Ecological Informatics 5 (2010) 400–409

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Estimating root zone soil moisture at distant sites using MODIS NDVI and EVI in a semi-arid region of southwestern USA

Mark T. Schnur [a,*], Hongjie Xie [a], Xianwei Wang [b]

[a] Department of Geological Sciences, University of Texas at San Antonio, Texas 78249, USA
[b] Department of Earth System Science, University of California, Irvine 92697, USA

ARTICLE INFO

Keywords:
MODIS
NDVI
EVI
SCAN
Soil moisture
Correlation
Regression

ABSTRACT

This study investigates the potential of using Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) to estimate root zone soil moisture at native in-situ measured sites, and at distant sites under the same climatic setting. We obtained in-situ data from Soil Climate Analysis Network (SCAN) sites near the Texas-New Mexico border area, and NDVI and EVI products from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on board the Terra satellite. Results show that soil moisture values of the same depth are highly correlated ($r = 0.53$ to 0.85) among sites as far as 150 km apart, and that NDVI and EVI are highly correlated at the same site ($r = 0.87$ to 0.91). Correlation based on raw time series of NDVI and soil moisture is in general higher than that based on deseasonalized time series at every depth. The correlation reaches maximum value when vegetation index (VI) lags soil moisture by 5 to 10 days. NDVI shows a slightly higher correlation with soil moisture than EVI does by using the deseasonalized time series of NDVI and soil moisture. It is found that deseasonalized time series of NDVI and soil moisture are correlated at native sites ($r = 0.33$ to 0.77), but not at sites where soil moisture is very low. Regression analysis was conducted using deseasonalized time series soil moisture and deseasonalized time series NDVI with a 5-day time lag. Regression models developed at one site and applied to a similar distant site can estimate soil moistures, accounting for 50–88% of the variation in observed soil moistures.

1. Introduction

Soil moisture is the quantity of water contained in soil on a volumetric or gravimetric basis. At the surface, soil moisture is a critical factor in the interaction with the atmosphere. Beneath the surface, root zone soil moisture controls surface vegetation health conditions and coverage, especially in arid and semi-arid areas, where water is one of the main controlling/limiting factors for vegetation growth (Magagi and Kerr, 2001). The root zone is the region of the soil penetrated by vegetation roots, and varies with climate, soil, and vegetation characteristics (Guswa, 2008). About 95% of shrub roots are located in the upper half-meter (0.5 m) in a semi-arid environment in West Africa (Kizito et al., 2006). The majority of roots of blue grama and black grama grasses are concentrated from 5 to 10 cm deep in the grasslands of the Central Great Plains (Vinton and Burke, 1997). Next to our study sites, Kurc and Small (2007) report that the majority of grass roots are near the surface and that the majority of shrub roots are concentrated between 20 and 40 cm, although roots can be present at soil depths up to 1 m in both grassland and shrubland in their study area in central New Mexico. Plant roots directly or indirectly affect many hydrologic components and processes, including soil moisture, ground water, evaporation, transpiration and opening surface-connected hydraulic pathways for rainfall penetration (England, 1975).

Root zone soil moisture links surface phenology and subsurface water storage in vegetated regions and strongly influences surface water balance and energy partitioning due to evapotranspiration (Song et al., 2000). The phenological patterns in vegetation in a semi-arid ecosystem in Mexico are correlated with soil moisture (Pavon and Briones, 2001). The distribution of desert grasses is controlled by soil (type and water content) heterogeneity in central New Mexico (Buxbaum and Vanderbilt, 2007). Within one ecosystem, vegetation self-adjusts its spatial density to match the local climate condition and water availability (Wu et al., 1985; Walker et al., 1989; Miina and Pukkala, 2002). Temporal deviation of soil moisture causes a change in vegetation characteristics (either by leaf condition, or by surface coverage). This temporal vegetation change associated with vegetation status and fractional vegetation cover could be captured by optical remote sensing measurements in that vegetation has different reflectances in the near-infrared band and the visible band. Vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), were derived to reflect the spectral signature of vegetation status and have been widely applied in all kinds of studies (Rouse et al., 1974; Huete et al., 1997; Miura et al., 2001; Huete et al., 2002, Wang et al., 2007).

* Corresponding author. Tel.: +1 210 827 1905.
E-mail address: mtschnur@gmail.com (M.T. Schnur).

1574-9541/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2010.05.001

Previous studies have found links between NDVI and surface moisture (Farrar et al., 1994), NDVI and water deficit and rainfall (Liu and Kogan, 1996), NDVI and root zone soil moisture (Adegoke and Carleton, 2002), and NDVI and crop water use (Hunsaker et al., 2009). Wang et al. (2007) investigated the potential of using the MODIS-derived NDVI to estimate root zone soil moisture and found a consistent and significant correlation ($r=0.46$–0.55) between deseasonalized NDVI and root zone soil moisture at two semi-arid sites and one humid site. They suggested that NDVI derived from space-borne optical sensors may provide a good proxy for root zone soil moisture mapping at large scale.

Various approaches have been developed to estimate soil moisture: from ground-based gravimetric sampling (e.g., Wilson et al., 2003), time-domain reflectometry (e.g., Topp et al., 1980; Roth et al., 1990), to air/space-borne remote sensing techniques (e.g., Engman and Chauhan, 1995; Dubois et al., 1995; Schmugge et al., 2002; Narayan et al., 2004). Most air/space-borne remote sensing techniques use microwave signals that are sensitive to soil water content, but could only penetrate the top few centimeters (Jackson et al., 1996, Schmugge et al., 2002) and could not retrieve the entire root zone soil moisture. In this study, we take the benefits of root zone soil moisture's impacts on vegetation, and use the optical remote sensing measurements (NDVI/EVI) of vegetation variation to estimate the root zone soil moisture variation.

Specifically, we investigate the feasibility of using MODIS NDVI and EVI and the in-situ measured seasonal soil moisture to estimate root zone soil moisture at increasingly distant sites under the same climatic setting with similar soil and vegetation type. We first investigate the correlation of NDVI/EVI and soil moisture at each site. Based on their correlation, we develop regression models at five depths at two native sites (hereafter, the “native” site refers to the site where the regression model is developed using the NDVI/EVI and soil moisture at that site). Finally, we apply the regression models to distant (neighboring) sites to estimate root zone soil moisture for validating the regression models.

2. Study area and data

2.1. Study area

The study area, a semi-arid region covering the Texas–New Mexico border, is primarily grass/herbaceous cover interspersed by shrubby rangeland (Soil Climate Analysis Network, SCAN, 2009). The specific study sites are the Soil Climate Analysis Network (Schaefer et al., 2007) sites named Reese Center (RC), Levelland (LL), Lehman (LM), Crossroads (CR), and Willow Wells (WW), from east to west (Fig. 1a). The five sites have grass/herbaceous primary cover, with grassland–rangeland secondary cover on the four eastern sites and shrubby rangeland secondary cover on the western site (Table 1). The three central sites (LL, LM, and CR) have Amarillo fine sandy loam soil type, with similar soil texture and available water capacity (Table 1, Fig. 1b). The pixel containing the LL site includes a roadway, airport runway, and cultivated cropland. The percentage of sand generally increases from east to west and the available water capacity generally decreases from east to west. Intersite distance ranges from 32 to 47 km, with a total east–west (RC-WW) distance of about 148 km. The five sites are ideal for investigating the potential to estimate root zone soil moisture using a VI at the same climatic setting and with similar vegetation and soil types.

