THESIS

SPATIAL PATTERNS FROM EOF/PC ANALYSIS OF SOIL MOISTURE AND THEIR DEPENDENCE ON SOIL, LAND-USE, AND TOPOGRAPHIC PROPERTIES

Submitted by
Summer Conklin
Department of Civil Engineering

In partial fulfillment of the requirements
For the Degree of Master of Science
Colorado State University
Fort Collins, Colorado
Fall 2005

COLORADO STATE UNIVERSITY

June 3, 2005

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY SUMMER CONKLIN ENTITLED SPATIAL PATTERNS FROM EOF/PC ANALYSIS OF SOIL MOISTURE AND THEIR DEPENDENCE ON SOIL, LAND-USE, AND TOPOGRAPHIC PROPERTIES BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

(Please print name under signature)

Freeman M Smith

FREEMAN M SMITH

Jorge A Ramirez JORGE A RAMIREZ

Jeffrey D. Niemann

Advisor

Jeffrey D. Niemann

Sandra Woods

Department Head

Sandra Woods

ABSTRACT OF THESIS

SPATIAL PATTERNS FROM EOF/PC ANALYSIS OF SOIL MOISTURE AND THEIR DEPENDENCE ON SOIL, LAND-USE, AND TOPOGRAPHIC PROPERTIES

Soil moisture is a highly variable parameter in both space and time, and accurate measurements are needed in hydrology and many other disciplines. While remote sensing techniques can measure near-surface soil moisture, such measurements are available at spatial resolutions that are too coarse for most applications. Thus, downscaling methods are needed. If regional characteristics that are readily available at a finer resolution are closely related to soil moisture patterns, then those characteristics could be used to downscale observations of soil moisture from remote sensing. We hypothesize that the variability in soil moisture patterns can be described by a relatively small number of spatial structures that are related to soil texture, land-use, and topographic characteristics. To test this hypothesis, an empirical orthogonal function and principal component (EOF/PC) analysis has been conducted using soil moisture data from the Southern Great Plains field campaign of 1997. This dataset contains 16 days of remotely sensed data on a 0.64 km^2 grid over nearly 10,000 km^2. From the EOF/PC analysis of spatial soil moisture anomalies, we identify one spatial structure (EOF) that explains 61% of the total variance, and find that three such structures explain 87% of the variance. To identify the regional characteristics that are most influential in determining soil moisture patterns, each of these EOFs has been compared to regional characteristics using a correlation analysis. The primary EOF is most highly correlated with the percent sand in the soil. Similar analyses were conducted for wet, average, and dry days, and the role of percent sand is greatest for wet days. As the soil becomes more dry, percent clay becomes more important than percent sand. We have also considered temporal soil moisture anomalies, which identify locations with more or less dynamic soil moisture. The spatial patterns for the temporal anomalies are more complex than those for the spatial anomalies. One EOF is only able to explain 50% of the total variance. Percent sand is also related to the primary EOF for the temporal anomalies, but percent clay becomes unimportant. Topographic characteristics are usually not important over the range of scales we consider, although elevation may play a role in identifying locations with more dynamic soil moisture.

Summer Dawn Conklin
Civil Engineering Department
Colorado State University
Fort Collins, CO 80523
Fall 2005

ACKNOWLEDGMENTS

I gratefully acknowledge financial support from the U.S. Army Research Office Terrestrial Sciences Program.

1 Introduction

Soil moisture is an important variable in hydrology and many related disciplines. In hydrology, soil moisture is a key variable determining runoff production from precipitation events as well as evapotranspiration and groundwater recharge rates between events. Soil moisture is also an important variable in perpetuating droughts and wet periods through precipitation recycling [6]. As a result of these roles in hydrology, soil moisture affects flood warning systems, land management practices, irrigation requirements, etc.

Despite its practical importance, limited soil moisture observations are available. A primary reason for the lack of data is the variability of soil moisture by geographical location, time, and depth in the soil. Rainfall events distribute soil moisture in a highly variable pattern throughout a watershed. Soil moisture is further redistributed during inter-storm periods through evaporation, transpiration, lateral flows, and groundwater recharge. Ground based soil moisture measurement techniques include the use of impedance probes for manual measurements and time domain reflectometers (TDRs) for continuous measurements at variable depths. Manual soil moisture measurements allow for ample spatial coverage over a field, but they are not practical for capturing the variability of soil moisture throughout a soil profile through time. Soil moisture measurements taken by a stationary TDR probes provide an accurate picture of the soil moisture dynamics in the soil column, but they do not provide information about the spatial variability throughout the watershed. Airborne and satellite remotely sensed soil moisture data provides an average near-surface soil moisture value over a 0.64 km^2 to 2500 km^2 footprint. Although this remotely sensed data provides ample spatial coverage, the information is too coarse for most practical applications.

In order to overcome the disconnection between available soil moisture observations and the needs of practical applications, numerous downscaling or interpolation methods have been developed. These methods range from strictly mathematical approaches to methods based on statistical relationships between soil moisture and regional characteristics to physically-based modeling of soil moisture. Mitasova and Mitas [13] present the spline with tension method of interpolation that could be applied to soil moisture. This method is strictly mathematical, but has been applied successfully to terrain modeling and analysis [12]. Kim and Barros proposed a fractal interpolation method based on contraction mapping [9], which is based on the observed scaling invariance of soil moisture. This method generates unique fractal surfaces and includes spatially and temporally varying scaling functions. Pellenq et al. [15] developed a method for downscaling remotely sensed soil moisture data for use in hydrologic modeling. Their model employs a simple soil vegetation atmosphere transfer model coupled with "topmodel" and was able to simulate both near-surface and deep soil moisture. Wilson et al. [20] developed a downscaling method by examining the changing influence of topographic and regional characteristics on soil moisture through cycles of wetting and drying. Prior knowledge of soil moisture in the area allowed the relationships between soil moisture and regional characteristics to be established. These relationships were then exploited to develop a physically-based method for downscaling.

Development of reliable downscaling and interpolation methods requires a sound understanding of the variables that control soil moisture patterns. Because many topographic and regional characteristics are available at a relatively fine scale, an understanding of the dependence of soil moisture on these characteristics could be exploited to develop a method for downscaling. Western et al. [19] investigated the relationship between soil moisture and topographic characteristics. Their research indicated that soil moisture on wet days exhibits a high degree of organization and is best predicted using the natural log of specific upslope area. Conversely, soil moisture on dry days exhibits little spatial organization and is best predicted using the potential solar radiation index. Kim and Barros [8] examined the relationship between soil moisture and regional soil and land-use characteristics. Their results indicated that soil moisture is most strongly connected to soil texture attributes (i.e. percent sand and percent clay).

One way to further analyze the spatial patterns of soil moisture and their connection to regional characteristics is by using Empirical Orthogonal Function and Principal Component (EOF/PC) analysis. EOF/PC analysis is a statistical method that can identify a relatively small number of spatial and temporal patterns that explain a large amount of the variance in a dataset. These patterns then can be correlated with regional characteristics to identify whether these characteristics have a strong influence on the most important tendencies of the soil moisture patterns. Soil moisture data has been examined using EOF/PC analysis at certain scales in the past. EOF/PC analysis was used by Yoo and Kim [22] to analyze patterns in soil moisture at the field scale (over an area of 0.64 km^2). Their research used gravimetric soil moisture data from two areas in the Southern Great Plains field campaign of 1997 (SGP97). By analyzing one month of daily soil moisture readings, they were able to distill patterns (or EOFs) that explain a large amount of the variance. The primary EOF for each field explained over 60% of the variability in that field. These patterns were correlated to elevation, slope, permeability, porosity, and Soil Conservation Service (SCS) curve number. They determined that the two most important EOFs were most highly correlated with elevation at one field site. At the other field site, the most important EOF was most highly correlated with slope and the second most important EOF was most highly correlated with elevation and SCS curve number. By examining the PC time series from their data it was also demonstrated a shifting importance of the underlying patterns through cycles of wetting and drying. At the basin scale (~144 km^2), Verhoest et al. [18] applied EOF/PC analysis to a wintertime sequence of eight raw remotely-sensed synthetic aperture radar images. Through this method, they were able to distinguish soil moisture information from other backscatter information and noise in the satellite images. Kim and Barros [8] used EOF/PC analysis to investigate soil moisture patterns using remotely sensed soil moisture data from SGP97 across ~10,000 km^2. They combined one day of soil moisture data with vegetation water content (VWC), elevation, and percent sand and used EOF/PC analysis to determine the pattern that explains the largest amount of variance between these fields. No matter which day of soil moisture data was selected, the most important EOF was most closely represented by percent sand. However, the second most important pattern was most similar to soil moisture on wet days and most like VWC on dry days. Wittrock and Ripley [21] utilized EOF analysis to examine a 34 year dataset of annual soil moisture measurements in the Canadian prairie. Their data included 23 locations that stretched across the agriculture region of Manitoba. They determined that the most important EOF, which represented 34% of the variance in the dataset, was a response to remote boundary forcing, while the second and third EOFs were responses to regional and local forcing. Liu [10] utilized a coupled version of EOF/PC analysis to analyze monthly to seasonal soil moisture and precipitation variability throughout east Asia. The two fields were generated using the National Center for Atmospheric Research (NCAR) regional climate model. Liu's analysis demonstrated the importance of spatial patterns of soil moisture on precipitation prediction.

In this research, we explore the possibility of using EOF/PC analysis to reduce the complex patterns of soil moisture to a relatively small number of spatial structures that capture a large amount of the variability in the dataset. These spatial structures will be compared to other known regional characteristics to determine if these characteristics are useful predictors of the soil moisture patterns. In the future, the relationship between the EOF patterns and regional characteristics could be exploited for downscaling of remotely-sensed soil moisture data. The relationships between regional characteristics and soil moisture will be evaluated across a range of spatial scales and for both wet and dry conditions. Furthermore, this research will determine if information from the regional characteristics can be distilled into a relatively small number of spatial structures that can be used to forecast or interpolate soil moisture. The following section (Section 2) describes the dataset that we use in this analysis (the SGP97 data), and Section 3 describes the EOF/PC methodology in more detail. Section 4 describes our results at the fine spatial resolution when all days of data are considered. Section 5 discusses how the results change as the spatial scale of observation changes, and Section 6 discusses how the results change between wet and dry days. Section 7 examines the distillation of regional characteristics, and Section 8 summarizes our main conclusions.

2 SGP97 Dataset

The Southern Great Plains 1997 (SGP97) field campaign was a collaboration between National Aeronautics and Space Administration (NASA), United States Department of Agriculture (USDA), other governmental agencies, and several universities. It had the goal of characterizing soil moisture dynamics across a range of scales. From June 18 to July 17, 1997, satellite, airborne, and ground-based measurements were collected for a variety of atmospheric, hydrologic, and surface variables. The SGP97 region stretches from nearly the northern border of Oklahoma to its southern border and includes the Little Washita experimental watershed [1]. The region is dominated by silt loam soils and is primarily agricultural lands or grasslands. Our analysis uses data from the airborne electronically scanned thinned array radar (ESTAR), which indirectly measures moisture over a soil depth of 0-5 cm [7]. ESTAR measurements were taken throughout the one month period along seven flight paths. Problems with instrumentation and weather limited the results to 16 days of soil moisture typically covering an area of about 10,000 km^2. In this analysis, we only consider the sub-region that has data available on all 16 days (9,463 km^2). The data are available at a 800 m grid resolution and were verified through an investigation of sub-pixel variability [5]. During SGP97, rain events occurred between the observations on June 25 and 26, June 29 and 30, and July 10 and 11. The first two rainfall events were centered on the northern half of the region; the third included the entire region but was concentrated in the south [7]. Figure 1 shows the volumetric soil moisture data (total volume of moisture divided by total volume of soil). Before the first rain event (i.e. June 18 – 25), the soil is relatively wet in the northern and central portions of the mapping region and dry in the south. The first storm causes the northern portion to become even wetter (June 25 – 26), but this same region subsequently dries between the first and second storms (June 27 – 29). During the second storm (June 29 – 30), the soil becomes wet again in the north, and after the second storm (July 1 – 3) there is drying throughout the entire region. The third storm (July 10 – 11) causes wetter values throughout the region. Finally, after the third storm (July 12 – 16), there is drying in the north with persistent wetness in the southwest.

