCM clustering technique used to find the adequate number of clusters for a strategic sampling and to regionally predict the target value (e.g., soil moisture).
75x220mm (300 x 300 DPI)



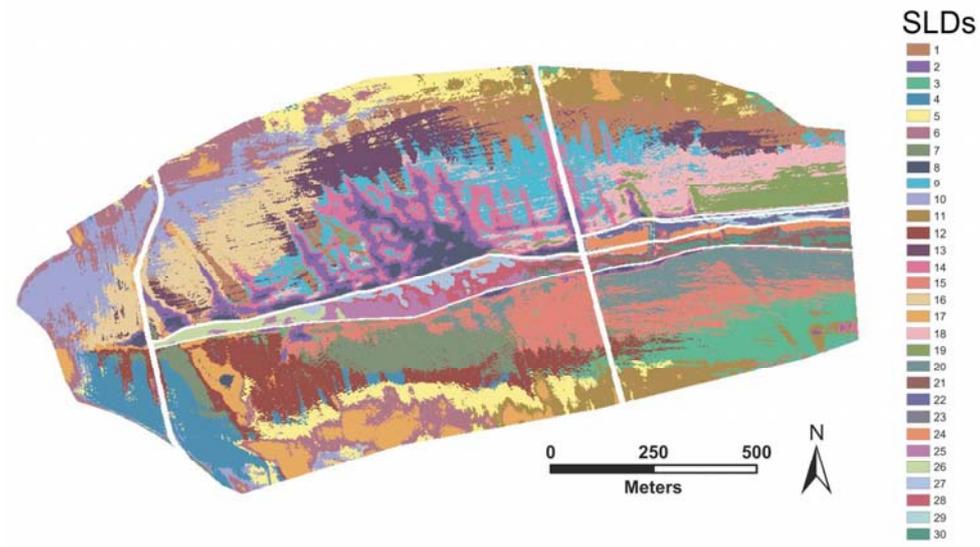
Total absolute difference (TAD) for different number of clusters. Red line indicates appropriate number of clusters according to elbow method.
82x67mm (300 x 300 DPI)



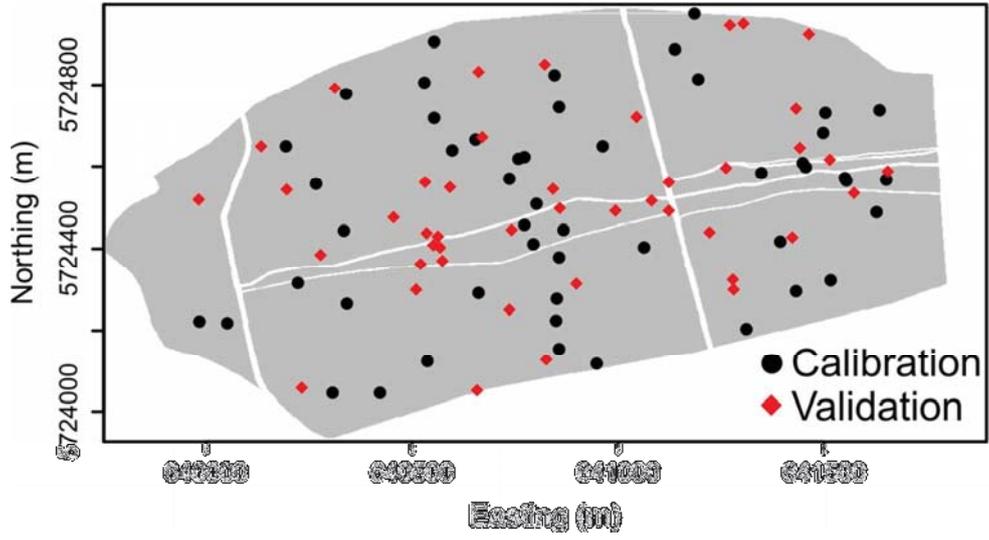
Location, topography and initial land use of the Schäfertal catchment. Meanwhile the grassland between the two roads has been transformed into arable land. Topographical data: DGM1 © GeoBasis-DE / LVermGeo LSA [2012, A13-6001119-2012] 209x148mm (300 x 300 DPI)