2.2. NDVI and EVI

The NDVI was proposed by Rouse et al. (1974) based on differences in pigment absorption features in the red (~0.660 μm) and near-infrared (~0.860 μm) regions of the electromagnetic spectrum (Eq. (1)). The EVI is a modified index combining blue, red, and near-infrared bands from the MODIS sensor (Eq. (2)) to minimize atmospheric and canopy background effects on NDVI (Huete et al., 1997; Miura et al., 2001; Huete et al., 2002). EVI is an ‘optimized’ index designed to enhance the vegetation signal with improved sensitivity in high biomass regions, and improve vegetation monitoring through a de-coupling of the canopy background signal and a reduction in atmospheric influences (Huete et al., 2002; Matsushita et al., 2007).

$$NDVI = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{Red}}} \tag{1}$$

$$EVI = G\frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + C_1\rho_{\mathrm{Red}} - C_2\rho_{\mathrm{Blue}} + L} \tag{2}$$

where ρ are atmospherically-corrected or partially atmosphere corrected (Rayleigh and ozone absorption) surface reflectances, L is the canopy background adjustment that addresses non-linear, differential NIR and red radiant transfer through a canopy, and C_1, C_2 are the coefficients of the aerosol resistance term, which use the blue band to correct for aerosol influences in the red band. The coefficients adopted in the MODIS-EVI algorithm are: $L=1$, $C_1=6$, $C_2=7.5$, and G (gain factor) $=2.5$ (Terrestrial Biophysics and Remote Sensing Lab, TBRS, 2008).

While the NDVI is chlorophyll sensitive, the EVI is more responsive to canopy structural variations, including leaf area index (LAI), canopy type, plant physiognomy, and canopy architecture. EVI is nearly always lower than NDVI to provide sensitivity throughout the valid NDVI/EVI range (0 to 1.0) (Terrestrial Biophysics and Remote Sensing Lab, TBRS, 2008). The two VIs complement each other in global vegetation studies and improve upon the detection of vegetation changes and extraction of canopy biophysical parameters (Terrestrial Biophysics and Remote Sensing Lab, TBRS, 2008). Semi-arid environments have lower biomass than humid environments (Sowell, 2001), and NDVI was reported to be more sensitive to vegetation in semi-arid regions (Didan et al., 2004), suggesting that NDVI may be a better estimator of root zone soil moisture in semi-arid environments. Senay et al. (2007) chose NDVI for this reason in their study on irrigated agriculture in an otherwise dry land environment in Afghanistan. BEN-Ze'Ev et al. (2006) reported an NDVI-EVI correlation (R^2) of 0.86 in the absence of smoke in a desert environment. Fensholt et al. (2006) reported an NDVI-EVI correlation (R^2) of 0.96 between Terra MODIS NDVI and EVI in the semi-arid northern part of Senegal, using MOD09 series data with 250 m resolution.

In a short time frame (hours), the VI may decrease due to sudden soil moisture increase (rainfall), since increasing top-layer soil moisture would result in a larger decrease of near-infrared reflectance compared to the red reflectance of vegetation (Jensen, 2007). However, in a longer time frame (such as 16 days in this study), it is expected that VI increases as soil moisture increases over the growing season (Wang et al., 2003).

3. Methodology

3.1. Data quality control

Global MODIS NDVI and EVI are designed to provide consistent spatial and temporal comparisons of vegetation conditions. The MOD13Q1 series data are 16-day composites at 250-m spatial resolution and are provided as a gridded level-3 product and include both NDVI and EVI products. The VI algorithms are described in detail at the Terrestrial Biophysics and Remote Sensing Lab (Terrestrial Biophysics and Remote Sensing Lab, TBRS, 2008). MODIS tile h09v05 from March 2005 (the beginning date of SCAN data at the study sites) through December 2008, downloaded from the Oak Ridge National Laboratory Distributed Active Archive Center (Oak Ridge National Laboratory Distributed Active Archive Center, ORNL DAAC, 2009), are used in this study.

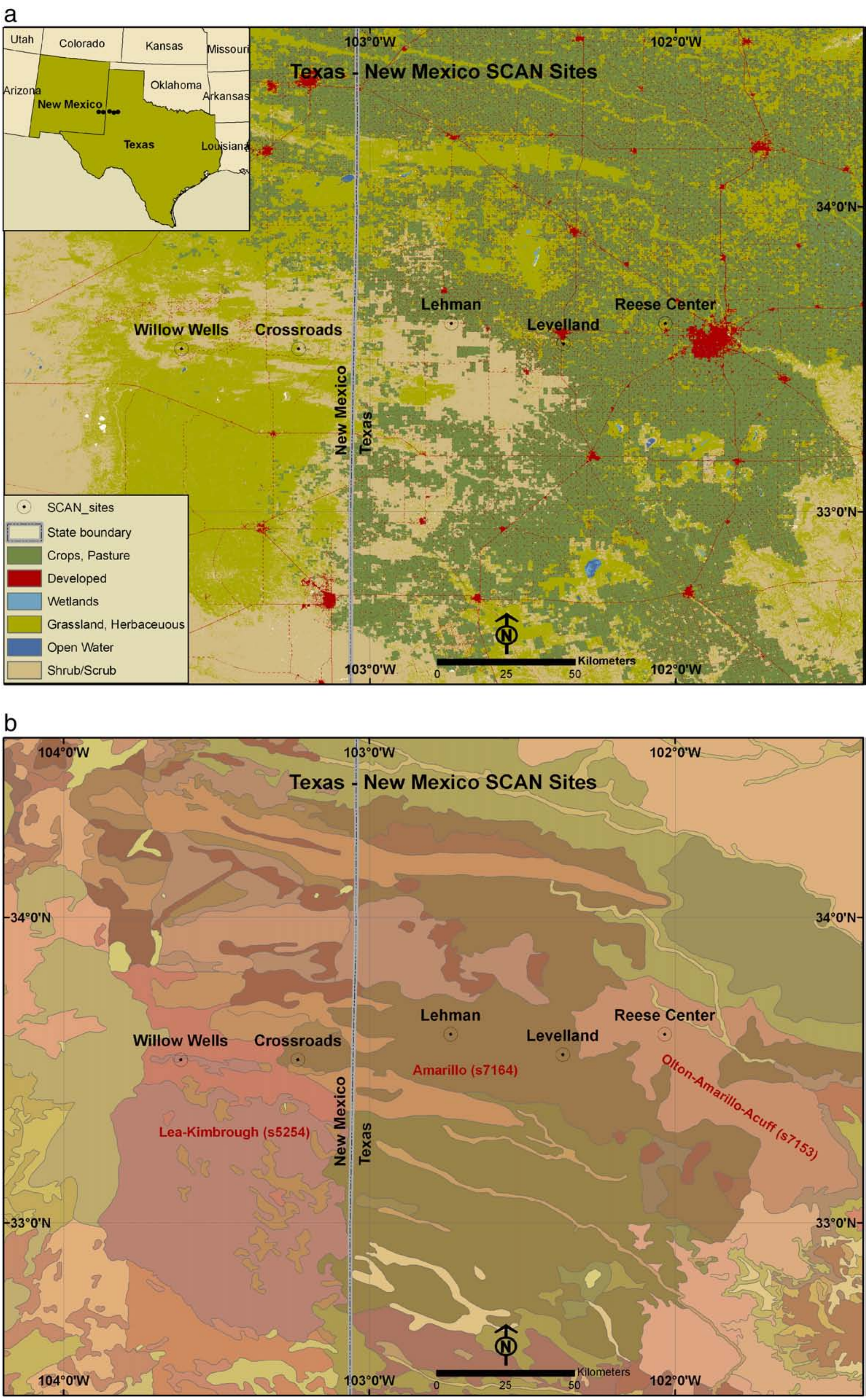
a
Texas - New Mexico SCAN Sites
Utah
Colorado
Kansas
Missouri
Arizona
New Mexico
Oklahoma
Arkansas
Texas
Louisiana
103°0'W
102°0'W
34°0'N
33°0'N
Lehman
Levelland
Reese Center
Willow Wells
Crossroads
New Mexico
Texas
SCAN_sites
State boundary
Crops, Pasture
Developed
Wetlands
Grassland, Herbaceuous
Open Water
Shrub/Scrub
N
Kilometers
0
25
50
b
Texas - New Mexico SCAN Sites
104°0'W
103°0'W
102°0'W
34°0'N
33°0'N
Lehman
Levelland
Reese Center
Willow Wells
Crossroads
Amarillo (s7164)
Olton-Amarillo-Acuff (s7153)
Lea-Kimbrough (s5254)
New Mexico
Texas
N
Kilometers
0
25
50

Table 1
Geoenvironmental characteristics of SCAN sites. Soil texture and available water capacity are the mean of reported values. (Natural Resources Conservation Service (2008), soil texture and available water capacity: SSURGO (2008)).