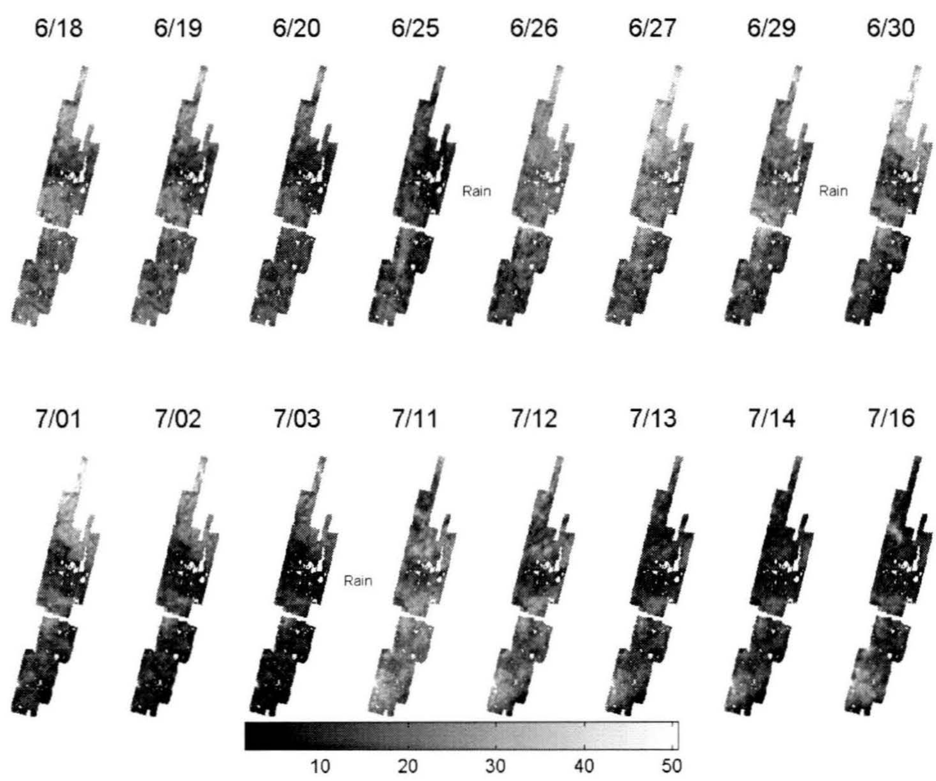

Figure 1: Soil moisture data for the SGP97 study area as collected.

In this analysis, we will examine soil moisture in both spatial and temporal anomaly forms. To calculate the spatial anomalies, the mean soil moisture for each day is subtracted from the local soil moisture values. Thus, the spatial anomalies describe the deviation of the local soil moisture from the regional mean (Figure 2). To calculate the temporal anomalies, the temporal mean soil moisture for each grid cell is subtracted from the soil moisture value for each day (Figure 3). Temporal anomalies do not identify wet or dry locations. Instead, they identify points that are wet or dry relative to their long-term tendency. Patterns observed in the spatial anomaly data are the same as those observed in the original dataset, but the temporal anomaly data shows very different patterns. Whereas the spatial anomaly data are characterized by patterns of above average soil moisture in the north and central regions before the third rain event (July 10-11), the temporal anomaly data initially shows above average soil moisture across the entire region with uniform drying between events, after the third rain event it shows persistent wetness in the central and southern areas.

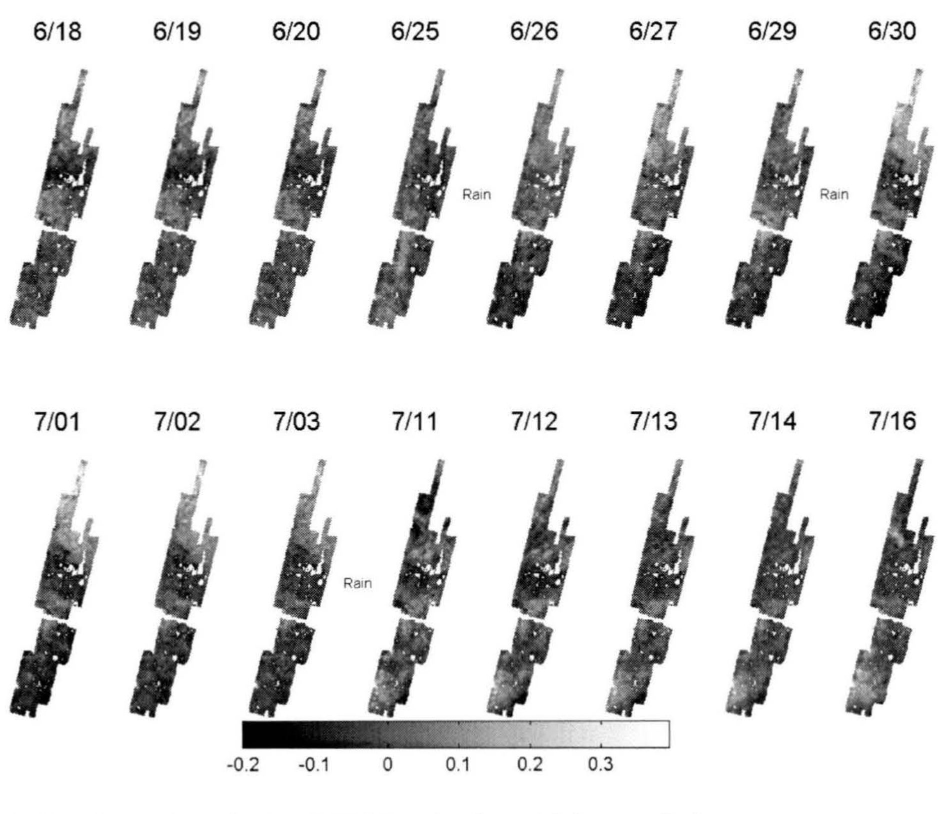

Figure 2: Soil moisture data for the SGP97 region in spatial anomaly form.

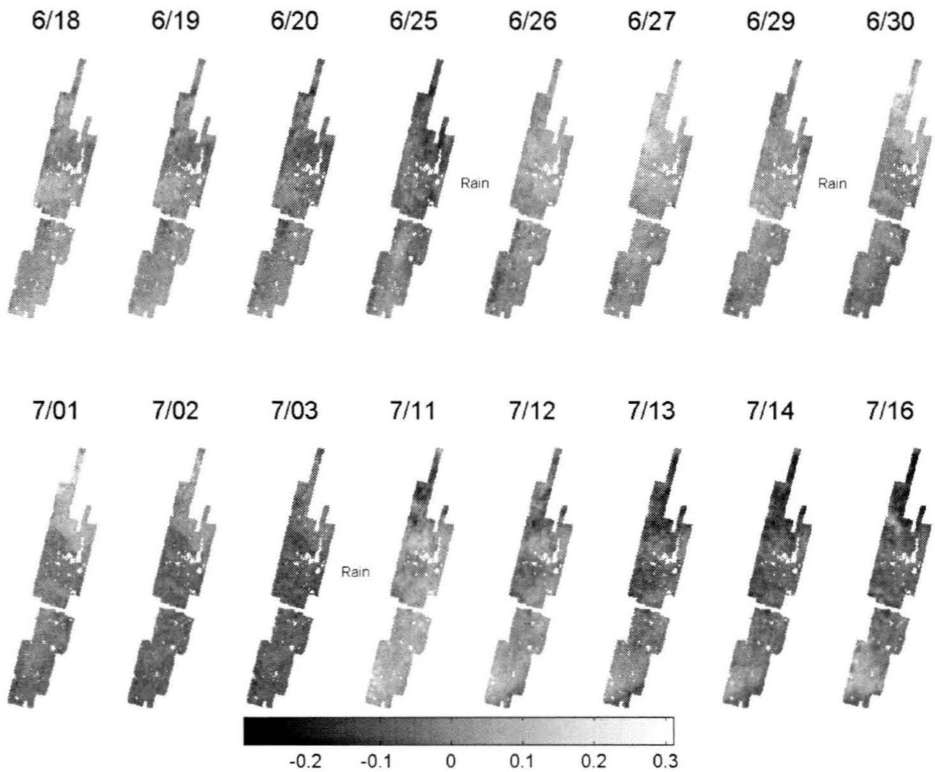

Figure 3: Soil moisture data for the SGP97 region in temporal anomaly form.

Numerous regional characteristics have been observed for the SGP97 site including: percent sand, percent clay, bulk density, vegetation water content (VWC), and surface roughness (Figure 4). These characteristics were used in the procedure to estimate soil moisture from the ESTAR measurements [7]. Percent sand and percent clay were estimated from soil texture categories from the conterminous United States multi-layer soil characteristics dataset (CONUS-SOIL) [11]. Bulk density is the mass of soil and water per cubic centimeter of soil. Sample measurements were taken throughout the SGP97 region, these samples were used to estimate bulk density for each land cover category. VWC represents the mass of water stored in the vegetation per square meter of land area. It was estimated by classifying the vegetation in each grid cell and using the Normalized Difference Vegetation Index (NDVI) to describe the density or "greenness" of the vegetation [7]. The surface roughness describes the variance of the surface elevation relative to the wavelength of the radar used [3]. For SGP97, the surface elevation variance was estimated from land-use information (e.g. bare agricultural areas, fields with corn, urban areas, etc.) [7].

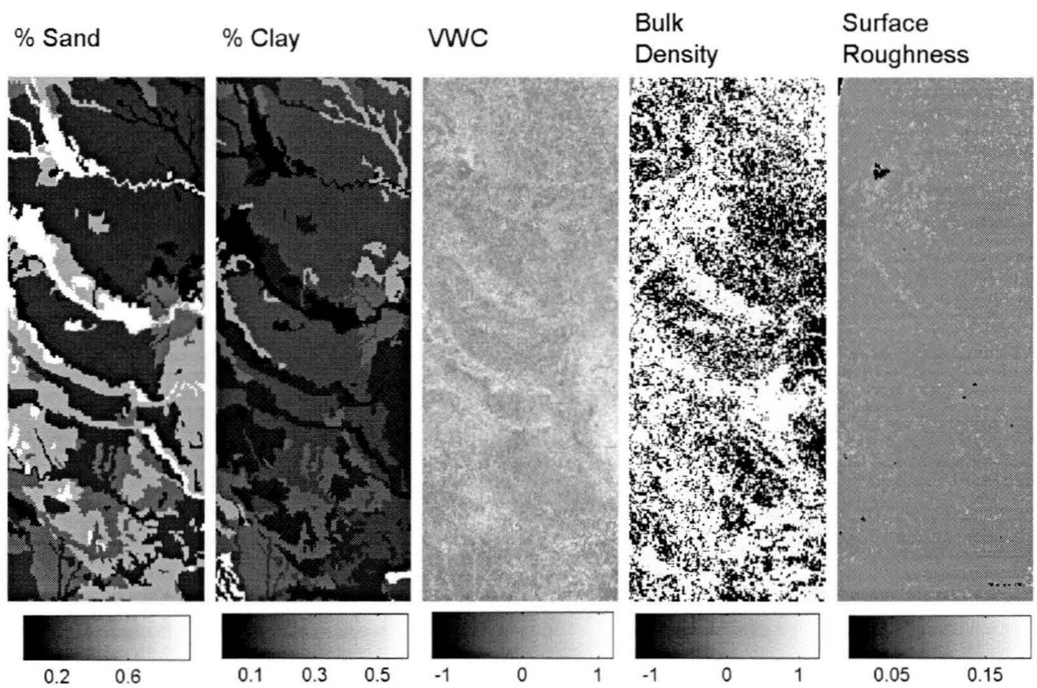

Figure 4: Soil and land cover characteristics of the SGP97 study area.

In addition, we have calculated numerous topographic properties from a Digital Elevation Model (DEM) available from Pennsylvania State University at a 100 m grid resolution. This DEM was translated to UTM zone 14 and trimmed to match the soil moisture data for the SGP97 site. From the elevation data, we calculated: slope, drainage angle, contributing area, wetness index, and curvature (Figure 5). Slope is defined as the steepest downward slope between the grid point being considered and the eight surrounding grid points. Drainage angle is the direction of the steepest downward slope measured in radians in a counterclockwise direction from east. Contributing area is computed as the number of grid cells that would contribute flow to the grid cell being considered in a rain event. In many cases, the area that contributes flow to a point extends well beyond the limits of the SGP97 site. While we extended our topographic analysis well beyond the boundaries of the SGP97 site in an attempt to quantify these areas, some contributing areas could not be determined. Such points are excluded from this analysis and can be seen in Figure 5 as three streams running west to east across the mapping region. Wetness index is defined as the natural log of contributing area per unit contour, which is also called the specific contributing area, divided by the slope [2]. In our analysis, we calculated the specific contributing area by simply dividing the contributing area by the linear size of the grid cell. Curvature is calculated as the rate of change of slope in the x direction plus the rate of change of slope in the y direction.