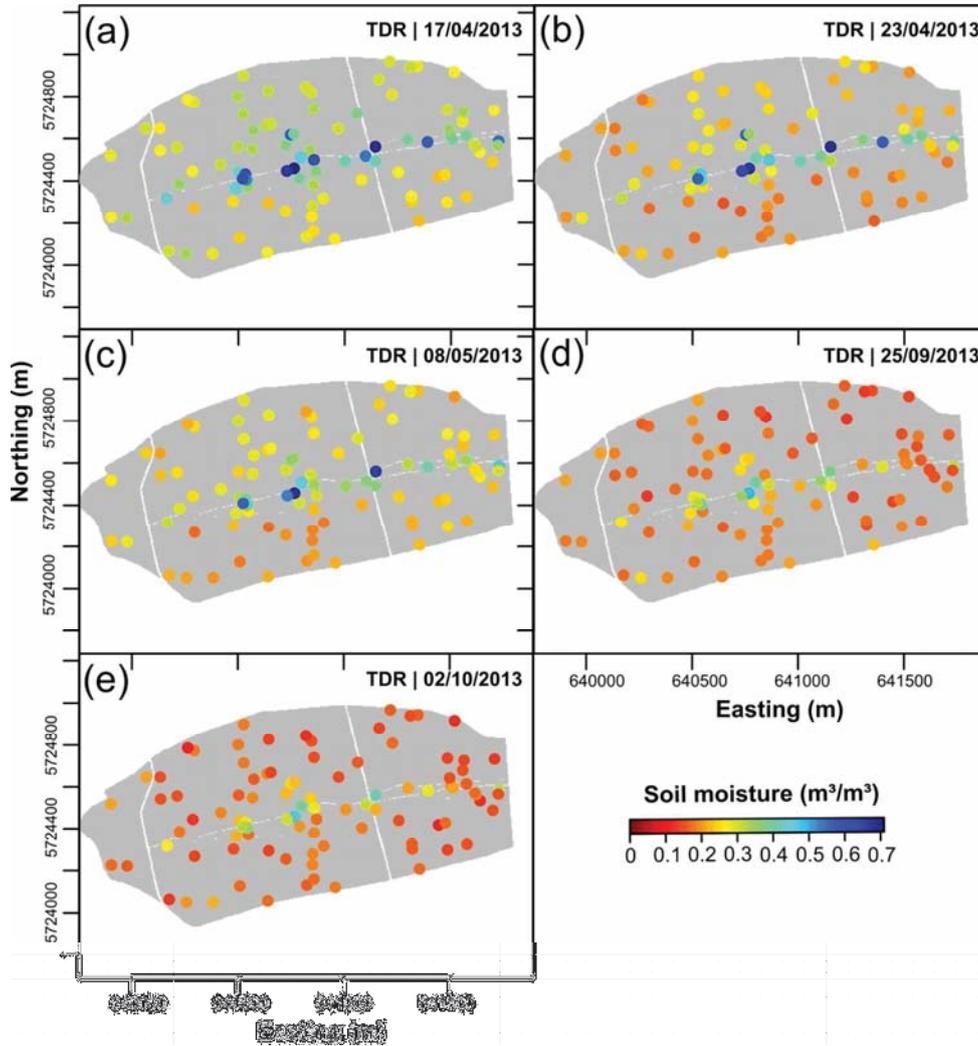
Terrain attributes (a) elevation, (b) SAGA wetness index , (c) slope, and (d) annual potential incoming solar radiation derived from a 2-m DEM for the Schäferfirtal catchment. White lines represent creek and roads that were masked and not used in the analysis. Topographical data: DGM1 © GeoBasis-DE / LVermGeo LSA [2012, A13-6001119-2012] 180x88mm (300 x 300 DPI)



