An approach of spatially- and temporally-extensive soil moisture data combination based on EOF analysis

Ying Zhao[1], Fei Li[2], Rongjiang Yao[3], Wentao Jiao[3], and Robert Hill[4]

[1]Institute of Plant Nutrition and Soil Science
[2]Chinese Academy of Agricultural Sciences
[3]Chinese Academy of Sciences
[4]University of Maryland

July 13, 2020

Abstract

Modeling and prediction of soil hydrologic processes require the identification of soil moisture spatial-temporal patterns and effective methods allowing the data observations to be used across different spatial and temporal scales. This work presents a methodology for the combination of spatially- and temporally-extensive soil moisture data obtained in the Shale Hills Critical Zone Observatory (CZO) from 2004 to 2010. The soil moisture data sets were decomposed into spatial Empirical Orthogonal Function (EOF) patterns, and their relationship with various geophysical parameters was examined to determine the dominant factors contributing to the profiled soil moisture variability. The EOF analyses indicated that one or two EOFs of soil moisture could explain 76-89% of data variation. The primary EOF pattern had high values clustered in the valley region, and conversely low values located in the sloped hills. We suggest a novel approach to integrate the spatially-extensive manually measured datasets with the temporally-extensive automated monitored datasets based on the EOF analyses. Given the data accessibility, the current data merging framework has provided the methodology for the coupling of the mapped and monitored soil moisture datasets, as well as the conceptual coupling of slow and fast pedologic and hydrologic functions. This successful coupling implies that a combination of different extensive moisture data has provided interesting insights into our understanding of hydrological processes at multiple scales.

An approach of spatially- and temporally-extensive soil moisture data combination based on EOF analysis

Ying Zhao[1,2,*], Fei Li[3], Rongjiang Yao[4,*], Wentao Jiao[5], Robert Lee Hill[6]

[1]College of Resources and Environmental Engineering, Ludong University, Yantai 264025, China;[2]Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA;[3]Grassland Research Institute, Chinese Academy of Agricultural Sciences, Huhhot 010010, China; [4]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China;[5]Research center for Eco-environment sciences, Chinese Academy of Sciences, Beijing, 100085, China;[6]Department of Environmental Science and Technology, University of Maryland, College Park, MD 20742, USA;[*]Corresponding author: yzhaosoils@gmail.com (Y. Zhao) and rjyao@issas.ac.cn

(R.J. Yao).

Abstract Modeling and prediction of soil hydrologic processes require the identification of soil moisture spatial-temporal patterns and effective methods allowing the data observations to be used across different spatial and temporal scales. This work presents a methodology for the combination of spatially- and

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

temporally-extensive soil moisture data obtained in the Shale Hills Critical Zone Observatory (CZO) from 2004 to 2010. The soil moisture data sets were decomposed into spatial Empirical Orthogonal Function (EOF) patterns, and their relationship with various geophysical parameters was examined to determine the dominant factors contributing to the profiled soil moisture variability. The EOF analyses indicated that one or two EOFs of soil moisture could explain 76-89% of data variation. The primary EOF pattern had high values clustered in the valley region, and conversely low values located in the sloped hills. We suggest a novel approach to integrate the spatially-extensive manually measured datasets with the temporally-extensive automated monitored datasets based on the EOF analysis. Given the data accessibility, the current data merging framework has provided the methodology for the coupling of the mapped and monitored soil moisture datasets, as well as the conceptual coupling of slow and fast pedologic and hydrologic functions. This successful coupling implies that a combination of different extensive moisture data has provided interesting insights into understanding of hydrological processes at multiple scales.

Keywords: soil moisture, data merging, Empirical Orthogonal Function

1. Introduction

Soil moisture is a key variable within earth system dynamics from regional to pedon scales (Famiglietti et al., 1998; Korres et al., 2015). Identification and prediction of soil moisture patterns at different scales are important in a wide range of agronomic, hydrological, pedological, and environmental studies (e.g., Grayson et al., 2002; Lin et al., 2006). However, obtaining accurate information on soil moisture at an appropriate temporal and spatial scales is still challenging (e.g. Crow et al., 2012; Vereecken et al., 2008). This shortage of information has impeded the modeling, predication, and management of water resources (Owe et al., 1982; Grayson et al., 1997; Zhu and Lin, 2010; Shi et al., 2015). Linking soil moisture mapping with monitoring can provides a more integrated approach to understanding soils and water resources (Dunn and Lilly, 2001; Joshi and Mohanty, 2010; Ma et al., 2017), and provides a way of combatting the general decline of field hydrology relative to modelling (Beven et al., 2020).

Previous studies have indicated that the spatial pattern of soil moisture is highly dependent on various controlling factors such as parent material, soil, landuse/vegetation, topography and climate (Grayson et al., 1997; Famiglietti et al., 1998; Western et al., 1999; Baldwin et al., 2017). However, separating the relative importance of individual factors is usually extremely difficult as some of the influencing factors are interdependent due to the scale-dependence, dry-wet cycles, and interface exchange from top to deep soil layers (Zeleke and Si, 2006; Famiglietti et al., 1998; Western et al., 1999). During wet conditions, subsurface lateral flow may be important, whereas during dry conditions, the surface vertical flow at the onset of precipitation has more influence (Grayson et al., 1997; Gómez-Plaza et al., 2001). The correlation length of soil moisture ranges from near zero to a thousand meters. It has been found to be highly dependent on the scale of the study and on the density of the measurements (Western and Blöschl, 1999), with a clear direct proportionality of increasing correlation lengths with increasing scales having been observed. For the on-site soil moisture monitoring system or networks, the correlation measurements typically include the several sites, whereas for the measurements characterizing soil moisture variation within the field, the area can cover several ha areas, or even much large areas for satellite-based investigations. It is evident that the potential exists for more accurate estimations of soil moisture by using synergistic approaches among a variety of earth observation methodologies.