Study sites	Location	Vegetation	Soil texture	Soil texture (%)			Available water capacity (cm^3 water/cm^3 soil)	General description
				Clay	Silt	Sand		
Reese Center (RC)	33.62°N 102.03°W	Grass/herbaceous cover, grassland rangeland	Clay loam	23.7	22.1	54.2	0.08–0.19	pH 7.9 slope 0.5 well drained
Levelland (LL)	33.55°N 102.37°W	Grass/herbaceous cover, grassland rangeland	Fine sandy loam	20.9	18.9	60.3	0.10–0.17	pH 8.0 slope 0.5 well drained
Lehman (LM)	33.62°N 102.73°W	Grass/herbaceous cover, grassland rangeland	Fine sandy loam	20.8	20.6	58.6	0.10–0.18	pH 8.0 slope 0.5 well drained
Crossroads (CR)	33.53°N 103.23°W	Grass/herbaceous cover, grassland rangeland	Fine sandy loam	18.8	12.2	69.0	0.10–0.18	pH 8.1 slope 0.5 well drained
Willow Wells (WW)	33.53°N 103.62°W	Grass/herbaceous cover, shrubby rangeland	Loamy fine sand	13.0	8.8	78.2	0.05–0.17	pH 7.9 slope 1.5 well drained

Note: pH and soil texture % are the average of reported soil samples.

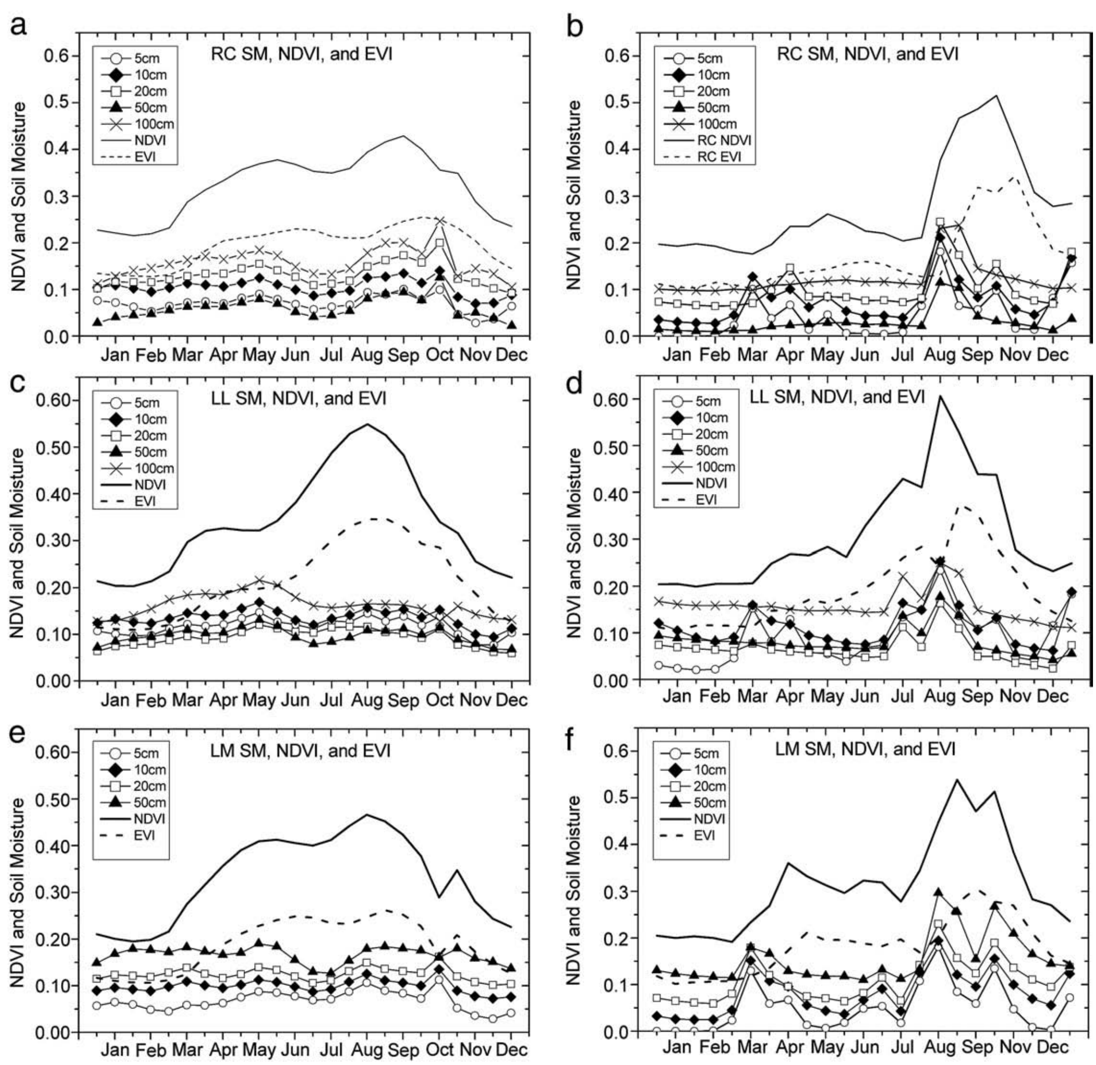

Fig. 2. Seasonal (left column) and raw (right column) time series of soil moisture, NDVI, and EVI for 2005–2008 and for 2006 at the 5 sites, respectively. Soil moisture, NDVI, and EVI have no unit. The *X*-axis is labeled as month, but is actually the 16-day mean. The LM 100 cm depth was not analyzed due to excessive missing values.

Soil moisture values used in this study are point-scale data from the Soil Climate Analysis Network (SCAN, accessed December 2008) sites. Volumetric soil water content (referred to as soil moisture hereafter) was measured hourly at 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm depths using a Stevens Hydra Probe II (Stevens Water Monitoring Systems Inc., 2008) at the SCAN sites. The hourly soil moisture was processed to match the 16-day MODIS data, i.e. the 16-day average is the average of the current day and the following 15 days. This produces eighty seven 16-day observations at each site and depth.

Fig. 1. a. Land cover in the SCAN sites region. The inset map shows the study area location in the southwestern United States. Land cover data: USGS National Land Cover database 2001. b. Soil types in the SCAN sites region. Soil type regions in the map are STATSGO for clarity, but analysis was based on SSURGO soil type for precision.