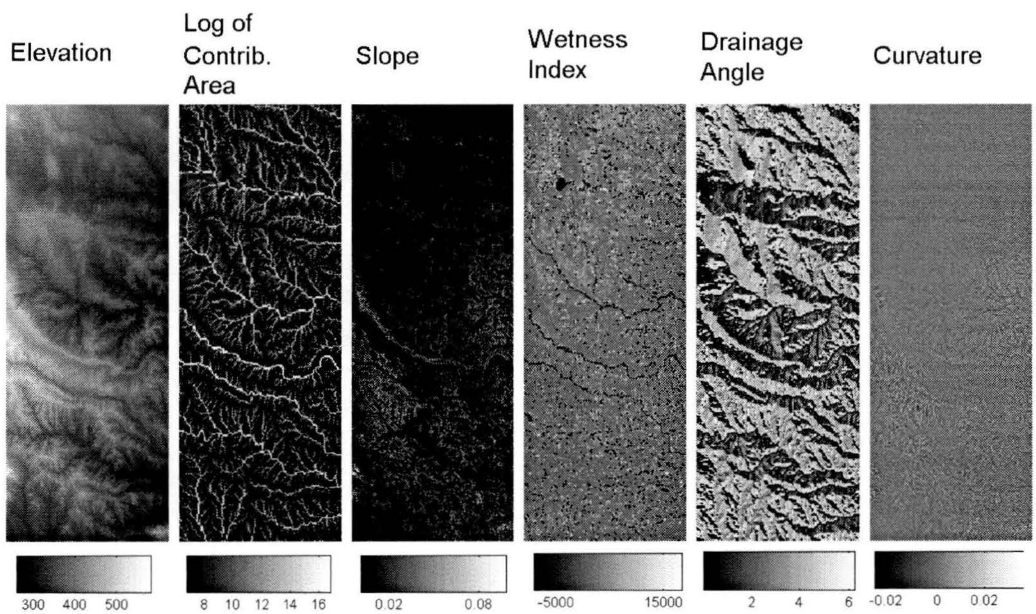

Figure 5: Topographic characteristics of the SGP97 study area.

3 EOF/PC Methodology

Consider a dataset that describes a spatial pattern including n locations, and each location is observed m times. This data is fully described by an m x n matrix. The matrix can alternatively be thought of as n points in m dimensional space, where we have a dimension or coordinate axis associated with each time. For example, consider the simple case with observations at many locations but only two times (Figure 6). Each point in Figure 6a represents the observation at a single location. The x-coordinate indicates the value at one time, and the y-coordinate indicates the value at the second time. Instead of using the values at the two times as our coordinate axes, we could alternatively define two new axes that describe the data more efficiently. For example, the first axis might fall along the main trend of the data. This particular axis would explain the maximum variance possible. If we constrain the second axis to be orthogonal to the first, the second axis would be identified as shown in the figure. If we had more than two dimensions to our problem, we would select each axis in the direction that explains the most remaining variance. Notice that the selection of the new axes amounts to a particular rotation of the data as shown in Figure 6b.

The procedure we have described is a simple example of EOF/PC analysis. The coordinate transformations are called the PCs. In general, a PC has a value for each observation time that indicates how well aligned the transformed axis is with the variability at that time. Thus, PCs are usually time series. PCs are linear functions of the observations and orthogonal to each other. Each PC has a corresponding EOF that represents the projection of the original data onto the PC. Thus, EOFs are usually spatial patterns.

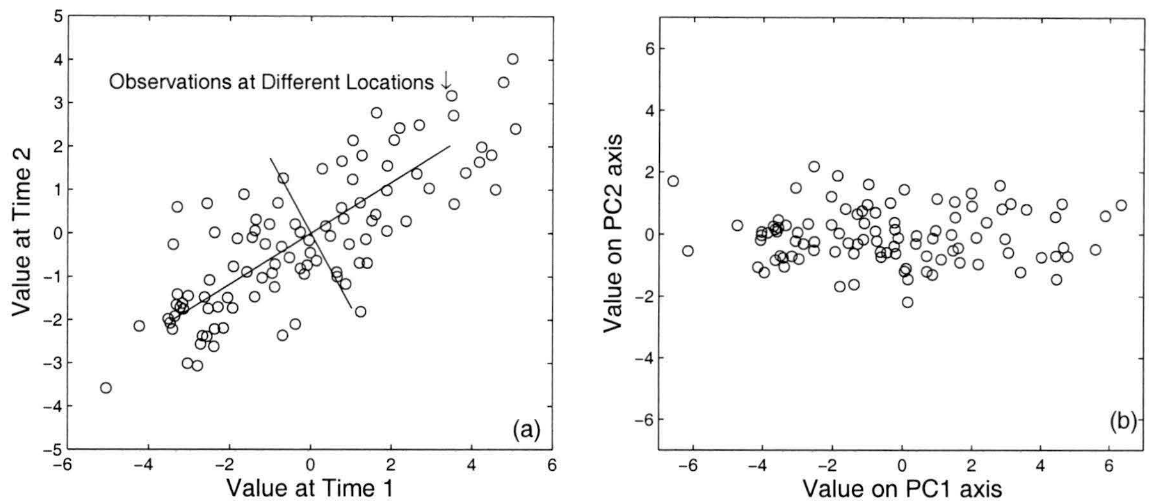

Figure 6: A conceptualization of the EOF/PC coordinate transformation.

In practice, to find the coordinate transformations, the covariance matrix R of the m x n data matrix X must first be found. R (m x m) is calculated:

$$R = 1/n\left(XX^T\right) \tag{1}$$

where the superscript T denotes the transpose of the matrix. The data in the X matrix must be in anomaly form (i.e. have zero mean) in order for Equation (1) to produce the covariance. Either spatial or temporal anomaly form can be used, depending on the variance that is of interest.

One can use eigenanalysis of the covariance matrix to obtain the PCs and EOFs. Eigenanalysis is based on the fact that for most any square symmetrical matrix R:

$$RE = LE \tag{2}$$

where E (m x m) is a matrix whose columns contain the orthogonal PCs and L (m x m) is a diagonal matrix containing the associated eigenvalues. The EOFs can be found by projecting the PC matrix E onto the anomaly data matrix X:

$$Z = E^T X \tag{3}$$

where the EOFs are the rows of the Z (m x n) matrix. An individual observation vector (i.e. all the values measured at a given time) X_{in} can be reconstructed from the PCs and EOFs using:

$$X_{in} = \sum_{j=1}^{m} E_{ij} Z_{jn} \,. \tag{4}$$

The variance explained by the EOF/PC pair is equal to its corresponding eigenvalue divided by the sum of all the eigenvalues. EOF/PC pairs are usually ranked according to the amount of variance that they explain. The primary EOF/PC pair explains the largest amount of variance, the second EOF/PC pair explains the second most, and so on. A test for the significance of each EOF/PC pair based on the number of independent observations at each location was established by North et al. [14]. In this method, error bars are established for the variance explained by each EOF/PC pair. An EOF/PC pair is considered to be significant when there is no overlap of its error bars with those of the previous EOF/PC pair. Error bars extend a distance Δl_i above and below the variance explained (l_i) which can be found from:

$$\Delta l_i = 0.5 l_i (2/m_*)^{0.5} \tag{5}$$

where m_* is the number of independent observations.

In the previous discussion of EOF/PC analysis, we have considered a dataset that includes observations of the same variable through time (m repetitions at n locations). It should be noted that EOF/PC analysis can also be conducted with a data matrix containing observations of m different variables at the same n locations. In such case, the spatial mean and standard deviation should be removed from each characteristic so that they can be analyzed together. This alternative analysis finds patterns that efficiently explain the variability of all of the included variables. We will use this alternative approach in Section 7 when we attempt to efficiently describe the patterns of the regional characteristics.

4 Results at the Fine Spatial Resolution

4.1 Spatial Anomalies

Analysis of the spatial anomaly soil moisture data yields a series of sixteen EOF/PC pairs. The spatial structures of the first five EOFs are displayed in Figure 7. The primary EOF shows a pattern with above average values in the northern and central portions of the study area. The southern portion of this EOF has primarily low values. Recalling the discussion of the daily soil moisture patterns in Section 2, EOF1 resembles the soil moisture patterns that occur through July 3. The second EOF shows a very different pattern that is characterized by low values throughout the study region with one cluster of high values in the southwestern corner. This pattern is consistent with the soil moisture observed after the third rainfall event (July 10-11). The third EOF is highly variable throughout the region. This EOF highlights the above average soil moisture values that were observed in the central portion of the region on nearly all days.

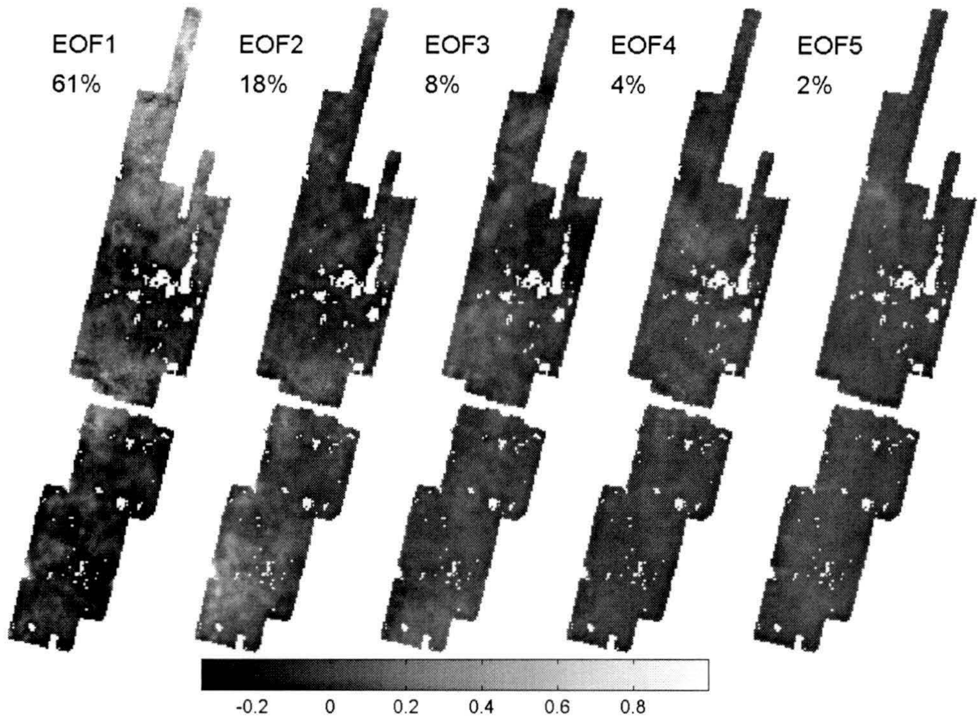

Figure 7: Significant EOFs of spatial anomaly soil moisture data.

The eigenvalues associated with the EOF/PC pairs indicate that a large amount of the total variability in this dataset can be explained through relatively few spatial patterns (Figure 8). Because there are 16 days of observation, there are a total of 16 EOF/PC pairs that together explain 100% of the variability. However, EOF1 alone explains 61% of the total variance, and together the first three EOFs explain 87% of the variance. Using the significance criteria from North et al. [14] and assuming that each day of observation is independent, the first five EOFs are considered significant and explain 93% of the total variance (Figure 8). If the daily soil moisture patterns have significant interdependence, the number of significant EOFs would be reduced. These results indicate that the seemingly complex patterns of soil moisture in the SGP97 field campaign can largely be explained by a very small number of underlying spatial structures. Specifically, the fact that 61% of the total variability has been captured in EOF1 indicates that a single spatial structure, which is invariant in time, can explain much of the overall soil moisture pattern. Although the relative importance of EOF1 on daily patterns of soil moisture waxes and wanes during cycles of wetting and drying (as discussed later in this section), the spatial pattern of EOF1 is consistent for all days of data. Yoo and Kim [22] also found that a single spatial structure, EOF1, could explain 60 – 65% of the variability in their 25 day dataset (depending on the field being analyzed). However, their subsequent spatial structures (EOF2, EOF3, EOF4, ect.) each explained less than 10% of the total variability.