The procedures for associating the relatively stable spatial patterns of soil moisture and specific soil hydrologic processes to precipitation events are sorely lacking. The temporal variations associated with field measurements of soil moisture have been displayed within numerous studies that have shown the static, terrain-derived indices rarely explained more than 50% of the soil moisture variability (Famiglietti et al. 1998; Western et al. 1999; Baldwin et al., 2017). Although climate seasonality and localized vs. non-local hydrological fluxes exert a strong influence on surface or near-surface soil moisture distribution (Grayson et al., 1997; Western et al., 1999; D'Odorico and Porporato, 2010), less is understood regarding how the subsurface soil moisture spatial patterns change with climate seasonality and seasonal changes in hydrological fluxes. By integrating spatially-distributed soil physical properties, topographical variables and meteorological data,

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

Schwärzel et al. (2009) developed a novel approach in model-based mapping of soil moisture within forested sites. This approach provided a more objective description of soil moisture variability than the traditional mapping by an integrated ecological approach. Other research using a modification of this approach showed that both the spatial distribution and temporal evolution of soil moisture may be investigated at multiple scales (e.g., Deiana et al., 2008). Another potential improvement was the integration of geophysical scanning with real-time soil moisture monitoring and dye staining (Guo et al., 2019). Real-time soil moisture monitoring may provide information concerning actual hydrologic dynamics and the timing of preferential flow occurrence, while dye staining patterns may validate the spatial distribution of preferential flow pathways as revealed by geophysical scans. The synthesis of these methods may enhance the data accuracy and associated processes, thus increasing our understanding of hydrological processes in various soils and landscapes (Ma et al., 2019).

Spatial dependency is commonly characterized and quantified by geostatistical methods, such as autocorrelation and variogram analysis (Zhao et al., 2011). An additional way to analyze the spatial patterns of soil moisture and their connection to regional characteristics is through an empirical orthogonal function (EOF) analysis (Jawson and Niemann, 2007; Zhao et al., 2012). The EOF analysis decomposes a dataset into a series of orthogonal spatial patterns. These patterns may be correlated with regional characteristics to identify whether the characteristics have an influence on the most important tendencies of the soil moisture. Utilizing correlation analyses, these underlying (stable) patterns of soil moisture variations can be connected to parameters derived from topography, soil, vegetation, land management and meteorology. Yoo and Kim (2004) investigated the spatial and temporal variability of field-scale soil moisture and concluded that there is no simple and unique mechanism that can be applied to explain the evolution of the soil moisture field. Wagenet (1998) summarized the main factors that influence soil moisture from the pore scale to the global scale, and recognized soil and topography as important local controls of soil moisture variations. During dry periods, soil moisture distribution may be conceptualized as being controlled predominantly by soil properties, whereas during wet periods the topography is the controlling factor within a landscape. However, the pedon/plot scale results have not been easily transferable to hillslope and catchment scales, and both theoretical and empirical approaches have been used to quantify hydrological dynamics based on such 'point-scale' data that tends to over- or underestimate parameters and fluxes (Sidle et al., 2017). While the temporal frequency for the manual measurement of soil moisture has been found to normally be about weekly-or monthly-based, the automated monitoring data may be hourly-based or daily-based (Korres et al., 2015).

In this study, part of our soil moisture dataset contains manually measured data that offers better spatial coverage with many sites, but with limited temporal frequency as weekly measurements were made. Another part of our soil moisture dataset contains automatically monitored data that offers better temporal frequency with repetitive ten-min to one-hour measurements, but limited spatial coverage with a limited number of selected sites. It is expected that the product of this combination captures the temporal variation of soil moisture from the automatically monitored data and improved spatial resolution using the information from manual measurements. The objectives of this study were to: (1) derive the dominant soil moisture spatial-temporal patterns based on multi-year datasets; (2) determine how these patterns are controlled by terrain, soil, and vegetation as a function of scale, wetness, and depth; and (3) provide a possible way to integrate spatial-extensive datasets with temporal-extensive datasets.

2. Materials and Methods

2.1 Study area

The forested Shale Hills catchment has an area of 7.9-ha, and is located in Huntingdon County, PA, on the eastern coast of the USA. The catchment is V-shaped overall, with a first-order stream in the valley and moderately steep slopes (up to 25%–48%) on both sides of the stream (Fig. 1). Swales are inter-dispersed within the catchment, with five on the south-facing slope and two on the north-facing slope. Elevation ranges from 256 m at the outlet of the catchment to 310 m at the highest ridge. The catchment is underlain by approximately 300-m thick, steeply bedded, highly fractured Rose Hill Shale. The soils were formed from shale colluvium or residuum, with many channery shale fragments throughout most of the soil profiles. Five

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

soil series were identified in the catchment according to the USDA classification (Takagi et al., 2011; Fig. 2). The Shale Hills catchment has a typical humid continental climate, with mean monthly temperatures minimum of -3degC in January and maximum of 22degC in July, and an annual precipitation of about 980 mm (National Weather Service, State College, PA). The precipitation is roughly evenly distributed throughout the year, during the summer months typically occurs in convective weather fronts that can produce high intensity, short duration storm events.