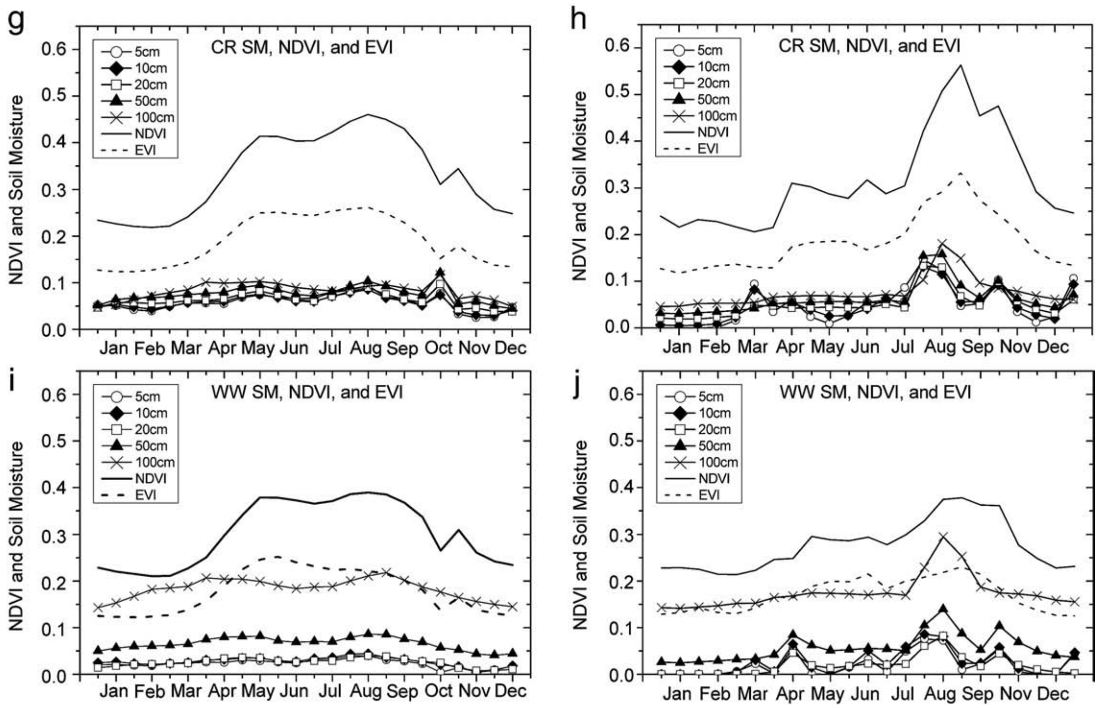

Fig. 2 (*continued*).

All of the soil moisture data sets have missing values. The LM 100 cm depth has 53.3% missing values, from 30 August 2006 through the end of the time series. Mertler and Vannatta (2001) recommend deleting any quantitative variable with more than 15% missing values, and thus the 100 cm soil depth variable at the LM site are removed from the analysis. MODIS NDVI and EVI QA data for the 16-day period beginning 17 January 2007 indicated clouds, and the VI values are interpolated for this time period.

3.2. Seasonality and time series

The seasonal component of time series variables can be identified using a simple moving average: a 47-point moving average for the daily soil moisture and a 3-point moving average for the 16-day NDVI and EVI (Wang et al., 2007). The deseasonalized time series is then produced by subtracting the seasonal time series from the raw time series.

3.3. Statistical analysis and model validation

Pearson Product Moment correlation coefficients (r) are calculated between NDVI/EVI and soil moisture at a native site, for five depths, during the growing season (April to September) for both raw and deseasonalized time series (excepting LM at 100 cm, which was removed due to missing values). Wang et al. (2007) report that r between NDVI and soil moisture peaks when NDVI has 0–5 day time lag over measured soil moisture in a semi-arid climate. This study uses 0 to 30 day time lags between VIs and soil moisture time series to identify the best possible VIs' time lag associated with maximum correlation between VI and soil moisture. For example, in a 5-day time lag, the NDVI of the 8th day (which is actually a composite from the 8th to the 23rd day of the month) is paired with soil moisture of the 3rd day, which is constructed from averaging daily soil moisture from the 3rd to the 18th day.

Regression analysis is a statistical method for studying the relationship between a single dependent variable and one or more independent variables and for estimating the dependent variable using independent variable(s) (Allison, 1999). Soil moisture at the same depth at different sites across the study area are highly correlated, therefore we use a simple linear regression model (Meyers et al., 2006; Allison, 1999). Our hypothesis is that the soil moisture-NDVI regression model developed at one site can be used to estimate root zone soil moisture using the NDVI at a distant site, provided that the two sites have similar soil type, vegetation, and climate regime.

We tested this hypothesis by developing a regression model using deseasonalized NDVI with a 5-day time lag as the independent variable and deseasonalized soil moisture as the dependent variable at two native sites (LL and LM) within the growing season from April to September. Regression models are developed for each depth at the two native sites and then applied to distant sites by substituting the deseasonalized NDVI from the distant site, to estimate the deseasonalized soil moisture of each depth at the distant site. The estimated value at each depth is then added into the seasonal soil moisture of the same depth at the distant site to obtain the estimated soil moisture of the same depth at the distant site. The estimated soil moisture values are then validated with the observed soil moisture of the same depth at the same distant site.

The model validation considers three factors within a similar semi-arid climate setting: distance, vegetation type, and soil type that may affect application of the regression models. First, we consider the impact of distance on the regression model. The LL model is developed using the deseasonalized soil moisture and NDVI at the LL site, and then applied to the LM and CR sites, which have similar soil type (Fig. 1b and Table 1) and vegetation type (Fig. 1a and Table 1), but different distances (35 km to LM and 80 km to CR). Second, the ideal case, the LM model is developed using the deseasonalized soil moisture and NDVI at LM, and then applied to LL and CR, which have similar distance (35 and 47 km) to LM, and similar soil type (Fig. 1b and Table 1) and vegetation type (Fig. 1a and Table 1). Third, the LM model is applied at both RC and WW, which have similar (65 and 82 km) distance to LM, but have different soil type than LM. WW has a different vegetation type than LM, but RC has similar vegetation type to LM.

4. Results

4.1. Soil moisture and VI variation

The seasonal and raw time series of soil moisture, NDVI and EVI for 2005–2008 and for 2006 are shown in Fig. 2. The VI values have the largest seasonal changes and are usually bimodal, with a seasonal increase in April–May (spring) and the second (usually the largest) increase in August–October (fall). The NDVI values are usually larger than the EVI values by 0.1 or more. NDVI and EVI are highly correlated with r ranging from 0.87 to 0.91 ($p<0.001$) for the five sites (not shown). For the seasonal time series (left panel of Fig. 2), soil moisture values at each depth are very similar with small variation monthly or seasonally, while the soil moisture values at different depths have large variation, except for the CR site where soil moistures are around 0.05 with minimal variation. For the raw time series (right panel of