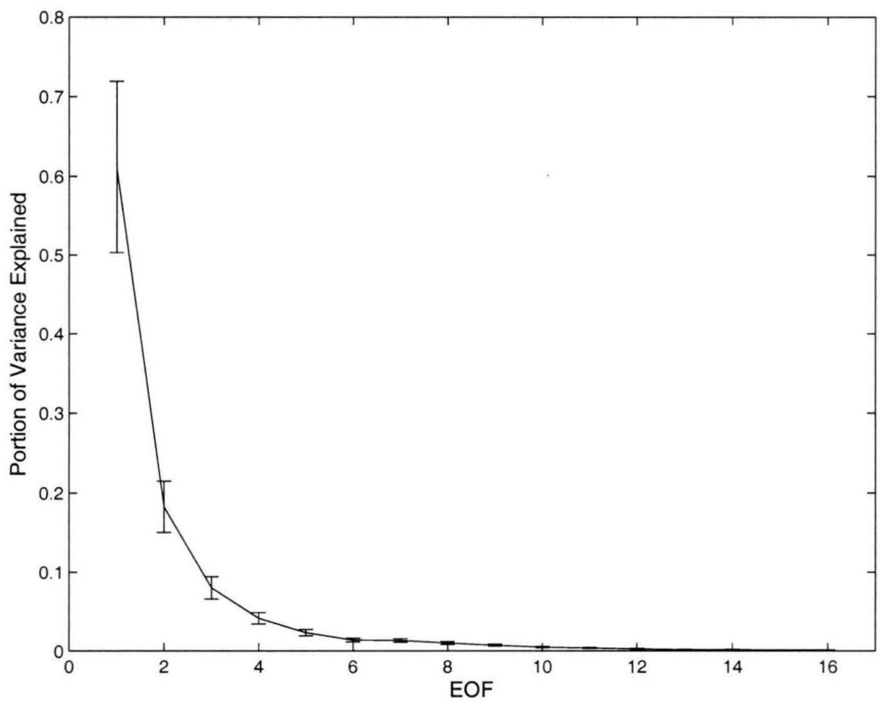

Figure 8: The portion of the total variance of the spatial anomalies of soil moisture that is explained by each EOF along with error bars to judge the significance of the EOFs.

Multiplying the PCs by the amount of the total variance they explain gives the weighted PCs, which indicate the relative importance of each EOF in describing the variance in each day's soil moisture pattern (Figure 9). The weighted PCs show clear cycles of wetting and drying, which can be related to the occurrence of precipitation events. During the first several days, when the above average soil moisture values are concentrated in the northern and central portions of the region, EOF 1 is dominant. As the soil dries before the first rain event, the first three EOFs become almost equally important in predicting soil moisture. After the first and second rain events, EOF1 becomes dominant again, and shows a decrease in its relative importance on predicting soil moisture as each period of drying progresses. After the third rain event (July 10-11), when the above average soil moisture values are in the southwestern corner of the data, EOF2 becomes the most important pattern in determining soil moisture while the importance of EOF1 and EOF3 nears zero. Yoo and Kim [22] also considered the weighted PCs of their soil moisture data. Their results showed small shifts in the relative importance of the EOFs during cycles of wetting and drying. However, their results did not show any shifts in the dominant EOF, as observed for EOF2 after the third rain event in this analysis. The consistency of the PCs with periods of wetting and drying is important because it suggests that the EOF/PC rotation of the data captures some physically-based tendencies in the underlying data.

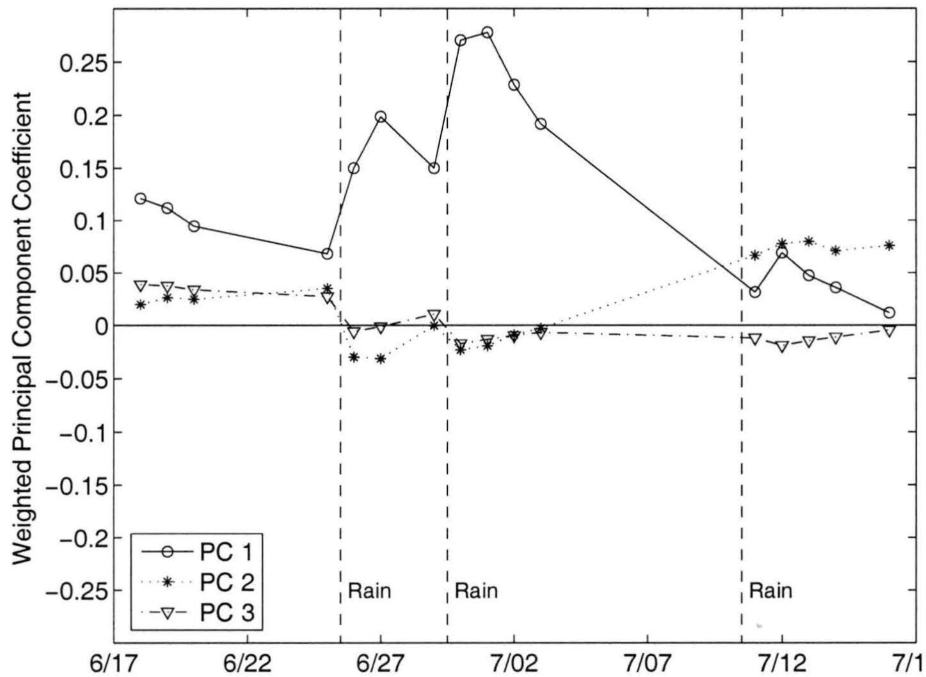

Figure 9: Variation of the weighted PCs through time. These PCs are associated with the EOFs in Figure 7.

In many previous EOF/PC analyses, the first EOF is dominated by one positive anomaly over the entire domain, and the second EOF has positive and negative anomalies occurring in north-south or east-west pairs. Richman [16] suggests that such EOFs are largely products of the domain shape. The EOFs derived from the SGP97 soil moisture data do not conform to these common patterns. The first EOF is dominated by a positive anomaly in the northern and central regions rather than a single positive anomaly throughout. The second EOF has a positive anomaly in the south and a negative anomaly in the north, which is more consistent with typical EOFs. We evaluated the possible effects of the domain shape by dividing the study area into northern and southern halves. These halves were then analyzed individually by recalculating the spatial anomalies and repeating the EOF/PC analysis. The resulting EOFs for the sub-regions are very similar to those found when analyzing the whole domain (Figure 10). The variability explained by each EOF does change. For the northern region, EOF1 explains 73% of the variance, EOF2 explains 10%, and EOF3 explains 6%. For the southern region, EOF1 explains 54% of the variance, EOF2 explains 26% and EOF3 explains 6%. These changes are expected because the first two rain events were concentrated in the northern region, and EOF1 is most associated with the soil moisture patterns produced by that event. Overall, however, this simple analysis shows that the domain shape has little effect on the spatial patterns identified by the EOF/PC analysis in this case.

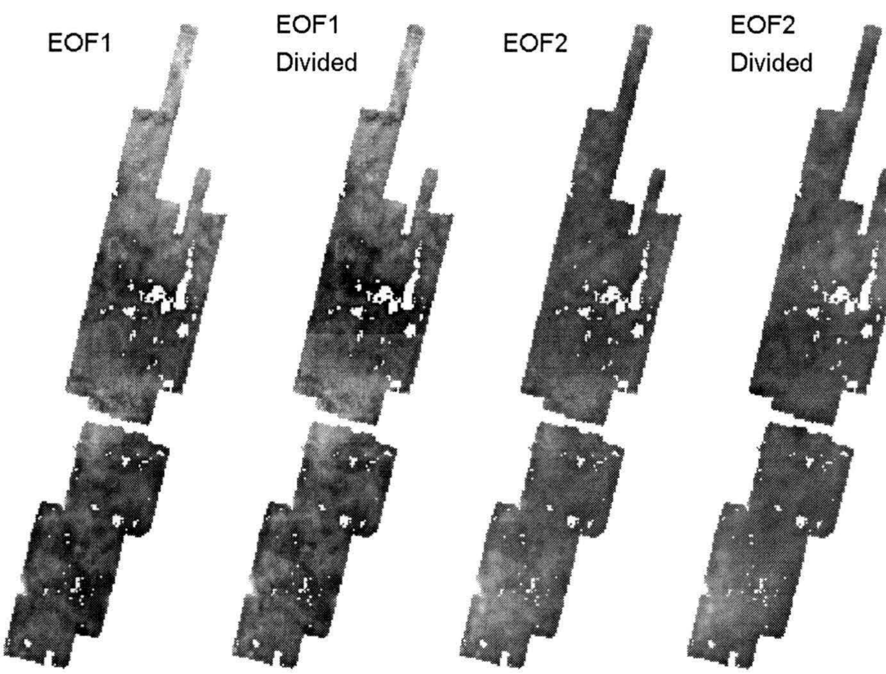

Figure 10: Comparison of EOFs generated from full dataset, and EOFs generated from data divided into northern and southern sections.

One of the main benefits of the EOF/PC analysis is that we have now identified a small number of orthogonal spatial patterns (e.g. EOF1, EOF2, and EOF3) that together explain a large portion of the total variability of soil moisture. We now examine how closely these underlying patterns resemble regional characteristics that might influence the soil moisture. For this analysis, we calculate the correlation coefficient r between the EOFs and the available regional characteristics. The correlation coefficient can be calculated:

$$r = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(A_{ij}-\overline{A}\right)\left(B_{ij}-\overline{B}\right)}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left(A_{ij}-\overline{A}\right)^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left(B_{ij}-\overline{B}\right)^2\right)}} \tag{6}$$

where A and B are two generic matrices and $\overline{A}$ and $\overline{B}$ are their associated means.

The results of the correlation analysis between the first five EOFs and the available regional and topographic characteristics are shown in Table 1. We recognize that some of the characteristics being correlated (percent sand, percent clay, vegetation water content, surface roughness and bulk density) were used to derive soil moisture from the ESTAR measurements. However, because this analysis is considering the patterns of variability, not soil moisture directly, the correlation will not be a direct result of the algorithm used to find soil moisture. While none of the correlations are very high, EOF1 has a relatively high correlation with percent sand, and moderate correlations with percent clay and slope. This indicates that some of the variability of the EOF1 pattern is related to soil type and slope. This is an intuitive result because percent sand and clay when coupled with slope determine how quickly a soil will drain after a rain event. Very sandy areas drain quickly and therefore tend to be drier than other locations. Conversely, very clayey areas tend to hold water and be wetter than other locations. This result is also consistent with the plot of the weighted PCs in Figure 9, which shows that the influence of EOF1 waned as the soil dried. The second EOF, which has a cluster of above average soil moisture values in the southwestern corner, is most correlated with elevation, and moderately correlated with percent clay and vegetation water content. It is difficult to determine a physical process that would produce higher soil moisture values at higher elevations. Recall that EOF2 is most like the pattern of soil moisture after the third rain event (July 10-11,) which was concentrated in the south. We believe that the correlation of EOF2 with elevation is non-physical and occurs simply because the rainfall happened to occur in an area of high elevation. For the subsequent EOFs, the correlations drop off quickly, but percent sand, percent clay, vegetation water content, and elevation usually remain the most influential characteristics.

A rank correlation analysis was also performed between the EOFs of soil moisture and the regional characteristics. The rank correlation analysis sorts the values in two selected datasets from their highest to lowest values and examines whether a relationship is present between the two rankings assigned to the data for each location in the region. This approach does not rely on the Gaussian assumption that underlies the simple correlation analysis. The results of the rank correlation analysis showed little difference from the results of the simple correlation analysis.

Table 1: Correlations between regional and topographic characteristics and the EOFs of the spatial anomaly soil moisture data.

	EOF1	EOF2	EOF3	EOF4	EOF5
% Sand	-0.42	0.02	-0.09	-0.23	0.09
% Clay	0.22	0.20	-0.01	0.10	-0.17
Vegetation Water Content	-0.07	0.24	-0.12	-0.02	0.07
Bulk Density	-0.15	0.12	0.05	-0.04	0.10
Surface Roughness	0.12	0.04	-0.06	0.01	0.01
Elevation	-0.14	0.47	0.21	-0.11	0.14
Contributing Area	0.04	-0.01	-0.02	-0.05	-0.02
Slope	-0.21	0.12	-0.03	-0.11	0.01
Wetness Index	0.11	-0.06	0.01	-0.01	0.02
Drainage Angle	-0.02	0.04	-0.07	-0.11	0.07
Curvature	0.00	-0.01	0.00	0.00	-0.01

4.2 Temporal Anomalies

The EOF/PC analysis was also performed on soil moisture data in temporal anomaly form. The analysis of temporal anomaly data tends to highlight areas with dynamic soil moisture. For example, a location that is always wet relative to other locations has no temporal anomaly. But a location that transitions from wet to dry (very dynamic) will have significant temporal anomalies.