2.2 Data collections

At the Shale Hills Catchment, a total of 106 manual soil moisture measurement locations were recorded on about a weekly basis over a 6-yr period (2004 to 2010; Fig. 1). A portable TRIME-FM Time Domain Reflectomery (TDR) Tube Probe (IMKO, Ettlingen, Germany) was used to determine soil moisture contents while being placed at specific depth (i.e., 0.1-, 0.2-, 0.4-, 0.6-, 0.8-, and 1.0-m) intervals using a PVC access tube installed at each sampling site. The number of locations measured on each measurement day varied due to the number of actual locations, personnel availability, and also weather conditions. The subsequent analysis includes 36 days where at least 65 soil moisture annual measurement locations were measured, as a trade-off between having a sufficient number of measuring sites to adequately represent spatial coverage and sufficient sampling days to have a more complete temporal coverage. Among these 36 measurement days, it can be classified as13 wet (>22%), 14 moist (<22%, >15%), and 9 dry (<15%) times based on the field averaged wetness condition, respectively.

At each sampling site of soil water content, we also measured other properties that were anticipated to influence the water transport and storage. The intact soil cores were collected throughout the catchment during the installation of soil monitoring tubes. Each soil core was first described using standard soil survey procedures, including horizon thickness, color, texture, structure, roots, and amount of redoximorphic features (Fig. 2). A digital elevation model (DEM) of the catchment was interpolated from light detecting and ranging (LiDAR) elevation point clouds collected by an airplane flown over the catchment in 2006 (PAMAP, PA Department of Conservation and Natural Resources). Approximately 40,000 LiDAR elevation points were converted into a 1x1 m DEM using ArcGIS 9.2 (ESRI Inc., Redland, CA, USA). The smoothed DEM was then use to derive the four primary topographic attributes (elevation, slope, curvature, and upslope contributing area) and one composite topographic attribute (topographic wetness index). Terrain attributes were then extracted at all soil moisture monitoring locations using the coordinates obtained from the total station survey.

In addition to spatially extensive manual measurement, soil moisture has been monitored in real-time at 5 representative sites, with 4-10 depths at each site depending on soil thickness and horizonation. These sites are located from the ridge top (site 74) to the valley bottom (site 61), and both planar (site 51)/convex (site 53) hillslopes and concave swales (site 15) (Fig. 1). At each long-term monitoring site, a pit was excavated, and capacitance-type probes of Decagon Devices, Pullman WA (EC10 or EC5) were used to monitor the profiled soil moisture at 10 min intervals. A Pluvio load cell rain gauge (OTT Hydrometry, Kempten, Germany; precision) was located in a clearing on the north ridge of the catchment and automatically recorded precipitation every 10 minutes automatically. Further details on the soil moisture probe installation can be found in Lin and Zhou (2008).

2.3 EOF analysis

Empirical orthogonal function (EOF) analysis has been widely applied for the analysis of the spatial and temporal variability of large multidimensional datasets (Zhao et al., 2012). The EOF, also known as a type of principal component analysis, decomposes the observed variability of a dataset into a set of orthogonal spatial patterns (EOFs) or a set of time series called expansion coefficients (ECs). This procedure is accomplished by transforming the original data set into a new set of uncorrelated variables, and then ordered in a manner so that the first few of the new variables explain most of the variation existing in the original data set. For example, it is possible to construct various second moment statistics linking one point to another in geophysical data maps. The resulting correlation matrix is real and symmetric, and therefore possesses a set

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

of orthogonal eigenvectors with positive eigenvalues. If there are geophysical data maps that are time series with any m x n matrix, A , square or rectangular, there uniquely exists two orthogonal matrices, U and V and a diagonal matrix L such that,

$A = U$ x L x V^{T} (1)

where V^{T} is the transpose of a matrix V . Note that L is padded with zeros to make the square diagonal matrix into an m x n matrix. This assumption also implies that L has at most $M = \min(m\ ,\ n\ )$ nonzero elements. The columns of U are called the EOFs of A and the corresponding diagonal elements of L are called the eigenvalues. Each row of V serves as a series of time coefficients that describes the time evolution of the particular EOF. The map associated with an EOF represents a pattern, which is statistically independent and spatially orthogonal to the others. The eigenvalue indicates the amount of variance accounted for by the pattern (Zhao et al., 2012).

While single soil moisture patterns might be affected by random processes (e.g., rainfall shortly before measurement), significant EOFs represent stable patterns of a dataset. The existing degree of randomness of a single soil moisture pattern is reflected by the associated EC, since the EC value represents the proportion of the significant EOF pattern within the soil moisture pattern of each date. In consequence, single soil moisture patterns (which might be random) were not used but the EOF patterns used for the subsequent correlation analysis. That is, the EOF patterns can be further correlated to the geophysical characteristics of the region to determine the dominant physical controls. For the EOF analysis, we used the spatial anomalies of the soil moisture dataset instead of the soil moisture which excludes the temporal variations from consideration (Perry and Niemann, 2007). The spatial anomalies are calculated by subtracting the mean soil moisture for a given sampling day from all the soil moisture observations collected on that day.