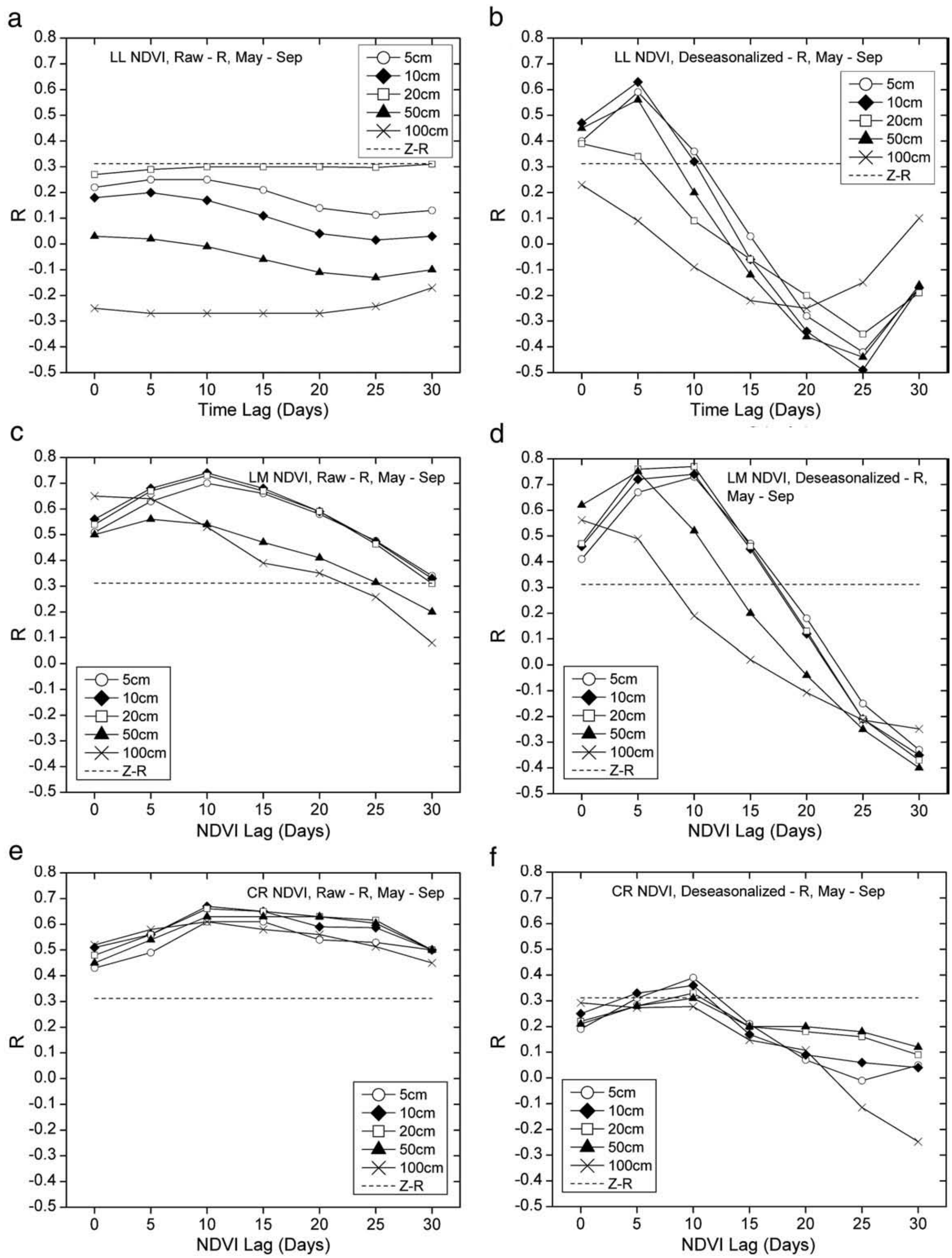

Fig. 3. VI-soil moisture correlation coefficient vs. time lag of NDVI for raw (left column) and deseasonalized (right column), at the LL, LM, and CR sites for 5 depths, during the growing season (April–September). The Z–R line indicates the threshold of statistical significance. The LM 100 cm depth was not analyzed due to excessive missing values. MAD is mean absolute difference between estimated soil moisture and observed soil moisture.

Fig. 2), the monthly variation of soil moisture is much larger than the seasonal time series especially in the rainy season from August to October.

The vertical variation of seasonal soil moisture at RC and LL sites (Fig. 2a, c) have a similar pattern, which first increases with soil depth (to 20 cm for RC and 10 cm for LL), then decreases with soil depth (with minimum value at depth of 50 cm for RC and 20 or 50 cm for LL) and finally increases again with soil depth. At LM (Fig. 2e), soil moisture increases from 0.04 at the top layer to 0.20 at 50 cm depth. CR is the driest site among the five sites, and its soil moisture (Fig. 2g) has a vertical variation pattern similar to LM, but with much less variation in magnitude than LM. The soil moisture at the WW site (Fig. 2i) has similar low values from 5 cm to 20 cm, and then increases with soil depth from 20 cm through 100 cm. Overall, the seasonal patterns are similar at the eastern three sites (RC, LL, and LM), while soil moisture is lower at the western sites (CR and WW).

The raw time series of soil moisture values in 2006 are lower than the seasonal average, meaning that 2006 was drier than the three year average. The 5 cm and 10 cm soil moistures tend to show greater variation throughout the year than the 20 cm, 50 cm, and 100 cm soil moistures. At the RC site (Fig. 2b), 5 cm and 10 cm soil moisture varies more than the seasonal pattern, while the 20 cm, 50 cm, and 100 cm soil moisture varies similar to the seasonal pattern. At the LL site (Fig. 2d), 5 cm and 10 cm soil moisture varies more than 20 cm, 50 cm, and 100 cm in the spring peak, while all 5 depths increase in August. At the LM site (Fig. 2e), the 5 cm to 50 cm depths show an increase in soil moisture in spring, and peaks in August and again in October, so this site has a trimodal pattern in 2006. Soil moisture data at the 100 cm depth at the LM site is missing after August 2006, and thus not shown in the figure. Soil moisture at the CR site (Fig. 2g) in 2006 shows increases in the spring at 5 cm to 20 cm, and at all depths in the fall, with increases in August and October at 5 cm to 50 cm depths. At the WW site (Fig. 2i), there is a slight spring increase and other increases in August and October.

Correlation analysis of soil moisture values at the same depth at the two sites closest together (RC and LL, 32 km) and the two sites farthest apart (RC and WW, 148 km) reveal that same depth soil moistures are strongly correlated. For instance, the soil moisture r values at the same depth were 0.52 to 0.89 between RC and LL, and 0.58 to 0.68 between RC and WW (not shown). Correlation analysis of soil moisture values at the same depth at the LL to CR sites (80 km apart) with similar soil type show that the r values are 0.62 to 0.84 (not shown). This result is encouraging and supports the hypothesis that VIs should reliably estimate soil moisture at distant sites up to ~100 km examined with similar soil type, vegetation, and climate regime.

4.2. Soil moisture-VI correlation at native sites

The growing season for the study area is from April to September, when VIs have the highest correlation with soil moisture. Therefore, forty one 16-day observations for the period of May–September in each year from 2005 to 2008 are used for the analysis. Fig. 3 shows the correlation examples between (raw and deseasonalized) NDVI and soil moisture at LL, LM and CR sites, which have the same soil type, Amarillo (s7164) (Fig. 1b). Generally, the deseasonalized soil moisture (except at 100 cm) and NDVI demonstrate consistent and significant correlations with maximum value at a 5 (LL and LM) to 10-day (CR) time lag for NDVI. The raw soil moisture and NDVI do not show significant correlation at the LL site, have similar correlation with the deseasonalized time series at the LM site, and have higher correlation values than the deseasonalized time series at the CR site. This indicates that only the deseasonalized NDVI could reflect the response of vegetation to the soil moisture variation and thus could be used as a proxy to estimate the root zone soil moisture variation (Wang et al. 2007).

Table 2
Maximum correlation value (r) and time lag (days) for correlation between deseasonalized NDVI and EVI with soil moisture. The threshold of statistical significance for r is 0.31. The LM 100 cm depth was not analyzed due to excessive missing values.

		RC		LL		LM		CR		WW	
		r	Time lag	r	Time lag	r	Time lag	r	Time lag	r	Time lag
NDVI	5 cm	0.58	10	0.59	5	0.73	10	0.39	10	0.19	20
	10 cm	0.56	10	0.63	5	0.74	10	0.36	10	0.16	0
	20 cm	0.55	10	0.39	0	0.77	10	0.33	10	0.22	20
	50 cm	0.50	10	0.56	5	0.75	5	0.31	5	0.19	20
	100 cm	0.45	5	0.23	0			0.29	0	0.26	20
EVI	5 cm	0.33	20	0.61	20	0.58	25	0.43	5	0.38	10
	10 cm	0.35	20	0.61	20	0.65	25	0.52	5	0.35	10
	20 cm	0.36	20	0.43	20	0.69	25	0.57	5	0.45	15
	50 cm	0.41	20	0.49	20	0.43	20	0.56	0	0.44	15
	100 cm	0.33	20	0.33	10			0.37	0	0.41	15

Table 2 shows maximum correlation value (r) and time lag in days for correlation between deseasonalized NDVI and EVI with soil moisture. Since semi-arid regions typically do not have high biomass, NDVI should not saturate and the two VIs should perform equally well, as the strong correlation between the two VIs (r = 0.87–0.91) suggest. However, when the VIs are correlated with soil moisture, EVI has a maximum correlation with soil moisture with a longer time lag than NDVI at RC, LL, and LM, and with a somewhat shorter time lag at CR and WW. EVI has a lower maximum correlation than NDVI at RC and LM, a similar maximum correlation at LL, and a higher maximum correlation at CR and WW. Since the study area is not a high biomass region, EVI is not necessary, and appears to give inconsistent results when correlated with soil moisture compared to NDVI (Table 2). Therefore, deseasonalized NDVI is used for the regression model to estimate root zone soil moisture at distant sites.