The analysis of temporal anomaly data produced very different results than those for the spatial anomalies. As many as five EOFs may be significant, and together these five EOF/PC pairs explain 98% of the total variance (Figure 11). The primary EOF, which explains 50% of the total temporal variability, highlights the soil moisture pattern observed between July 12 and 16. This EOF has above average values in the southwestern corner of the region, and below average values in the northern portion of the region. This pattern is similar to, but more exaggerated than EOF2 of the spatial anomaly data. The second temporal anomaly EOF, which explains 28% of the variance, has an above average anomaly over most of the mapping region, but the anomaly is most pronounced in the northern portion. This pattern is most consistent with the first and second rainfall events and their subsequent drying patterns. The third EOF explains 9% of the total variance and is dominated by a strong above average signal over the entire study area. The differences between the spatial and temporal anomaly EOFs demonstrates that the most important patterns in determining temporal variability are different from the most important patterns in determining the spatial variability. Also, the variance is distributed differently among the temporal anomaly EOFs than the spatial anomaly EOFs. Specifically, the first spatial anomaly EOF can explain 63% of the total variance while the first temporal anomaly can explain only 50%. This result suggests that a more complicated pattern underlies the temporal anomalies than the spatial anomalies.

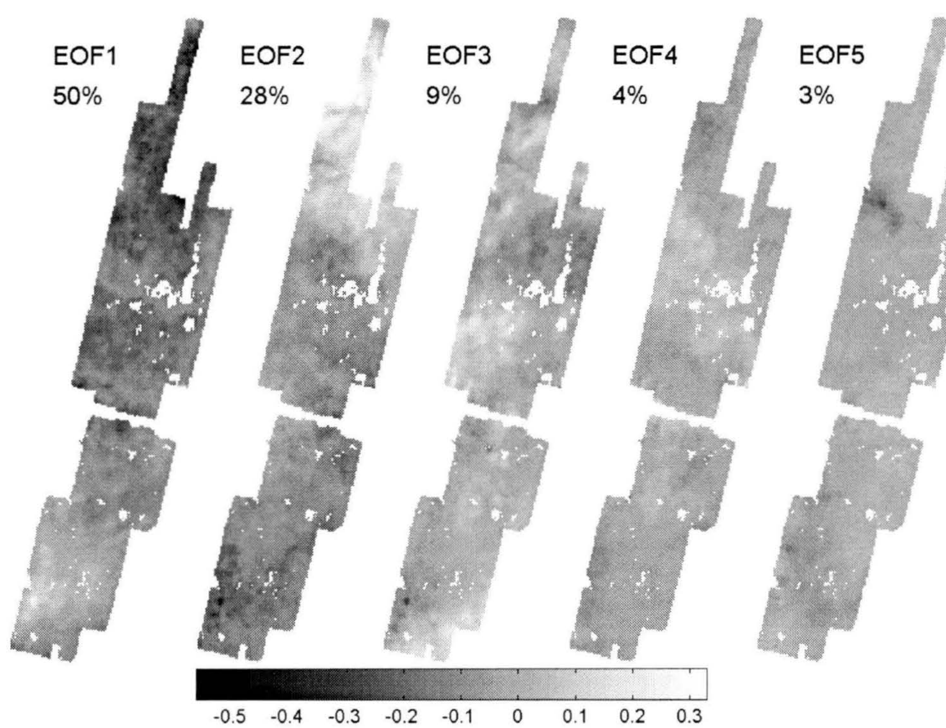

Figure 11: Significant EOFs generated from temporal anomaly data.

To determine how the patterns of temporal variability are related to physical characteristics and processes, the temporal anomaly EOFs were correlated with the regional characteristics (Table 2). EOF1 is most highly correlated with elevation and percent sand, but this EOF shows relatively high correlations with several other characteristics. Notice that percent clay has very little correlation with EOF1 for the temporal anomalies (it is correlated with EOF1 for the spatial anomalies). Locations with values near zero in the temporal anomaly EOFs tend to be locations with little variability during the 16 day dataset. In contrast, locations with high or low values in these EOFs usually correspond to high variability sites. Thus, the high correlation between EOF1 and percent sand suggests that sandy locations tend to have more dynamic soil moisture. These points may be more dynamic because sandy soils can quickly drain water received from precipitation events. In contrast, clayey sites are not correlated with more dynamic locations because clay tends to retain water during the period of observation. Locations with steeper slopes and higher elevations also tend to be more dynamic. Steeper slopes may promote drainage and produce higher temporal variability, and higher elevation points may be more disconnected from stable sources of moisture such as rivers or regional aquifers and thus be more dynamic. Alternatively, higher elevation points may be more dynamic simply because the third precipitation event occurred over a region of higher elevations.

Table 2: Correlation of temporal anomaly EOFs with regional and topographic characteristics.

	EOF1	EOF2	EOF3	EOF4	EOF5
% Sand	0.39	-0.23	-0.02	0.06	-0.09
% Clay	-0.06	0.01	-0.09	-0.18	0.13
Vegetation Water Content	0.25	-0.15	-0.15	-0.02	-0.04
Bulk Density	0.21	-0.17	0.04	0.02	-0.08
Surface Roughness	-0.06	0.08	-0.07	-0.03	-0.01
Elevation	0.44	-0.35	0.12	-0.21	-0.15
Contributing Area	-0.01	0.05	-0.01	-0.04	0.01
Slope	0.25	-0.19	-0.02	-0.06	-0.05
Wetness Index	-0.11	0.12	0.01	0.00	-0.02
Drainage Angle	0.07	0.01	-0.06	-0.04	-0.10
Curvature	0.00	0.00	0.00	0.00	0.01

5 Results At Different Spatial Resolutions

Inherent to the question of what regional and topographic characteristics influence soil moisture variability is the question: "at what scales are the different characteristics important?" It is possible that the observed patterns of soil moisture, and the subsequent EOF patterns, are imbedding information from different regional characteristics at different scales. Research has shown that the relationship of mean soil moisture to the variability of the sample depends on the scale being observed [4]. Additionally, the variance and spatial correlation of soil moisture have been shown to follow a power law decay function, which is an indication of scaling processes [17]. To investigate the influence of regional and topographic characteristics on patterns of soil moisture at different scales, the EOF/PC and correlation analyses were repeated at different spatial resolutions with the SGP97 data. The soil moisture data were aggregated by identifying new grid cell sizes and averaging the available data within the boundaries of each cell. This aggregation technique allows us to observe soil moisture from a finest resolution of 0.64 km^2 (the original data) to a final grid cell size of 256 km^2. The largest cell size averages data from 400 pixels (20 by 20) to determine the new soil moisture value. Our aggregation scheme is similar to Kim and Barros [8] because we include some rectangular grid cells to gain more levels of aggregation. After completing the aggregation procedure, the EOF/PC analysis was repeated for each level of aggregation. The regional characteristics, percent sand, percent clay, VWC, bulk density, surface roughness and elevation were aggregated in the same way. It is important to note that the signs determined for the EOFs and PCs in the eigenanalysis are arbitrary (i.e. the positive direction for an axis can be defined in either direction). Thus, the signs of the EOFs and PCs have the potential to flip at different scales. Therefore, our analysis in this section considers only the absolute values of the correlations in order to simplify comparisons among different scales.

5.1 Spatial Anomalies

The results of the correlation analysis for the spatial anomaly EOFs and the regional characteristics across scales are shown in Figure 12. The correlation of EOF1 and percent sand increases with increasing scale. This tendency indicates that the occurrence of a large scale dry area is more dependent on a large area with above average sand content than a small scale dry area is dependent on a small area with above average sand content. A location may be dry relative to its neighbors for any number of reasons including a steep slope that promotes lateral drainage, that fact that precipitation did not occur at that location, or the evapotranspiration processes is especially efficient due to certain land cover conditions. However, the results of this analysis indicate that large dry regions tend to require large sandy regions. While the correlation of land-cover dependent characteristics with EOF1 does increase with scale (VWC and surface roughness patterns overtake percent clay), they remain secondary to soil properties (percent sand and bulk density). For EOF2, the correlation with elevation and VWC drop slightly at large scales, while the correlations of percent clay and bulk density become larger. Surface roughness and percent sand remain essentially uncorrelated with EOF2 at all scales.

The correlation analysis of EOF1 across a range of scales also provides some insight into the role of lateral fluxes. It is possible that the correlations observed for slope and percent sand occur because they promote lateral fluxes of water from one location to an adjoining location in the dataset. Given the large spatial extent of SGP97, lateral fluxes would be confined to relatively short distances within the region so the dependence of EOF1 on percent sand would be expected to drop off at increasing scales. However, this is not the case. Thus, the dependence on percent sand is probably not associated with the movement of water between adjacent pixels in our dataset.

The contrast of pattern correlations found by Yoo and Kim [22], and those presented in this paper, further solidifies the proposition that different regional characteristics are influencing soil moisture at different scales. The correlation of EOFs from the Little Washita SGP97 data indicated that topographic characteristics were the primary source of variability within their 0.64 km^2 region. For the first field they considered, EOF1 and EOF2 were most highly correlated with elevation (0.47 and 0.44 respectively). For the second field they considered, EOF1 was most correlated with slope while EOF2 was most correlated with elevation (0.59 and 0.762 respectively).

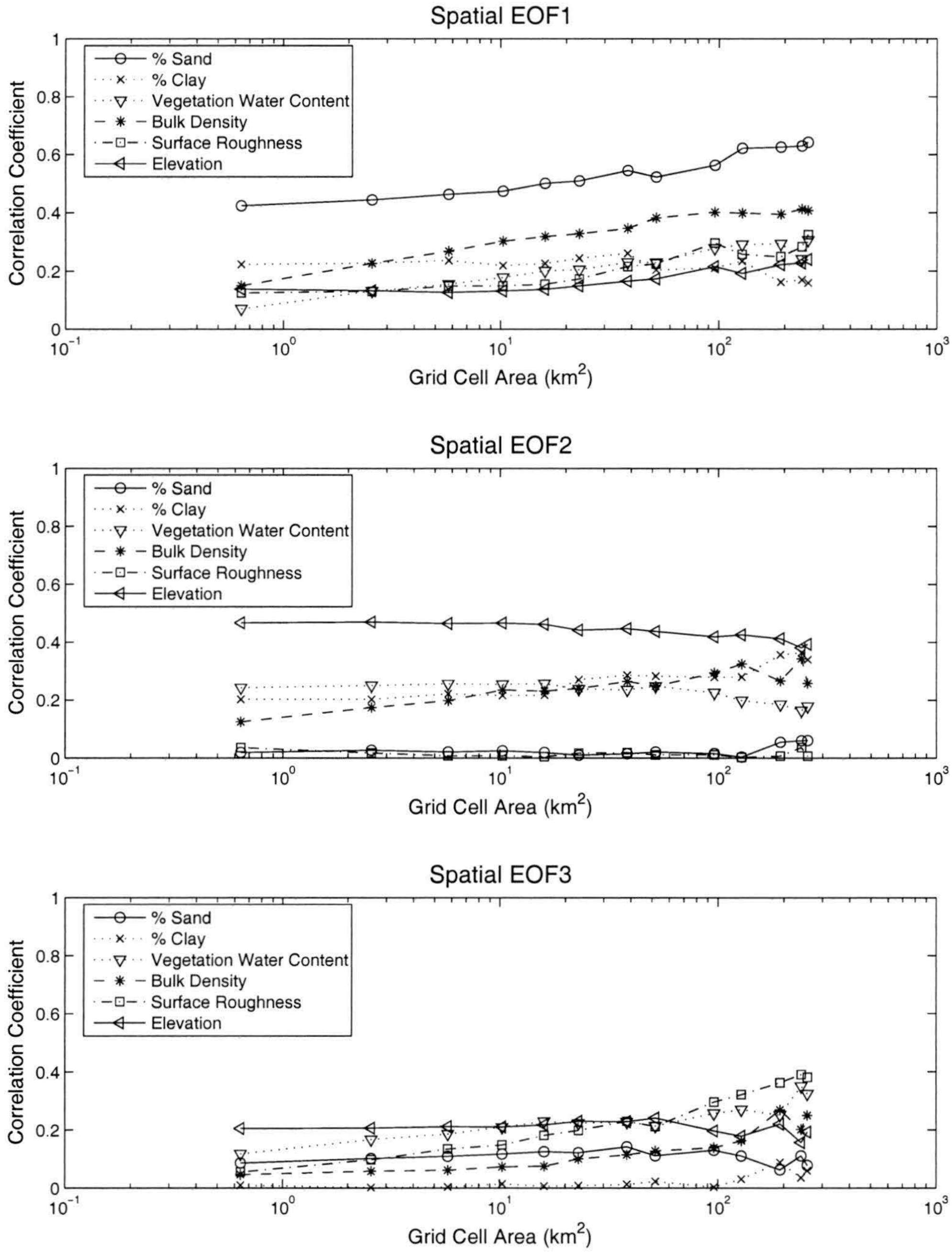

Figure 12: Correlation of spatial anomaly EOFs and regional characteristics across a range of scales.