2.4 Data combination

One of the primary benefits of the EOF analysis is a small number of orthogonal spatial patterns was identified that together explain a large proportion of the total variability of the soil moisture data. We now examined how closely these underlying patterns resembled regional characteristics that might dominate the spatial variability of soil moisture. For this analysis, we used the correlation coefficient between the EOFs and the available regional characteristics. Statistical analyses were conducted using SPSS for Windows (SPSS Inc., Chicago). The Pearson's correlation was used to investigate the correlation between soil moisture and soil-terrain attributes. Linear regression was used to predict soil moisture (as a dependent variable) based on using either soil properties and/or terrain attributes (as independent variables).

Furthermore, given that the first EOFs exhibited the general patterns of soil moisture across the whole investigated spatial coverage, if the EOFs of the manual soil moisture measurements at some sites had the strong correlations with its automatically recorded values, then it was possible to apply this relationship to the first EOFs to derive the soil moisture across the entire area at any automated monitoring time. Consequently, we considered that this method may have provided an appropriate way to integrate the spatially-extensive (but temporally-limited) manual datasets and the temporally-extensive (but spatially-limited) automated monitoring datasets. The following narrative describes in detail the four steps involved with the data merging methods.

First, an EOF decomposition was performed on the soil moisture dataset to identify the patterns of covariation (the EOFs) and their importance on each date (the ECs).

Second, statistical tests were used to determine the EOFs between whole spatially-extensive data and automated monitored data that were statistically significant and should be retained in the transferring method.

Third, a multiple linear regression was performed to identify empirical relationships between automated monitored soil moisture data and its EOFs.

Fourth, the identified empirical relationships were employed to calculate soil moisture at each manual measured site within the entire catchment via its EOFs.

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

3. Results and Discussion

3.1 Spatial-temporal patterns of soil moisture

The variogram analysis and histogram of soil moisture storage indicated that interpolated soil moisture maps exhibited seasonal alignments of soil moisture storage along topographic convergent areas (Fig. 3). We found the sample variograms with a clear sill and nugget and observed that the geostatistical structure of soil moisture was seasonally evolved. During the wet winter period, high sills (15-25 $(\%)^2$) and low correlation lengths (20-30 m) were observed, whereas during the dry summer periods, sills were smaller (10-15 $(\%)^2$) and correlation lengths were longer (30-40 m). Regardless of the wetness conditions, the wettest soil was always located within the swales and the valley floor (i.e., near-stream zone). These wet-up and dry-down patterns were consistent with the overall distribution of the soil types and the topographic wetness index within the catchment. There was an exponential increase in the catchment-wide soil moisture variability with increased averaged-catchment moisture contents (Zhao et al., 2012). These conditions were obvious due to the well-drained and steep-sloped soils within the catchment that confined saturated areas to the swales and the valley floor.

The soil moisture variability was explained by using only the first few EOF patterns within the Shale Hills (Table 1). At the soil depths, the first four EOFs together explained approximately 87% of the total variability, whereas only the first EOF (or EOF1) explained about 76% of the total soil moisture variance, indicating that a single spatial structure may explain much of the overall soil moisture pattern. With increased soil depths, the total variations explained by the derived EOFs also increased. These results indicated that the seemingly complex patterns of soil moisture within the Shale Hills may largely be explained by a very small number of the underlying spatial EOFs. In the EOF analysis of spatial patterns, the impacts of temporally variable factors, which do not affect the whole area uniformly, also resulted in noise and would also be expected to have decreased the amount of the variance explained by the significant EOFs.

A close examination of the EOF patterns associated with soil land units in Figure 4 reveals that the EOF1 displayed high values within the valley floor, and low values within the hillslopes, respectively. Obviously, the high EOF values indicated the clustered site with the above average soil moistures, and conversely low EOF values is equivalent to the sites of below average soil moisture values (Fig. 4). From the weighted EC series (Fig. 5), the variance explained by the EOF1 values closely followed the increased field mean moisture contents, e.g., the variance is sharply increased with increased moisture contents following rainfall recharge. Therefore, the EOF analyses seem to represent a very powerful set of tools that helped explain the patterns in the variance associated with the general spatial patterns, the indications of positional characteristics, and the temporal dynamics. Perry and Niemann (2007) applied an EOF analysis for a 10.5 ha Tarrawarra grassland catchment, and the first EOF in their study explained 55% of the soil moisture spatial variability. The explained variances found at the Shale Hills are higher than the previously mentioned studies that were about 55% to 70% of surface soil moisture variability that may be explained by the stable spatial patterns associated with the soil parameters and topography at their study sites (Perry and Niemann, 2007; Korres et al., 2010). Because of the strong combined soil-topographic effects, the observed soil moisture patterns in the Shale Hills was high, and can largely be explained by only a few underlying spatial structures or EOF patterns that are obviously correlated to the various geophysical characteristics.

3.2 Controls of primary soil moisture patterns

The spatial patterns of soil moisture show the higher Pearson correlation coefficients with soil-topography properties at each measurement depth (Table 1). The results generally indicated that terrain features were larger contributors to the variance in soil moisture than the soil properties. While most of topographical attributes (e.g., topography wetness index and slope) had strong correlations with the derived EOFs, only soil texture among the soil parameter showed significant correlations. Soil organic matter displayed lower correlations at the surface soil (0-20 cm). Depth to bedrock, which is related to both soil thickness and topography, seems to have had a large influence on soil moisture variability at all depths. This result may be confirmed by examination of the soil moisture values within the wet locations which are characterized by the