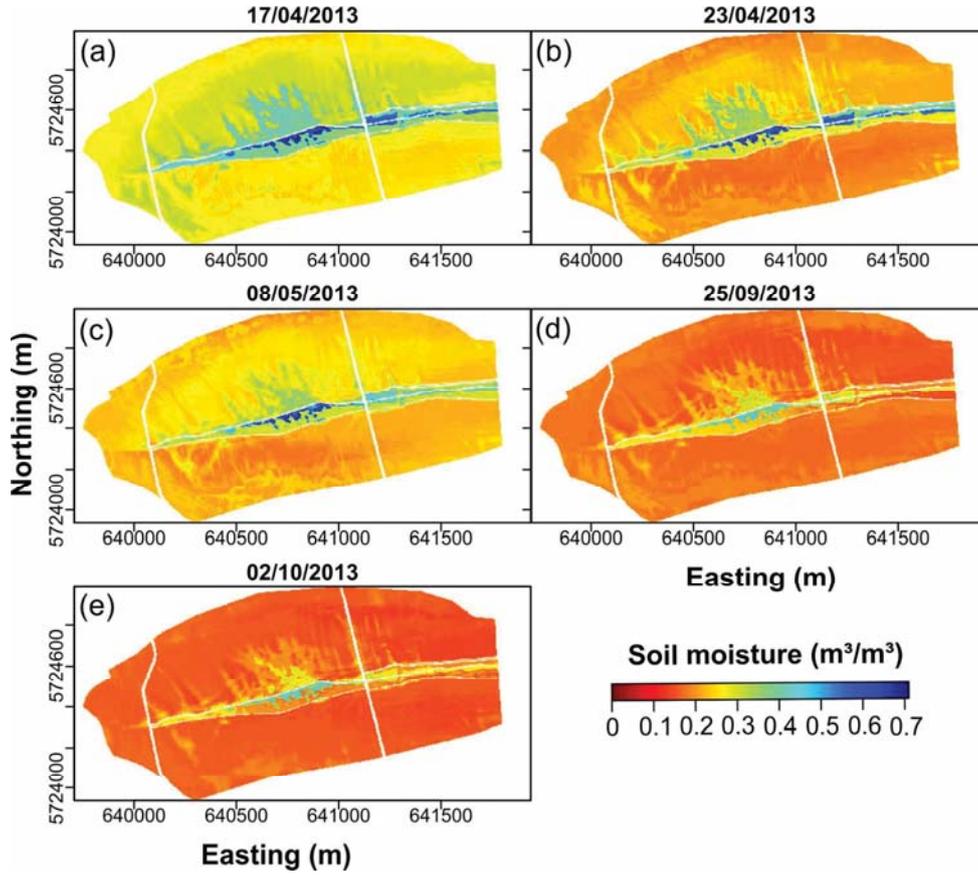
Map of the 30 SLDs obtained with the FCM SEA.
159x90mm (300 x 300 DPI)