5.2 Temporal anomalies

To determine what features control large scale dynamic areas, the correlation analysis across scales was also completed for the temporal anomaly data. As shown in Figure 13, temporal anomaly EOF1 shows a relative increase in the importance of bulk density and percent sand at large scales. At the largest scales, they are more important than elevation. Surface roughness is uncorrelated at small and medium scales, but has an increasing effect at large scales. Temporal EOF2 also shows an increase in the relative roles of bulk density, percent sand and surface roughness. Temporal EOF3 however shows the most dramatic increase in the role of surface roughness at large scales, while VWC remains important.

The shifting of the regional characteristic most correlated with EOF1 is an important result. This indicates that the EOF explaining the largest amount of the variance is most dependent on different characteristics at different scales. At small scales, high local elevations are relatively good indicators of locations with dynamic soil moisture, perhaps because these locations are associated with ridge tops. However, regional average elevation values are not as reliable as regional soil characteristics such as percent sand in determining locations with dynamic soil moisture.

A test for the significance of the correlations was done across the range of scales considered. All of the results we have highlighted are significant. At the fine scale, all correlations above 0.02 are significantly different from 0 with a 95% confidence limit. At the largest scale, all correlations above 0.24 are significantly different from 0 with a 95% confidence limit.

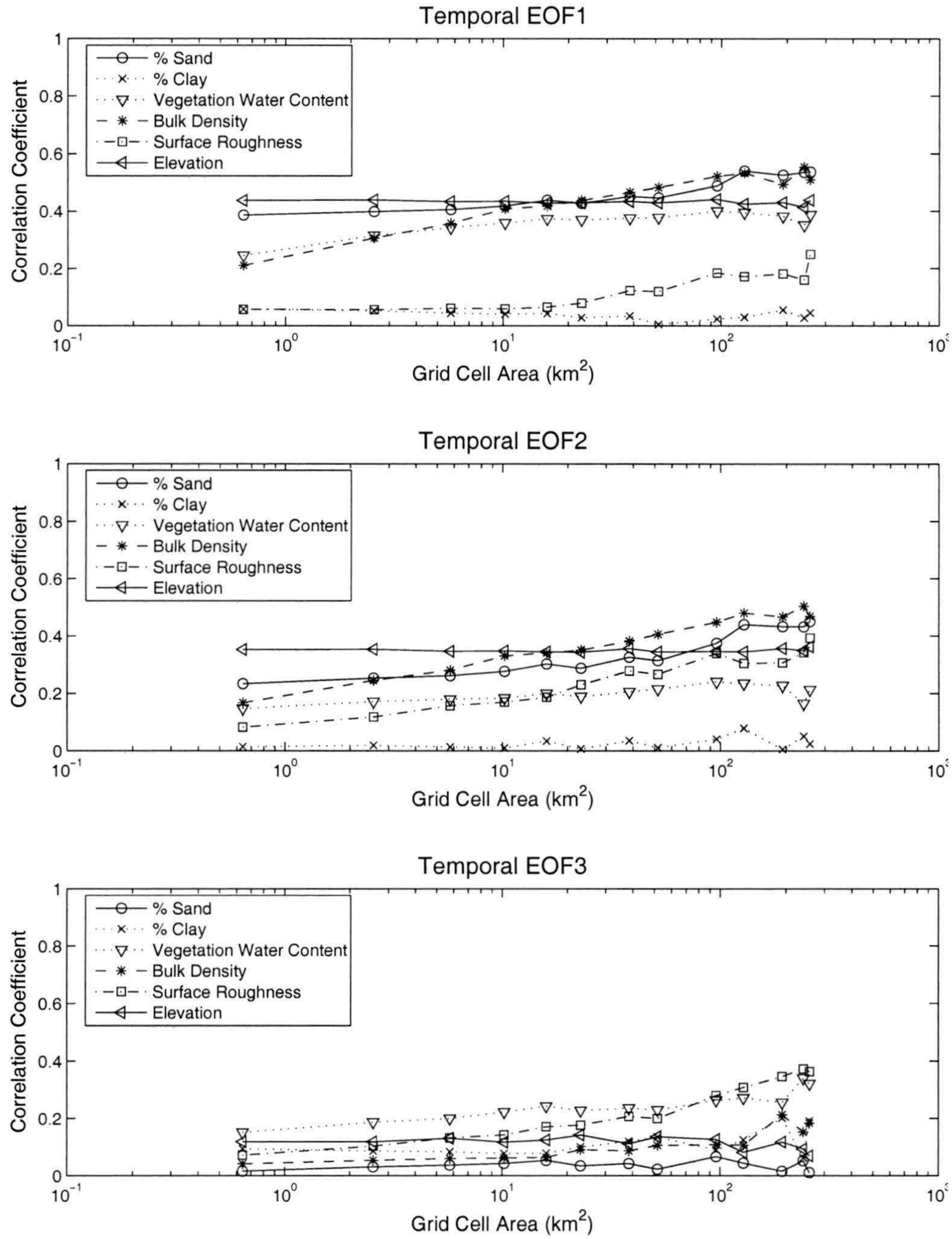

Figure 13: Correlation of temporal EOFs with regional characteristics across a range of scales.

6 Results for Wet, Average, and Dry Periods

It is also possible that different processes are important in the vadose zone water balance depending on the overall level of moisture. This could lead to different physical characteristics controlling the soil moisture patterns for wet and dry days. In order to assess this possibility, the soil moisture data were divided into wet, average, and dry days. To classify the data, the soil moisture measurements from all 16 days were combined, and the mean and standard deviation of the combined dataset were calculated (18.5% and 10%, respectively). The mean for each day was then compared with the mean of the entire dataset to determine whether the day was wet, average, or dry. If the daily mean is one-quarter standard deviation above or below the overall mean, then the day is labeled wet or dry, respectively. The remaining days are considered average (Figure 14).

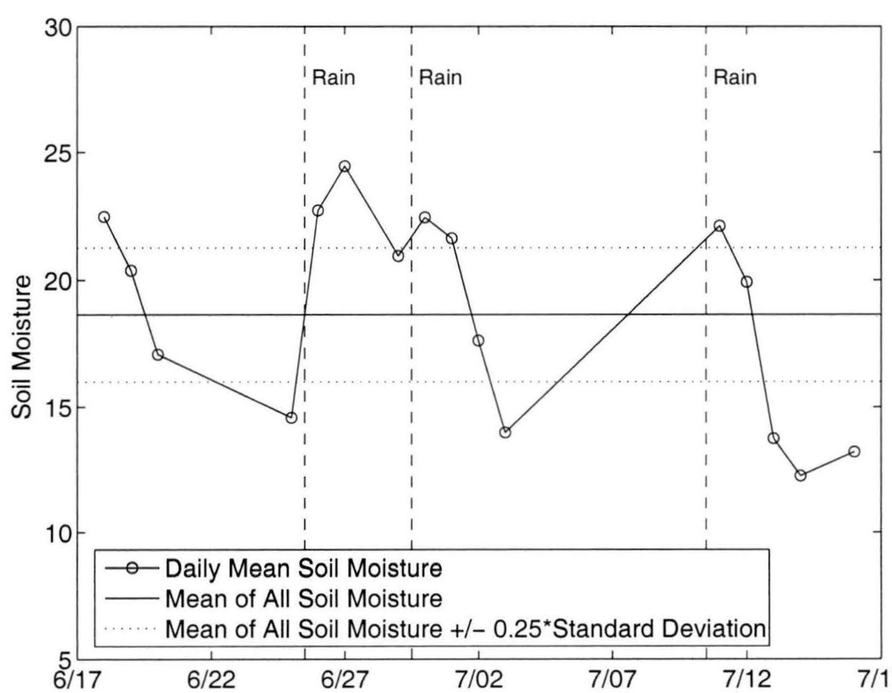

Figure 14: Daily mean soil moisture and the days classified as wet, average and dry.

The spatial anomaly EOF/PC analysis was repeated for the data in each of these categories, and the primary EOF for each category is shown in Figure 15. All three primary EOFs are reminiscent of the primary EOF for the entire dataset because they have a positive anomaly in the northern and central regions. However, as one considers drier conditions, the area of above average wetness in the southwestern corner becomes more pronounced. Correlation analysis of these primary EOFs at the fine scale produces interesting results, which are shown in Table 3. Most important is the role of percent sand and percent clay between wet, average, and dry days. Percent sand is most correlated with EOF1 on wet days, and the correlation drops as the sample becomes drier. This behavior occurs because percent sand is related to the ability of a soil to drain quickly. Thus, it is most important shortly after a rain event. Conversely, percent clay is the most correlated with EOF1 on dry days, and the correlation drops as the sample becomes wetter. Percent clay is an indication of soil moisture holding capacity, and will therefore identify areas of above average soil moisture long after a rain event.

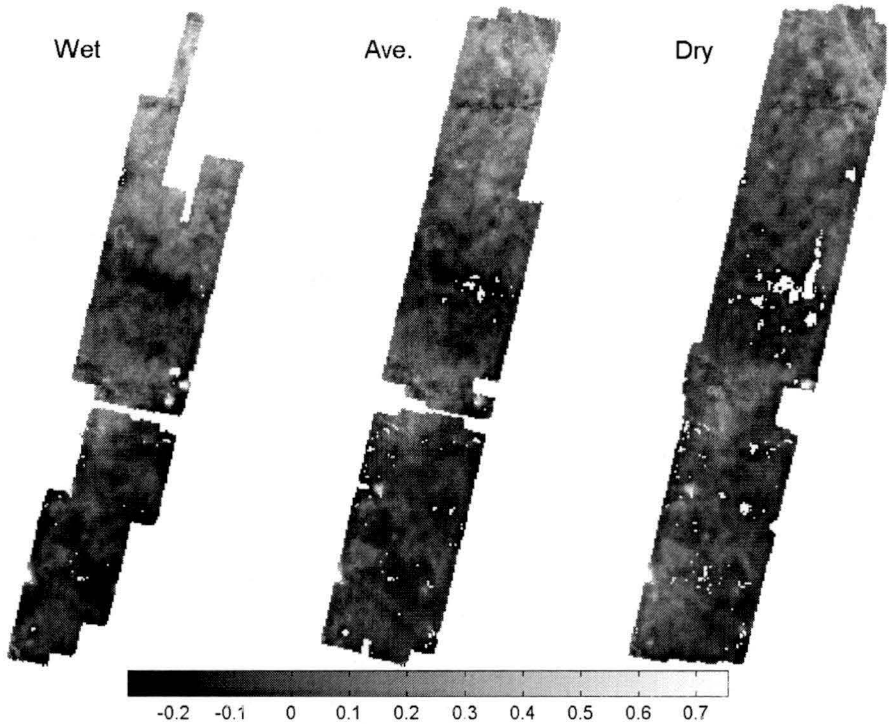

Figure 15: EOFs for soil moisture data divided along temporal scales.

Table 3: Correlation of the primary EOF of wet, average and dry days with regional and topographic characteristics.