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

soils with >1 m depth to bedrock. Elevation, slope, and curvature were negatively related to soil moisture contents, while upslope contributing area, depth to bedrock, topographic wetness index, and percent silt and clay were positively correlated to the soil moisture contents. This result is likely due to the fact that most soils with the deep soil profiles are generally limited to lower elevations and concave slope areas (i.e., valley floor and swales) where soil moisture is normally the highest. Henninger et al. (1976) reported that soil moisture increased toward the near-stream zone within a predominantly agricultural watershed, which was a result of both topographic convergence and moderately to poorly drained soils within the near-stream zone. Our regression analysis indicated that soil texture did not exert a strong influence on the soil moisture spatial distributions at the catchment scale, particularly at soil depth intervals between 0.4 and 0.6 m. This finding is due to the relatively small variations in soil textural properties throughout the measured locations for the different soil-landform units (Takagi and Lin, 2011). Famiglietti et al. (1998) reported that under wet conditions, the best correlations existed with porosity and hydraulic conductivity along a profile of a 200-m length; whereas under dry conditions, the relative elevation, aspect and clay content provided the best correlations.

The EOF analysis was repeated for the spatial anomaly data in two categories (depth/wetness), and the degree to which these factors affected the soil moisture distribution was calculated (Table 1). At the Shale Hills, the correlation coefficient values generally increased with soil depth for elevation, slope, and percent silt and clay content, whereas the highest values were observed at intermediate depths (0.4 m) for curvature, depth to bedrock, and topographic wetness index. These results indicate that soil moisture becomes strongly aligned with convergent topography and suggests that lateral flow processes may be important driver of soil moisture redistribution at these depths. The increased influence of the parameters with depth may relate to the seasonal changes of soil moisture which undergoes more dramatic changes near the soil surface. As indicated in Takagi and Lin (2012), the subsurface soil moisture exhibited weak temporal variability in the correlation coefficient values that suggested the dampened effects of climate and hydrological fluxes. Thus, the subsurface soil moisture distribution in this catchment is a function of both topographic parameters and soil depth. An observation that was reinforced by the transient hydrological fluxes such as the presence of the ephemeral shallow water table that seasonally exist within the valley.

3.3. Data combining method of spatially- and temporally-extensive soil moisture

In this study, we found that the first EOF can predict the general patterns of soil moisture, for example at the 20-cm soil depth with 85% of the variance explained. Although the relative importance of the first EOF on daily patterns of soil moisture waxes and wanes during cycles of wetting and drying, the spatial pattern of the EOF is invariant in time. Therefore, we considered it is an efficient way to integrate the spatially extensive (but temporally limited) manual measurement sites with other field long-term automatic monitoring datasets that are temporally extensive (but spatially limited). Conceivably, based on the EOFs for all spatially extensive sites, it was possible to predict the spatially-distributed soil moisture for all those sites based on the derived regressive equation between the EOFs of the temporally-extensive sites and its automated monitoring datasets. For instance, the first EOF of soil moisture at 20-cm soil depth was derived from a manual dataset with the higher explained variations, but there also was a strong-correlated regression coefficient between the soil moisture automated monitoring sites (e.g., five sites 15, 51, 55, 61, and 74) and their corresponding EOFs values. Based on the derived equations, all manual measurement values could be predicted by either the manual measurements or the monitored values at those five monitoring sites. To validate this assumption, we selected three wetness conditions on the same dates as used in Figure 3 (i.e. wet: ; moist: ; dry:). Remarkably, the predicted values via the manually-measured data (Fig. 6) have a strong linear correlation to the measurements with high confidence levels (95%). These results mean that the suggested method is practical to combine the manually-measured datasets with the automated- monitored datasets.

Note that the results showed a relatively large scatter when the automated- monitored values were used to predicate soil moisture values at the spatially extensive sites (Fig. 7). Whether this approach is accurate is also dependent on how well the manually-measured data and the automated-monitored data closely match

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

for the same soil depths at the same sites. Due to the differences in the measured thickness and horizonation, spatial dimensions and scales for the two methods (Gu et al., 2018), the values between manually-measured and automated-monitored datasets may not necessarily match well with one another. As indicated in Figure 8, except for site 55, there were large differences between manually-measured and automated-monitored soil moisture values. For instance, the manually-measured moisture contents are consistently higher than the automated- monitored values for the site 51 during the entire measurement period. Even worse, the trends between both datasets for sites 15 and 74 are somewhat irregular. These results challenged the suitability of this approach when the automated-monitored data, instead of the manually-measured data were used at the temporally-limited site. As shown in Figure 9, we found that the fit between manual-measured and auto-recorded soil moisture datasets were significant, but relatively weak. Therefore, to apply this method reasonably, it is important for the predicted data accuracy accounting the manually-measured and automated monitored data to be somewhat in agreement. It is expected that the EOF method could be a practical and efficient data merging method if the primary EOF explains >60% of the variation. Nevertheless, taking into account those differences, the EOF method as applied in this study could be quite valuable, and therefore provide an essential way to assimilate data from multiple sources.

Furthermore, we explored the EOF method to breakdown a more dynamic time series of soil moisture in to a lesser number of orthogonal spatial EOF patterns (that are invariant in time) and the corresponding EC components (that are invariant in space). This modification greatly simplifies our task as we can just deal with only a few spatial EOF structures instead of the whole data set. The higher-order EOFs are usually taken into account depending on the amount of the total variance explained by them. The associated EC components show the variation in the influence of the EOFs during the wetting/drying phases, which could be reasonably associated with the automated monitoring moisture dynamics and theoretically provided the basic for the data fusion. To determine the dominant physical controls, the EOF patterns were correlated to the geophysical characteristics of the region. From our analyses, we inferred that some of the variability of the soil moisture EOF patterns is related to both topography and soil texture. We assessed that, using the EOF analysis, it is particularly applicable to combine the manual datasets with the automatic datasets in terms of different resolutions for different data sources. The soil moisture dataset is currently providing either better spatial coverage or better temporal coverage. Our data assimilation approach provides an important way to combine both datasets together which certainly improved the explanations for the variation and data use.