4.3. Soil moisture estimation and validation

Table 3 summarizes regression model validation statistics for the LL and LM models applied to distant sites. Scatter plots for estimated vs. observed soil moisture at 10 cm soil depth (root zone, the best results) for the two regression models are shown in Figs. 4 and 5. When distance between native site and distant site is a factor, but soil and vegetation are similar (LL model applied to LM and CR, Fig. 4a, b), the r^2 values between estimated and in-situ measured soil moisture at different depths vary from 0.72 to 0.86 at LM, and vary from 0.55 to 0.67 at CR (Table 3) with longer distances and smaller correlation. The

Table 3
Coefficient of determination (r^2) and MADs between estimated and observed soil moisture for the regression models developed at the LL and LM sites and applied to distant sites, at 5, 10, 20, and 50 cm. Regression estimates were derived from deseasonalized time series of NDVI and soil moisture, with a 5-day time lag of NDVI. MAD is mean absolute difference between estimated soil moisture and observed soil moisture.

Native sites	Distant sites			Soil depth (cm)			
	Name	Distance (km)	Statistic	5	10	20	50
LL	LM	35	r^2	0.76	0.79	0.72	0.86
			MAD	0.02	0.02	0.02	0.02
	CR	80	r^2	0.55	0.58	0.67	0.66
			MAD	0.02	0.02	0.01	0.02
LM	LL	35	r^2	0.66	0.74	0.66	0.71
			MAD	0.03	0.02	0.03	0.02
	CR	47	r^2	0.50	0.53	0.52	0.55
			MAD	0.02	0.02	0.02	0.02
	RC	65	r^2	0.72	0.85	0.86	0.77
			MAD	0.02	0.02	0.02	0.01
	WW	82	r^2	0.62	0.76	0.85	0.88
			MAD	0.02	0.02	0.01	0.02

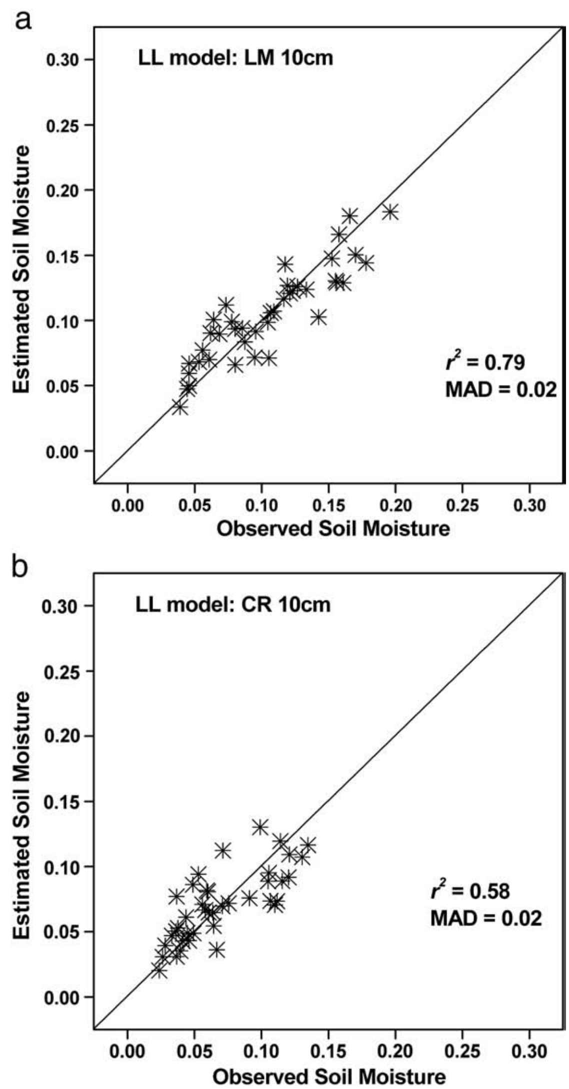

Fig. 4. Estimated vs. observed soil moisture for the regression model developed at the LL site and applied to the LM (a) and CR (b) sites, at 10 cm. Regression estimates were derived from deseasonalized time series of NDVI and soil moisture, with a 5-day time lag of NDVI. MAD is mean absolute difference between estimated soil moisture and observed soil moisture.

mean absolute difference (MAD) between estimated soil moisture and in-situ soil moisture is 0.02 for all depths at LM and varies from 0.01 to 0.02 at CR (Table 3).

When distance, soil type, and vegetation are similar between native site and distant site (LM model applied to LL and CR, Fig. 5a, b), the r^2 values between estimated and in-situ measured soil moisture at different depths vary from 0.66 to 0.74 at LL, and vary from 0.50 to 0.55 at CR (Table 3), which shows slightly different land cover than LL and LM (Fig. 1a). The estimated soil moisture using the LL model has higher r^2 values than the LM model at CR (Table 3). The mean absolute difference (MAD) between estimated soil moisture and in-situ soil moisture at different depths varies from 0.02 to 0.03 at LL and 0.02 for all depths at CR (Table 3).

When distance and vegetation are similar but soil type is different (LM model applied to RC, Fig. 5c), or when distance is similar but soil type and vegetation are different (LM model applied to WW, Fig. 5d) the r^2 values between estimated and in-site measured soil moisture at different depths vary from 0.72 to 0.86 at RC and vary from 0.62 to 0.88 at WW (Table 3). The mean absolute difference (MAD) between estimated soil moisture and in-situ soil moisture at different depths varies from 0.01 to 0.02 at both RC and WW (Table 3).

5. Discussion

5.1. Raw and deseasonalized time series

Both raw and deseasonalized time series have been used in previous studies and both can give good results. Wang et al. (2007) used raw and deseasonalized data and obtained results consistent with Kendall and Ord (1990), reporting that at the semi-arid Adams Ranch and Walnut Gulch SCAN sites, the Pearson's *r* between the raw time series of soil moisture and NDVI has larger values than the deseasonalized time series, while at the humid Prairie View SCAN site in east Texas, the Pearson's *r* between the raw time series of soil moisture and NDVI had lower values than the deseasonalized time series. We also find that deseasonalized NDVI has consistent and significant correlation with deseasonalized soil moisture, while the raw NDVI and soil moisture may have lower *r* values (LL) or larger *r* values (CR and LM) due to their different or similar seasonality. Since NDVI has a similar seasonal pattern with soil moisture at CR and LM, removing the seasonal component lowers the *r* value between the VI and soil moisture (Wang et al., 2007). In contrast, at LL, removing the different seasonal patterns of soil moisture and NDVI improve their *r* values, which represents the real response of vegetation (VIs) to the soil moisture change and reveals the physical mechanism to use our regression model to estimate the root zone soil moisture variation using NDVI. The LL 20 cm time series regression model had a much lower *r* value (0.34) that is not explained by available data, but still estimated soil moisture at LM and CR. The LL site behaved differently from the other sites, as previously noted. When NDVI was weakly correlated or not correlated with soil moisture, as occurs at the CR and WW sites, respectively, the regression model estimated soil moisture variation using deseasonalized NDVI has larger uncertainties and may not represent the realistic soil moisture variation. In this case, a conservative alternative is to use the seasonal soil moisture to fill the final estimated soil moisture.