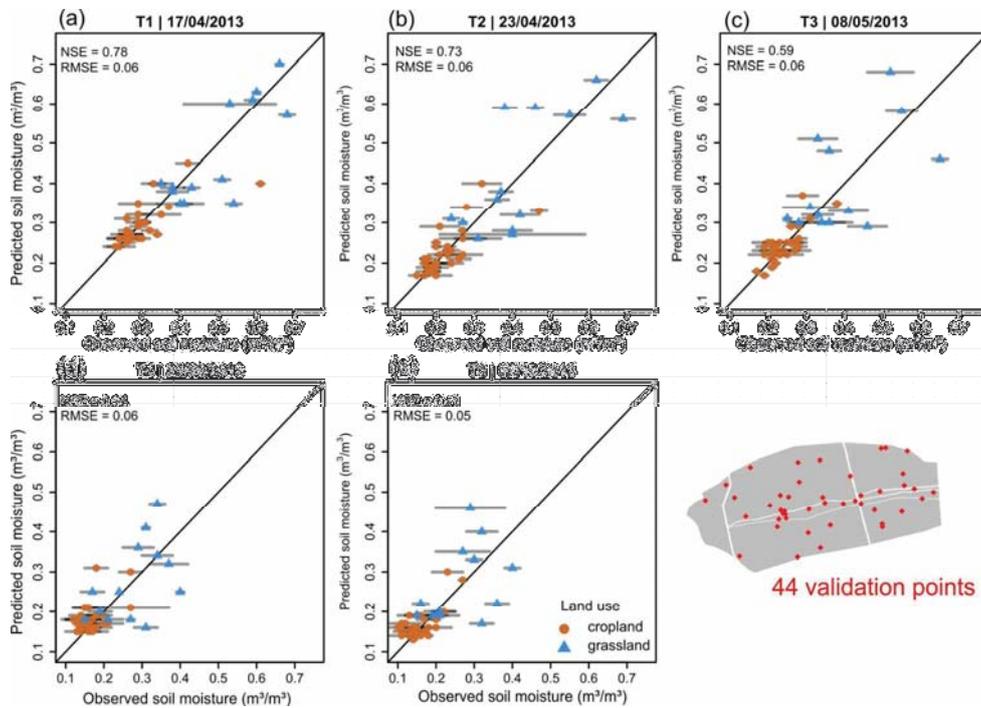
Distribution of the soil moisture measurement locations. Black: Locations obtained from the FCM clustering technique (see Fig. 2) used for calibration (50 points). Red: Independent locations used for validation (44 points). White lines represent creek and roads.
86x47mm (300 x 300 DPI)



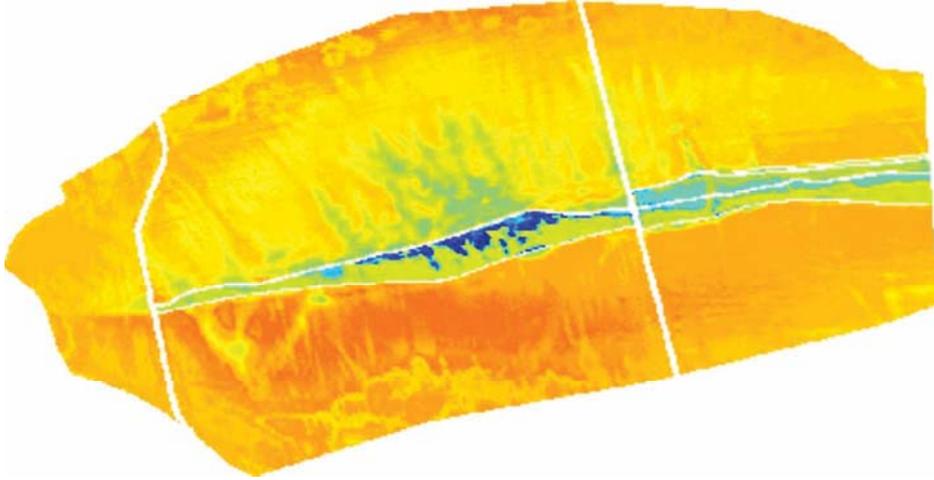
Observed soil moisture patterns in the Schäfertal catchment for five occasions. Each dot represents the average volumetric soil moisture of three replicate TDR measurements in the top 10 cm of the soil profile. 128x138mm (300 x 300 DPI)



Predicted maps of volumetric soil moisture content using the FCM interpolation method (Fig. 2) for different moisture states.
164x146mm (300 x 300 DPI)



Comparison between the predicted volumetric soil moisture using the FCM interpolation method and the observed soil moisture for the five sampling dates (a-d). Horizontal bars indicate "±1" standard deviation of the observed soil moisture. (NSE: Nash-Sutcliffe coefficient of efficiency; RMSE: root mean square error) 160x115mm (300 x 300 DPI)



Cover
94x46mm (300 x 300 DPI)