3.4 Implications

Existing methods for direct field measurement of soil hydraulic properties remain complex, time-consuming, costly, and significant spatial and temporal variability challenges the possibility of extensive measurements (Ma et al., 2018). However, the spatial patten, as indicated by the EOF analysis, is the relatively stable. This stability attribute of the spatial pattern has implied that it is possible to continuously assess the soil moisture distribution in a catchment. In addition to the spatial coverage maps, adding in the long-term monitoring of both surface and subsurface soil moisture provides a comprehensive picture of the spatial-temporal pattern of soil moisture dynamics across the whole area and allow the identification of factors which influence it through time. A unique long-term real time soil moisture data set was previously used to identify local dominant hydrological processes and its time dynamics. In this perspective, our approach becomes more effective given that the long-term monitored site is characterized as a time-stable location via a time-stability analysis (Zhao et al., 2010). Our approach also has the capability to assimilate additional data sources, e.g., remote sensed data at this time-stable site. Given the high accuracy of the soil moisture monitoring, the time-resolution soil moisture patterns over an area could be obtained by selecting a temporally stable monitoring site, which is useful in ground truthing of a remotely sensed footprint for validation of simulation modelling results (Zhao et al., 2010). Given the importance of soil moisture in Earth's land surface interactions and to a large range of applications, one can appreciate that its accurate estimation is critical in addressing key practical challenges such as food security, sustainable planning and management of water resources. The launch of new, more sophisticated satellites strengthens the development of innovative research approaches and scientific inventions that will result in ground-breaking advancements.

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

One goal of this study was to lay the foundation for the design of cost-effective real-time soil moisture monitoring networks that fill in the gap between point sensors and traditional manual measurements or even remote sensing values. Our study is representative of a novel approach with the potential benefits for an effective soil moisture monitoring network design within the study area, determining the spatiotemporal statistics of the observed soil moisture fields, and the use of a spatial regression procedure in data merging. It is more realistic to observe a difference between developed maps as surface conditions evolve. We believe that this combination reflects more adequately the basin heterogeneity and complex interactions between soil moisture and topographic attributes. Although there are the simple linear data transfer methods that have been applicable to this type of, our approach may accommodate different data analysis methods, such as a multi-step regression method based on the EOF analysis (Temimi et al., 2010). Once the tasks within our approach have been completed, the EOF-based transfer method may be used as a foundation any region and/or date under the assumptions that the identified empirical relationships will be valid for the application conditions.

The present soil-landscape has been shaped through a combination of long- and short-time processes, and this history can provide some clues to project future changes. However, linking long-term and slow processes with shorter-term and fast processes remains one challenge (Ma et al., 2018). While mapping depicts the spatial distribution of soil-landscape relationships, as indicated by the dominant EOF patterns; monitoring captures the temporal dynamics of pedologic and hydrologic properties, as indicated by the profiled data dynamics (Ma et al., 2018). Given the spatial-extensive data benefits of traditional mapping, and the temporal-extensive data benefits of traditional monitoring, the presented data-integrated method may provide a justifiable basis for the combination of mapped and monitored data, as well as a conceptual basis for the coupling of slow and fast processes. Firstly, bridging mapping with monitoring is very helpful in the dynamic mapping of hydropedologic functional units (Ma et al., 2018). Secondly, mapping provides information to aid in optimal site selection for monitoring (Zhao et al., 2010). Thirdly, mapping and monitoring supplies essential data for the calibration and validation of modeling, and may help provide additional information for a more holistic, refined and predictive management of soil and water resources (Guo et al., 2019). Our approach provides an essential set of tools to evaluate the improvement of data use. We assumed that the relationship between different data sources remains the same over time, but suggest that future studies verify this behavior.

4. Conclusions

In this study, we developed a data combination approach on the basis of EOF analysis of space-time soil moisture data at a reference Shale Hills catchment. We investigated the space-time characterization of soil moisture and found that the variation of soil moisture could be explained by using the first few EOFs. Results of the correlation analysis showed that topography and soil properties have mixed effects on the variability explained by the dominant soil moisture EOFs. Benefits based on the derived underlying stable EOF patterns of soil moisture, the relationships between site characteristics and the EOF patterns were examined to conduct a spatial-temporal dataset combination. Based on the long-term spatial extensive sampling campaign and the specific transect of real-time monitoring, this study investigated how to integrate the spatially extensive, but temporally limited manual datasets with the temporally extensive, but spatially limited automated monitoring datasets. This exercise was helpful for understanding the soil moisture spatio-temporal patterns and the hydrological responses at a small landscape scale, and are considered to be important factors for the effective measurement and the practical management of soil moisture at multiple scales.

Acknowledgments

This research was supported in part by the U.S. National Science Foundation through the Shale Hills Critical Zone Observatory grant (EAR-0725019), the Taishan Scholars Program (201812096) and the National Natural Science Foundation of China (41977009). Fei Li was supported by the Agricultural Science and Technology Innovation Program of CAAS (CAAS-ASTIP-2020-IGR-04). Dr. Hill was partially supported by the USDA National Institute of Food and Agriculture, Hatch project 1014496. Assistance in field data collections from the Penn State Hydropedology Group was gratefully acknowledged.