At the CR site, with relatively low soil moisture, deseasonalized NDVI has a weak correlation with 5, 10, 20, and 50 cm soil moisture with 10 days time lag. There is no significant correlation between deseasonalized NDVI and soil moisture at the WW site, suggesting that at sites with low soil moisture, the deseasonalized VIs are not effective at estimating root zone soil moisture. This study shows greater variation of correlations in deseasonalized time series at five sites (Table 2, Fig. 3).

Deseasonalized NDVI was correlated with soil moisture at up to 100 cm depth at RC and LM, and up to 50 cm at LL in this study, suggesting that the root zone extends at least to 50 cm at these sites. Root density profiles developed by Kurc and Small (2007) indicate that "roots are present to soil depths of 1 m in both the grassland and shrubland" in central New Mexico, near this study site, "with a majority of grass roots near the surface and a majority of shrub roots between 20 and 40 cm". This suggests that deseasonalized NDVI is estimating root zone soil moisture used by grasses near the surface and by shrubs at depths of 20 to 100 cm.

5.2. NDVI vs. EVI

Soil moisture is one of the major controlling factors for vegetation growth in semi-arid climates, and NDVI and EVI are closely tied to soil moisture, therefore NDVI and EVI will change closely with soil moisture. This study used Terra MODIS NDVI and EVI, enabling comparison of the two indexes in a semi-arid climate. Previous studies have used NDVI and EVI in semi-arid climates (Laneve and Castronuovo, 2005; Fensholt and Sandholt, 2005; Kawamura et al., 2005; Huete et al., 2002), but none have compared the two indices for sensitivity to root zone soil moisture. The raw NDVI and EVI time series are highly correlated at the 5 sites ($r = 0.87$ to 0.91), suggesting that the VI-soil moisture correlations would be very similar. However, this was not the case, as EVI has inconsistent maximum correlations with soil moisture (Table 2). EVI was developed to improve sensitivity in high biomass regions, where NDVI tended to saturate (Terrestrial

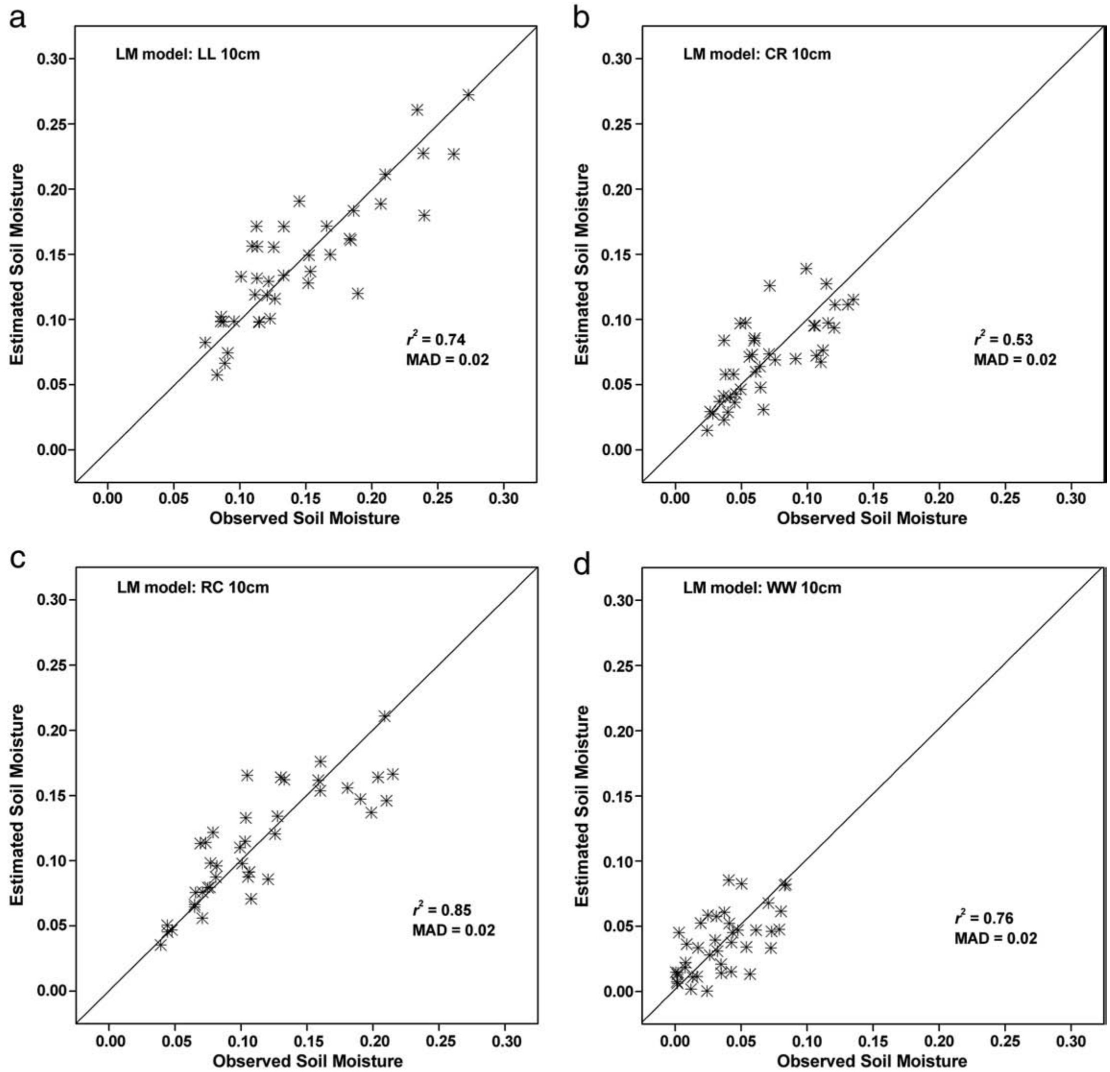

Fig. 5. Estimated vs. observed soil moisture for the regression models developed at the LM site and applied to the LL (a), CR (b), RC (c), and WW (d) sites, at 10 cm. Regression estimates were derived from deseasonalized time series of NDVI and soil moisture, with a 5-day time lag of NDVI. MAD is mean absolute difference between estimated soil moisture and observed soil moisture.

Biophysics and Remote Sensing Lab, TBRS, 2008). Since semi-arid regions typically do not have high biomass, NDVI is sufficient for the semi-arid study area. Therefore, in our regression model, only the NDVI-soil moisture regression model is examined.

5.3. Estimating soil moisture at distant sites

The two regression models produce consistent results when applied to distant sites and validated against observed soil moisture (LM: $r^2=0.55$–0.86; LL: $r^2=0.50$–0.88). This suggests that deseasonalized NDVI could serve as a proxy to estimate the root zone soil moisture variation in the examined areas. Distance appears at first to decrease the estimation capability as the LL model applied to CR has a lower coefficient of determination ($r^2=0.55$–0.66) than LL applied to LM ($r^2=0.72$–0.86). However, the LM model applied to CR ($r^2=0.50$–0.55) was also lower than LM applied to LL ($r^2=0.66$–0.74) when distance is similar, suggesting that distance is not a determining factor in this study area. Both models applied to CR produce lower r^2 values than when applied to the other distant sites, suggesting that there are some other factors at the CR site not observed from the soil moisture and vegetation maps (Fig. 1a, b) affecting the soil moisture estimate.

When the LM model is applied to RC with different soils, the soil moisture estimate compared to observed soil moisture was higher ($r^2=0.72$–0.86) than LM applied to WW with different soil and vegetation ($r^2=0.62$–0.88), suggesting that vegetation type is more important than soil type in estimating soil moisture.