	Wet	Average	Dry
% Sand	-0.44	-0.43	-0.32
% Clay	0.22	0.27	0.34
Vegetation Water Content	-0.10	-0.08	0.04
Bulk Density	-0.18	-0.10	0.03
Surface Roughness	0.11	0.15	0.19
Elevation	-0.23	-0.19	-0.11
Contributing Area	0.03	0.05	0.06
Slope	-0.22	-0.25	-0.17
Wetness Index	-0.19	-0.21	-0.14
Drainage Angle	-0.03	0.01	0.06
Curvature	-0.01	0.00	-0.01

The correlations between the topographic features and the primary EOFs also exhibit some interesting tendencies. Elevation is the most important topographic characteristic in wet periods, and slope is the most important in average periods. The fact that elevation becomes less important as the soil dries confirms our interpretation that elevation is important simply because it rained in a region with high elevations (Section 4.1). Surface roughness shows an increase in correlation as the sample becomes drier. None of the topographic characteristics are correlated above 20% with the dry period. This is consistent with the idea that topography is more important in determining soil moisture during wet periods than during dry periods [8].

The EOF/PC analysis of wet, average, and dry days was repeated across spatial scales (Figure 16). For wet and average periods, percent sand remains the most important variable at all scales. For the dry periods, percent sand and surface roughness become more important than percent clay at large scales. For all three divisions of data, the relative role of elevation becomes relatively less important at large scales.

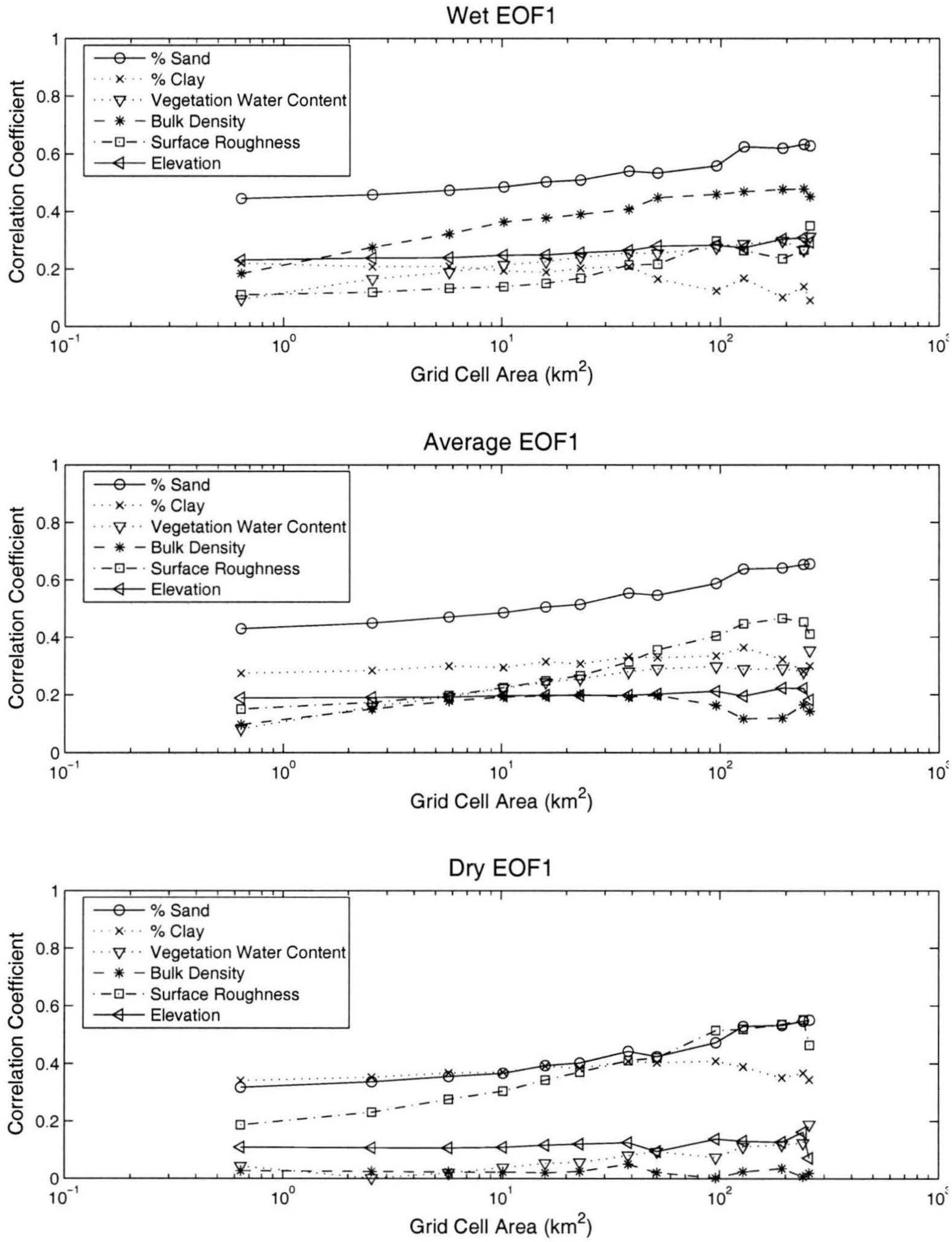

Figure 16: Correlation of EOFs divided into wet, average and dry with regional characteristics across spatial scales.

7 EOF Analysis of Regional Characteristics

The EOF/PC analyses in the previous sections have shown that the observed soil moisture patterns can be efficiently approximated by a small number of EOFs. Furthermore, these EOFs have varying degrees of correlation with regional characteristics such as percent sand and elevation. In this section, we examine whether the spatial patterns of the regional characteristics themselves can be approximated by a small number of orthogonal spatial structures and whether these spatial structures have any relation with soil moisture patterns. This approach is useful because regional characteristics are often correlated with each other. By identifying EOFs from the regional characteristics, we determine spatial patterns for the regional characteristics that are independent of each other. Analyzing these patterns may help clarify the relationships between soil moisture and regional characteristics identified above.

The EOF/PC analysis in this section was conducted as follows. The regional and topographic characteristics most related to spatial and temporal soil moisture were used as the dataset for the EOF/PC analysis. The included characteristics were selected based on the variance the characteristics explained in the spatial and temporal anomaly analyses above. The mean and standard deviation were removed from each characteristic and the EOF/PC analysis was used to distill patterns of regional variability. The EOFs produced by this type of analysis remain spatial structures similar to the previous sections, but the PCs now indicate the dependence of the patterns on the regional characteristics. The variability explained by each EOF is now meaningless, so the EOFs must be sorted according to the variables they most represent.

7.1 Spatial Anomalies

By looking at the correlation between the spatial anomaly EOFs and regional characteristics in Section 4.1, it was determined that percent sand, percent clay, bulk density, surface roughness and slope were the most important characteristics. These six spatial patterns were used as the dataset for the EOF/PC analysis. Figure 17 shows the EOFs generated from this analysis, and Table 4 shows the dependence of each EOF on the regional characteristics. As shown in this table, the EOF/PC pairs are combinations of all the characteristics to different degrees. For example, the first EOF/PC pair has been named Elevation because of its strong connection to this characteristic. However, it is also related to the patterns of percent sand, percent clay and slope.

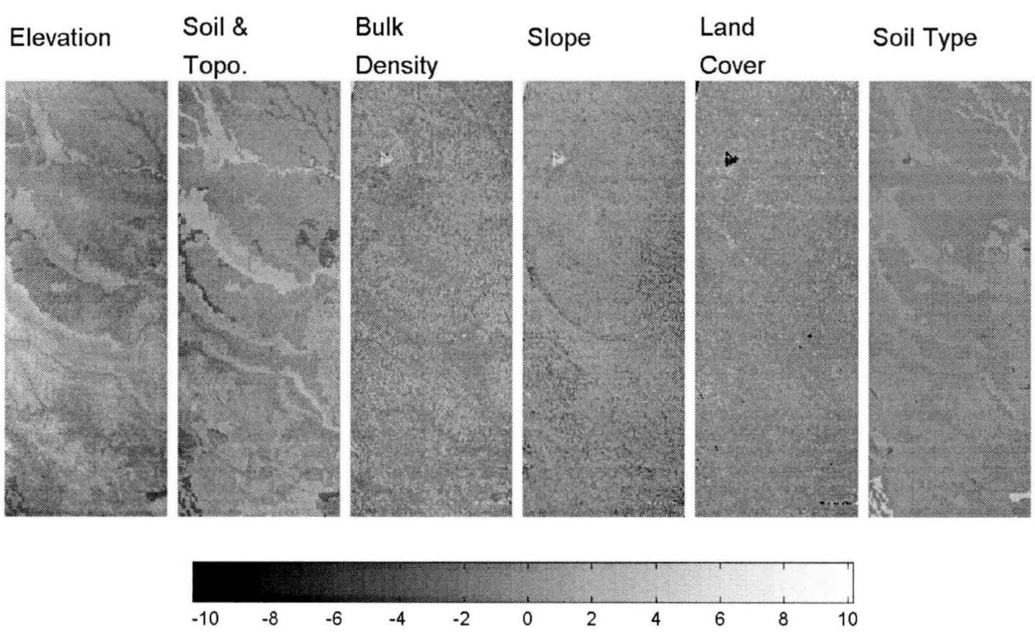

Figure 17: EOFs generated from ancillary data most related to spatial anomaly soil moisture patterns.

Table 4: PC coefficients showing how much each EOF depends on the regional characteristics.

	Elevation PC	Soil and Topo. PC	Bulk Density PC	Slope PC	Land Cover PC	Soil Type PC
% Sand	0.39	0.61	0.12	-0.06	0.03	0.67
% Clay	-0.31	-0.58	-0.14	-0.11	-0.05	0.73
Bulk Density	0.15	0.13	-0.81	-0.26	-0.48	-0.06
Surface Roughness	0.03	0.01	-0.37	-0.38	0.85	-0.04
Elevation	0.81	-0.48	-0.08	0.32	0.09	0.00
Slope	0.26	-0.20	0.41	-0.82	-0.20	-0.11

These EOFs have been correlated to the EOFs of the spatial anomalies of soil moisture as well as the daily soil moisture patterns. The results of these analyses are shown in Tables 5 and 6, respectively. Table 5 shows that spatial anomaly EOF1, which explains 61% of the variance, is most related to the "Elevation" EOF, which depends on elevation, percent sand, and percent clay. It also shows moderate correlations with the other EOFs that depend on soil type, topography, and land cover. Spatial anomaly EOF2 has a very similar tendency except that it is less correlated with the land cover EOF. As shown in Table 6, the soil moisture patterns before the third rain event are most strongly correlated with the elevation EOF. Soil moisture patterns after the third rain event are most strongly correlated with the soil type and topography EOF. In the previous sections EOF1 of spatial anomaly data was most related to percent sand across scales, and the role of topography was minimal. The primary EOF of soil moisture data shows moderate correlations to the patterns that combine topography and soil type. However, none of these combined patterns (ancillary EOFs) shows an improved correlation compared to the results in Section 4.1.

Table 5: Correlation of ancillary EOFs with the EOFs of spatial anomaly soil moisture.

Ancillary EOFs / Spatial Anomaly EOFs	Elevation	Soil Type & Topo.	Bulk Density	Slope	Land Cover	Soil Type
EOF1	-0.32	-0.23	-0.04	0.18	0.21	-0.24
EOF2	0.34	-0.32	-0.13	0.00	-0.05	0.19
EOF3	0.12	-0.14	-0.07	0.11	-0.04	-0.12

Table 6: Correlation of ancillary EOFs for spatial anomaly data with daily soil moisture patterns.

Ancillary EOF / Soil Moisture	Elevation	Soil Type & Topog.	Bulk Density	Slope	Land Cover & Use	Soil Type
6/18	-0.39	-0.27	-0.15	0.09	0.14	-0.04
6/19	-0.31	-0.30	-0.17	0.12	0.13	-0.02
6/20	-0.27	-0.28	-0.17	0.13	0.13	-0.04
6/25	-0.12	-0.29	-0.20	0..12	0.05	0.04
6/26	-0.52	-0.12	-0.12	0.08	0.20	-0.13
6/27	-0.49	-0.11	-0.10	0.12	0.21	-0.20
6/29	-0.24	-0.21	-0.11	0.19	0.17	-0.23
6/30	-0.43	-0.16	-0.11	0.12	0.20	-0.15
7/01	-0.45	-0.14	-0.10	0.11	0.20	-0.15
7/02	-0.45	-0.15	-0.14	0.10	0.19	-0.10
7/03	-0.42	-0.12	-0.10	0.10	0.19	-0.11
7/11	0.00	-0.36	-0.15	0.09	0.03	0.07
7/12	-0.21	-0.26	-0.25	0.03	0.13	0.07
7/13	-0.12	-0.24	-0.20	-0.01	0.07	0.11
7/14	-0.12	-0.21	-0.19	-0.03	0.06	0.12
7/16	0.13	-0.26	-0.22	-0.03	0.20	0.13

7.2 Temporal Anomalies

The correlations presented in Section 4.2 showed that the most important regional and topographic characteristic for determination of temporal anomalies in soil moisture were percent sand, vegetation water content, bulk density, elevation slope, and wetness index. So these characteristics were used in the EOF/PC analysis. Figure 18 shows the

EOFs derived from the regional characteristics and Table 7 shows the dependence of the EOFs on the regional characteristics.