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

Data Availability Statement

The data set is available at http://criticalzone.org/shale-hills/data/datasets/.

References

Bachmair, S., Weiler, M., Nutzmann, G., 2009. Controls of land use and soil structure on water movement: Lessons for pollutant transfer through the unsaturated zone. J. Hydrol. 369, 241–252.

Baldwin, D., Naithani, K.J., Lin, H. 2017. Combined soil-terrain stratification for characterizing catchment-scale soil moisture variation. Geoderma. 285, 260–269.

Beven, K., Asadullah, A., Bates, P., Blyth, E., Chappell, N., Child, S., . . . and T., Wagener. 2020. Developing observational methods to drive future hydrological science: Can we make a start as a community? Hydrol. Proc., 34, 868–873.

Bourennane, H., King, D., Couturier, A., 2000. Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. Geoderma. 97, 255–271.

Brocca, L., Morbidelli, R., Melone, F., & Moramarco, T. (2007). Soil moisture spatial variability in experimental areas of central Italy. J. Hydrol. 333(2-4), 356–373.

Deiana, R., Cassiani, G., Villa, A., Bagliani, Andrea, Bruno, V., 2008. Calibration of a vadose zone model using water injection monitored by GPR and electrical resistant tomography. Vadose Zone J. 7, 215–226. http://dx.doi.org/10.2136/vzj2006.0137.

D'Odorico, P., Porporato, A., 2004. Preferential states in soil moisture and climate dynamics. Proc. Natl. Acad. Sci. U. S. A. 101, 8848-51. doi: 10.1073/pnas.0401428101.

Dunn, S.M., Lilly, A., 2001. Investigating the relationship between a soils classification and the spatial parameters of a conceptual catchment-scale hydrological model. J. Hydrol. 252, 157–173.

Entin, J.K., Robock, A., Vinnikov, K.Y., Hollinger, S.E., Liu, S., Namkhai, A., 2000. Temporal and spatial scales of observed soil moisture variations in the extratropics. J. Geophys. Res. 105, 11865–11877.

Famiglietti, J.S., Rudnicki, J.W., Rodell, M., 1998. Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. J. Hydrol. 210, 259–281.

Famiglietti, J.S., Ryu, D., Berg, A.A., Rodell, M., and Jackson, T.J., 2008. Field observations of soil moisture variability across scales. Water. Resour. Res. 44, W01423, doi:10.1029/2006WR005804.

Grayson, R.B., Bloschl, G., Western, A.W., McMahon, T.A., 2002. Advances in the use of observed spatial patterns of catchment hydrological response. Adv. Water Resour. 25, 1313–1334.

Grayson, R.B., Western, A.W., Chiew, F.H.S., Bloschl, G., 1997. Preferred states in spatial soil moisture patterns: local and non local controls. Water. Resour. Res. 33, 2897–2908.

Gu, W.Z., Liu, J.F., Lin, H.S., Lin, J., Liu, H.W., Liao, A.-M., Wang, N., Wang, W.Z., Ma, T., Yang, N., Li, X.G., Zhuo, P., Cai, Z. 2018. Why hydrological maze: The hydropedological trigger? Review of experiments at Chuzhou Hydrology Laboratory. Vadose Zone J. 17:170174. doi:10.2136/vzj2017.09.0174.

Guo, L., H. Lin, B. Fan, J. Nyquist, L. Toran, G. Mount. 2019. Preferential flow through shallow fractured bedrock and a 3D fill-and-spill model of hillslope subsurface hydrology. J. Hydrol. 576, 430–442.

Ireson, A.M., Butler, A.P., 2011. Controls on preferential recharge to Chalk aquifers. J. Hydrol. 398, 109-123.

James, A.L., Roulet, N.T., 2009. Antecedent moisture conditions and catchment morphology as controls on spatial patterns of runoff generation in small forest catchments. J. Hydrol. 377, 351–366.

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.

Jawson, S.D., and J.D. Niemann. 2007. Spatial patterns from EOF analysis of soil moisture at a large scale and their dependence on soil, landuse, and topographic properties, Adv. Water Resour., 30(3), 366–381.

Joshi, C.B., Mohanty, B.P. 2010, Physical controls of near surface soil moisture across varying spatial scales in an agricultural landscape during SMEX02, Water Resour. Res., 46, W12503, doi:10.1029/2010WR009152.

Korres, W., T.G. Reichenau, P. Fiener, C.N. Koyama, H.R. Bogena, T. Cornelissen, R. Baatz, M. Herbst, B. Diekkruger, H. Vereecken, and K. Schneider. 2015. Spatio-temporal soil moisture patterns-A meta-analysis using plot to catchment scale. J. Hydrol. 520, 326–341.

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

Koster, R.D., Suarez, M.J., Higgins, R.W., Van Den Dool, H.M., 2003. Observational evidence that soil moisture variations affect precipitation. Geophys. Res. Lett. 30, 1241. doi: 10.1029/2002GL016571.

Li, X.Y., Z.P. Yang, Y.T. Li, and H.S. Lin. 2009. Connecting ecohydrology and hydropedology in desert shrubs: Stemflow as a source of preferential water flow in soils. Hydrol. Earth Syst. Sci. 13, 1133–1144.