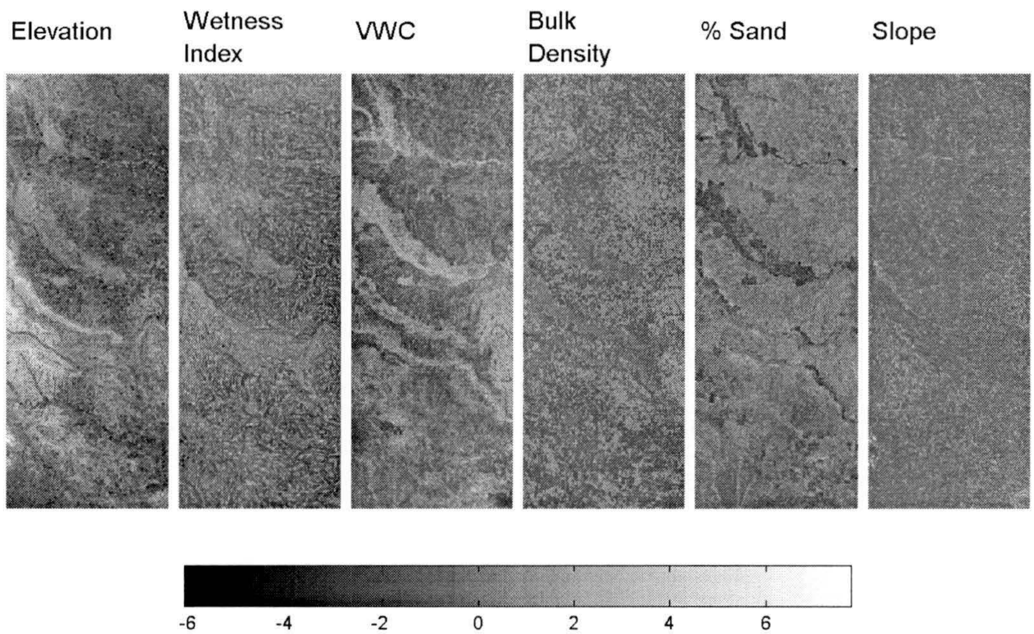

Figure 18: EOFs generated from ancillary data most related to temporal variability.

These EOFs have been correlated to the temporal anomaly EOFs and daily soil moisture patterns. The results of these analyses are shown in Tables 8 and 9, respectively. As shown in Table 8, the temporal anomaly EOF1, which explains 50% of the variance, is most highly correlated with the Elevation EOF, which combines the spatial patterns of elevation, slope, wetness index, and percent sand. In fact, the correlation with the Elevation EOF is larger than the correlation observed for any one variable alone (see Section 4.2). This suggests that the use of the EOF/PC method to distill patterns of regional variability might be useful for identifying dynamic areas of soil moisture. Table 9 shows that the Elevation EOF is moderately to highly correlated with soil moisture data before the third rain event. The ancillary Percent Sand EOF is moderately correlated with all daily soil moisture values. This distillation of ancillary data in connection to temporal anomaly EOFs has shown that the use of EOF/PC analysis to identify patterns of regional variability can identify a more efficient pattern for the description of temporal soil moisture variability.

Table 7: Dependence of ancillary EOFs for temporal anomalies on regional characteristics.

	Elevation PC	Wetness Index PC	VWC PC	Bulk Density PC	% Sand PC	Slope PC
% Sand	0.32	0.16	0.56	0.13	-0.74	0.01
VWC	0.24	0.00	0.68	0.21	0.65	-0.13
Bulk Density	0.12	0.33	0.17	-0.92	0.09	0.04
Elevation	0.75	0.44	-0.44	0.18	0.11	-0.04
Slope	0.39	-0.55	0.00	-0.11	0.04	0.73
Wetness Index	-0.34	0.61	0.09	0.22	0.10	0.67

Table 8: Correlation of temporal anomaly EOFs with EOFs of regional characteristics.

Ancillary EOF / Temp. Anomaly EOF	Elevation	Wetness Index	VWC	Bulk Density	% Sand	Slope
EOF1	0.51	0.12	0.17	-0.08	-0.05	0.12
EOF2	-0.39	-0.07	-0.07	0.09	0.02	-0.06
EOF3	0.02	0.08	-0.15	-0.04	-0.08	0.00

Table 9: Correlation of the ancillary EOFs generate for temporal anomalies with daily soil moisture.

Soil Moisture \ Ancillary EOF	Elevation	Wetness Index	VWC	Bulk Density	% Sand	Slope
6/18	-0.34	0.01	-0.19	-0.01	0.27	-0.11
6/19	-0.28	0.06	-0.22	-0.01	0.27	-0.10
6/20	-0.25	0.09	-0.19	0.00	0.26	-0.10
6/25	-0.11	0.15	-0.20	0.04	0.21	-0.07
6/26	-0.48	-0.04	-0.09	0.00	0.25	-0.13
6/27	-0.47	-0.03	-0.12	0.03	0.23	-0.14
6/29	-0.23	0.07	-0.18	0.05	0.29	-0.15
6/30	-0.41	0.01	-0.14	0.04	0.25	-0.10
7/01	-0.43	-0.01	-0.12	0.04	0.23	-0.11
7/02	-0.42	0.02	-0.11	0.02	0.24	-0.11
7/03	-0.40	0.00	-0.09	0.03	0.21	-0.11
7/11	0.05	0.11	-0.17	0.00	0.30	-0.07
7/12	-0.14	0.12	-0.04	-0.05	0.34	-0.07
7/13	-0.05	0.08	-0.01	-0.05	0.29	-0.05
7/14	-0.05	0.07	-0.03	-0.06	0.29	-0.04
7/16	8.18	0.16	0.00	-0.07	0.28	0.00

8 Conclusions

We have utilized EOF/PC analysis to determine spatial patterns that efficiently explain the variance in the SGP97 soil moisture dataset, and we have utilized correlation analyses to determine whether these patterns are related to regional characteristics. From this analysis, we can draw three main conclusions.

First, a seemingly complex dataset of soil moisture can be well approximated by a very small number of underlying spatial structures. For the spatial anomaly data, one EOF explains 61% of the variance in a sixteen day dataset, suggesting that a large amount of the variability in space is fixed in time. The role of this EOF is modulated by its associated PC, but the spatial pattern remains unchanging in time. The PC time series shows clear cycles of wetting and drying. The spatial patterns of temporal variability are slightly more complex. However, a single EOF can still explain 50% of the total variability. The ability to capture such a large amount of variability in just one pattern could be useful for downscaling large-scale remotely-sensed data.

Second, this analysis has identified soil texture as the most important factor in determining these EOF patterns. When analyzing the spatial anomalies, percent sand and percent clay were the most important factors in determining the primary EOF pattern. Percent sand tends to identify dry locations whereas percent clay tends to identify wet locations. As one considers soil moisture patterns at larger spatial scales, percent sand becomes even more important in determining areas with low soil moisture, although land-cover characteristics become increasing correlated with the primary EOF. Percent sand also plays a more significant role on wet days, whereas percent clay become more important on dry days. For the temporal anomalies, percent sand is still the most important characteristic in determining the primary EOF, but percent clay becomes unimportant. This result suggests that percent clay is useful in identifying sites that are wet relative to other locations, but it is not useful in determining locations that are wet relative to their 16 day average.

Third, topography was shown to be an unimportant characteristic in determination of soil moisture across the range of scales considered. Although elevation was moderately correlated with EOF2 for the spatial anomalies, we believe this correlation occurs simply because it rained in an area with high elevations. High elevations also are associated with locations with more dynamic soil moisture. While this result may also be coincidental as well, higher elevations may be more dynamic because they have poorer connections with stable water sources such as regional aquifers and rivers. Other topographic characteristics such as slope, wetness index, and curvature were often relatively unimportant in explaining the soil moisture patterns. These characteristics are expected to have an influence on the variability of soil moisture through lateral flows, which are not easily observed at a 0.64 km^2 resolution.

Overall, this research has demonstrated the usefulness of the EOF/PC method for identifying patterns of physically-based variability in soil moisture.

References

[1] Allen PB, Naney JW. Hydrology of the Little Washita River Watershed, Oklahoma: Data and Analyses. USDA, ARS-90, 1991.

[2] Beven KJ, Kirkby MJ. A physically based, variable contributing area model of basin hydrology. Hydrol Sci Bull 1979;24(1):43-69.

[3] Choudhury BJ, Schmugge TJ, Chang A, Newton RW. Effect of surface roughness on the microwave emissions from soils. Jl of Geoph Res 1979;84(C9): 5699-5706.

[4] Crow WT, Wood EF. Multi-scale dynamics of soil moisture variability observed during SGP'97. Geoph Res Letters 1999;26(23): 3485-3488.

[5] Famigletti JS, Deveraux JA, Laymon CA, Tsegaye T, Houser PR, Jackson TJ, Graham ST, Rodell M, van Oevelen PJ. Ground-based investigation of soil moisture variability within remote sensing footprints during Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Res Research 1999;35(6): 1839-1851.

[6] Giorgi F, Mearns LO, Shields C, Mayer L. A regional model study of the importance of local versus remote controls of the 1988 drought and the 1993 flood over the central United States. Jl of Clim 1996;9(5): 1150-1162.

[7] Jackson T J, Le Vine DB, Hsu AY, Oldak A, Starks PJ, Swift CT, Isham JD, Haken M. Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains hydrology experiment. IEEE Trans on Geoscience and Remote Sensing 1999;37(5): 2136-2151.

[8] Kim G, Barros AP. Space-time Characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data. Remote Sensing of Env 2002;81: 393-403.

[9] Kim G, Barros AP. Downscaling of remotely sensed soil moisture with a modified fractal interpolation method using contraction mapping and ancillary data. Remote Sensing of Env 2002;83: 400-413.

[10] Liu, Y (2003). Spatial patterns of soil moisture connected to monthly-seasonal precipitation variability in a monsoon region. Jl of Geoph Res 2003;108(D22): 8856, doi:10.1029/2002JD003124.

[11] Miller DA, White RA. A conterminous United States multi-layer soil characteristics dataset for regional climate and hydrology modeling. Earth Interactions 1998;2: 1-15.

[12] Mitasova H, Hofierka J. Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry." Mathematical Geology 1993;25(6): 657-669.

[13] Mitasova H, Mitas L. Interpolation by regularized spline with tension: I. Theory and implementation. Mathematical Geology 1993;25(6): 641-655.

[14] North GR, Bell TL, Chalan RF, Moeng FJ. Sampling errors in the estimation of empirical orthogonal functions. Monthly Weather Review 1982;110: 699-706.

[15] Pellenq J, Kalma J, Boulet G, Saulnier G-M, Wooldridge S, Kerr Y, Chehbouni A. A disaggregation scheme for soil moisture based on topography and soil depth. Jl of Hydro 2003;277: 112-127.

[16] Richman MB. Rotation of principal components. Jl of Clim 1986;6: 293-335.

[17] Rodriguez-Iturbe I, Vogel GK, Rigon R, Entekhabi D, Castelli F, Rinaldo A. On the spatial organization of soil moisture fields. Geoph Res Letters 1995;22(20): 2757-2760.

[18] Vorhoest NEC, Troch PA, Paniconi C, De Troch FP. Mapping basin scale variable source areas from multitemporal remotely sensed observations of soil moisture behavior. Water Resources Res 1998;34(12): 3235-3244.

[19] Western AW, Grayson RB, Bloschl G, Willgoose GR, McMahon TA. Observed spatial organization of soil moisture and its relation to terrain indices. Water Resources Res 1999;35(3): 797-810.

[20] Wilson DJ, Western AW, Grayson RB. A terrain and data-based method for generating the spatial distribution of soil moisture. Adv in Water Resources 2005;28: 43-54.

[21] Wittrock V, Ripley EA. The predictability of autumn soil moisture levels on the Canadian prairies. Intl Jl of Clim 1999;19: 271-289.

[22] Yoo C, Kim S. EOF analysis of surface soil moisture field variability. Adv in Water Resources 20041;27: 831-842.