Lilly, A., Boorman, D.B., Hollis, J.M., 1998. The development of a hydrological classification of UK soils and the inherent scale changes. Nutr. Cycl. Agroeco. 50, 299–302.

Lin, H., Zhou, X., 2008. Evidence of subsurface preferential flow using soil hydrologic monitoring in the Shale Hills catchment. Eur. J. Soil Sci. 59, 34–49.

Lin, H.S. 2006. Temporal stability of soil moisture spatial pattern and subsurface preferential flow pathways in the Shale Hills Catchment. Vadose Zone J. 5, 317–340.

Lin, H.S., Bouma, J., Pachepsky, Y., Western, A.W., Thompson, J.A., van Genuchten, M.T., Vogel., H., Lilly, A., 2006. Hydropedology: Synergistic integration of pedology and hydrology. Water. Resour. Res. 42, W05301, doi: 10.1029/2005WR004085.

Ma, Y., Li, X., Guo, L., Lin, H., 2017. Hydropedology: interactions between pedologic and hydrologic processes across spatiotemporal scales. Earth Sci. Rev. 171, 181–195.

Martini, E., U. Wollschlager, A. Musolff, U. Werban, and S. Zacharias. 2017. Principal component analysis of the spatiotemporal pattern of soil moisture and apparent electrical conductivity. Vadose Zone J. 16(10). doi:10.2136/vzj2016.12.0129.

Meyles, E., Williams, A., Ternan, L., Dowd, J., 2003. Runoff generation in relation to soil moisture patterns in a small Dartmoor catchment, Southwest England. Hydrol. Process. 17, 251–264.

Owe, M., Jones, E.B., Schmugge, T.J., 1982. Soil moisture variation patterns observed in Hand Country, South Dakota. Water. Resour. Bull. 18, 949–954.

Perry, M.A. and Niemann, J.D., 2008. Generation of soil moisture patterns at the catchment scale by EOF interpolation. Hydrol. Earth. Syst. Sci. 12: 39–53.

Shi, Y., Baldwin, D.C., Davis, K.J., Yu, X., Duffy, C.J., Lin, H., 2015. Simulating high-resolution soil moisture patterns in the Shale Hills watershed using a land surface hydrologic model. Hydrol. Process. 29, 4624–4637. http://dx.doi.org/10.1002/hyp.10593.

Shouse, P.J., Russell, W.B., Burden, D.S., Swlim, H.M., Sisson, J.B. and van Genuchten, M.Th., 1995. Spatial variability of soil water retention functions in a silt loam soil. Soil Sci. 159, 1–12.

Starr, J.L. and Timlin, D.J., 2004. Using high-resolution soil moisture data to assess soil water dynamics in the vadose zone, Vadose Zone J. 3, 926–935.

Takagi, K., Lin, H.S., 2011. Temporal evolution of soil moisture spatial variability in the Shale Hills catchment. Vadose Zone J. 10, 832–842.

Takagi, K., Lin, H.S., 2012. Changing controls of soil moisture spatial organization in the Shale Hills Catchment. Geoderma 173–174, 289–302.

Temimi, M., R. Leconte, N. Chaouch, P. Sukumal, R. Khanbilvardi, and F. Brissette. 2010. A combination of remote sensing data and topographic attributes for the spatial and temporal monitoring of soil wetness, J. Hydrol. 388(1-2), 28–40.

Wagenet, R.J., 1998. Scale issues in agroecological research chains. Nutr. Cycl. Agroecosys. 50, 23–34.

Western, A.W., Grayson, R.B., Bloschl, G., 2002. Scaling of soil moisture: a hydrologic perspective. Annu. Rev. Earth. Pl. Sci. 30, 149–180.

Western, A.W., Grayson, R.B., Bloschl, G., Willgoose, G.R., McMahon T.A., 1999. Observed spatial organization of soil moisture and its relation to terrain indices. Water. Resour. Res. 35, 797–810.

Yoo, C., Kim, S., 2004. EOF analysis of surface soil moisture field variability. Adv. Water Resour. 27(8), 831–842.

Zehe, E., and Bloschl, G. 2004. Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions. Water Resour. Res., 40(10), W10202. doi: 10.1029/2003WR002869.

Zhao, Y., Tang, J., Zhu, Q., Takagi, K., Graham, C.B., Lin, H., 2012. Hydropedology in the ridge and valley: soil moisture patterns and preferential flow dynamics in two contrasting landscapes. Hydropedology (Chapter 12). Edited by H. Lin. Academic Press. 381–411.

Zhao, Y., Peth, S., Horn, R., Hallett, P., Wang, X.Y., Giese, M., Gao, Y.Z., 2011. Factors controlling the spatial patterns of soil moisture investigated by multivariate and geostatistical analysis. Ecohydrology 4, 36–48.

Zhao, Y., Peth, S., Wang, X.Y., Lin, H.S., and Horn, R. 2010. Temporal stability of surface soil moisture spatial patterns and their controlling factors in a semi-arid steppe. Hydrol. Proc. 24, 2507–2519.

Zhu, Q., and H.S. Lin. 2010. Interpolation of soil properties based on combined information of spatial structure, sample size and auxiliary variables. Pedosphere 20, 594–606.

Hosted file

`Table and figure-ZY-Final-F.docx` available at https://authorea.com/users/341747/articles/468629-an-approach-of-spatially-and-temporally-extensive-soil-moisture-data-combination-based-on-eof-analysis

Posted on Authorea 13 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159466981.10661342 — This a preprint and has not been peer reviewed. Data may be preliminary.