The estimates compared to observed soil moisture are consistently higher at the 10 cm–50 cm depth than at 5 cm at all sites (Table 3), suggesting that the root zone is concentrated in the 10 cm–50 cm depth, supporting the root zone range reported by Kurc and Small (2007).

Overall in this study, a regression model developed at a native site and applied to a distant site with similar soil type and vegetation accounts for 50–86% of the variation in observed soil moisture. When a regression model developed at a native site and applied to distant sites with either different soil type and similar vegetation (e.g., LM to RC), or different soil type and different vegetation (e.g., LM to WW), but still in the semi-arid climate, the model accounts for 62–88% of variation in the observed soil moisture, at distances up to about 100 km.

The vegetation type reported by SCAN is similar at the three center sites (CR, LM and LL), but a detailed analysis of the vegetation type at each site is needed to determine the specific NDVI-soil moisture relationship. A multivariate regression model using soil type, soil texture, vegetation type, along with other climatic variables such as precipitation, soil temperature, etc. may improve the model and better explain the variation among the three sites.

Our results suggest that regression models derived from deseasonalized time series can be applied to distant sites in a semi-arid region to estimate root zone soil moisture within the growing season, given known soil type, vegetation, and climate regime. This increases the potential for regression models to produce root zone soil moisture maps, as suggested by Wang et al. (2007). This method assumes that soil types and vegetation types are known. If soil and vegetation types are not available, one needs to obtain them before any empirical equations can be used for estimating soil moisture.

6. Conclusions

This study examined the characteristics of soil moisture at five SCAN sites on the Texas–New Mexico border using time series data from March 2005 to June 2008.

Results of regression models developed at native sites find that MODIS deseasonalized NDVI can be used to estimate root zone soil moisture at distant sites up to 100 km examined with similar soil and vegetation characteristics at the 10-50 cm depth. More research is needed into the VI-soil moisture relationship at distant sites in arid, semi-arid, and humid regions using the 8-day VI products, detailed soil and vegetation characteristics, and a longer time series to improve the estimation.

Acknowledgements

Data provided by the US Soil Climate Analysis Network and the Oak Ridge National Laboratory Distributed Active Archive Center are sincerely acknowledged. All reviewers contributed helpful and high quality comments that greatly improve this manuscript.

References

Adegoke, J.O., Carleton, A.M., 2002. Relations between soil moisture and satellite vegetation indices in the US Corn Belt. Journal of Hydrometeorology 3 (4), 395–405.

Allison, P.D., 1999. Multiple Regression: a Primer. Sage Publications, Thousand Oaks, CA.

Ben-Ze'Ev, E., Karnieli, A., Agam, N., Kaufman, Y., Holben, B., 2006. Assessing vegetation condition in the presence of biomass burning smoke by applying the Aerosol-free Vegetation Index (AFRI) on MODIS images. International Journal of Remote Sensing 27 (15), 3203–3221.

Buxbaum, C.A.Z., Vanderbilt, K., 2007. Soil heterogeneity and the distribution of desert and steppe plant species across a desert-grassland ecotone. Journal of Arid Environments 69 (4), 617–632.

Didan, K., Huete, A., Jacobson, A., Yuan, Y., 2004. The MODIS Vegetation Index (MOD13) product series: specifications, performance and status. Terrestrial Biophysics & Remote Sensing Lab, Department of Soil, Water, and Environmental Science, The University of Arizona, Tucson, AZ. Accessed May 2008 from http://www.ntsg.umt.edu/MODISVeg.2004/posters/K_Didan_VI_Product_Montana_Worskho.pdf.

Dubois, P.C., Vanzyl, J., Engman, T., 1995. Measuring soil-moisture with imaging radars. IEEE Transactions on Geoscience and Remote Sensing 33 (4), 915–926.

England, C.B., 1975. Root depth as a sensitive parameter in a deterministic hydrologic model. Journal of the American Water Resources Association 11 (5), 1046–1051.

Engman, E.T., Chauhan, N., 1995. Status of microwave soil-moisture measurements with remote-sensing. Remote Sensing of Environment 51 (1), 189–198.

Farrar, T.J., Nicholson, S.E., Lare, A.R., 1994. The influence of soil type on the relationships between NDVI, rainfall, and soil-moisture in semiarid Botswana. 2. NDVI response to soil-moisture. Remote Sensing of Environment 50 (2), 121–133.

Fensholt, R., Sandholt, I., 2005. Evaluation of MODIS and NOAA AVHRR vegetation indices with in situ measurements in a semi-arid environment. International Journal of Remote Sensing 26 (12), 2561–2594.

Fensholt, R., Sandholt, I., Stisen, S., 2006. Evaluating MODIS, MERIS, and vegetation — vegetation indices using in situ measurements in a semiarid environment. IEEE Transactions on Geoscience and Remote Sensing 44 (7), 1774–1786.

Guswa, A.J., 2008. The influence of climate on root depth: a carbon cost–benefit analysis. Water Resources Research 44 (2), 11.

Huete, A.R., Liu, H.Q., Batchily, K., vanLeeuwen, W., 1997. A comparison of vegetation indices global set of TM images for EOS-MODIS. Remote Sensing of Environment 59 (3), 440–451.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83 (1–2), 195–213.

Hunsaker, D.J., El-Shikha, D.M., Clarke, T.R., French, A.N., Thorp, K.R., 2009. Using ESAP software for predicting the spatial distributions of NDVI and transpiration of cotton. Agricultural Water Management 96 (9), 1293–1304.

Jackson, T.J., Schmugge, J., Engman, E.T., 1996. Remote sensing applications to hydrology: soil moisture. Hydrological Sciences Journal 41 (4), 517–530.

Jensen, J., 2007. Remote Sensing of the Environment: an Earth Resource Perspective, second ed. Pearson Prentice Hall, New Jersey.

Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, S., 2005. Comparing MODIS vegetation indices with AVHRR NDVI for monitoring the forage quantity and quality in Inner Mongolia grassland, China. Grassland Science 51, 33–40.

Kendall, S.M., Ord, J.K., 1990. Time Series, third ed. Oxford University Press, New York.

Kizito, F., Dragila, M., Sene, M., Lufafa, A., Diedhiou, I., Dick, R.P., Selker, J.S., Dossa, E., Khouma, M., Badiane, A., Ndiaye, S., 2006. Seasonal soil water variation and root patterns between two semi-arid shrubs co-existing with Pearl millet in Senegal, West Africa. Journal of Arid Environments 67 (3), 436–455.

Kurc, S.A., Small, E.E., 2007. Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland. Water Resources Research 43 (6), 13.

Laneve, G., Castronuovo, M.M., 2005. Comparison between vegetation change analysis in Kenya based on AVHRR and SeaWiFS images. International Journal of Remote Sensing 26 (12), 2549–2559.

Liu, W.T., Kogan, F.N., 1996. Monitoring regional drought using the Vegetation Condition Index. International Journal of Remote Sensing 17 (14), 2761–2782.

Magagi, R.D., Kerr, Y.H., 2001. Estimating surface soil moisture and soil roughness over semiarid areas from the use of the copolarization ratio. Remote Sensing of Environment 75 (3), 432–445.

Matsushita, B., Yang, W., Chen, J., Onda, Y., Qiu, G.Y., 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: a case study in high-density cypress forest. Sensors 7 (11), 2636–2651.

Mertler, C.A., Vannatta, R.A., 2001. Advanced and Multivariate Statistical Methods: Practical Application and Interpretation. Pearson, Los Angeles.

Meyers, L., Gamst, G., Guarino, A., 2006. Applied Multivariate Research: Design and Interpretation. Sage Publications, Inc, Thousand Oaks, CA.

Miina, J., Pukkala, T., 2002. Application of ecological field theory in distance-dependent growth modelling. Forest Ecology and Management 161 (1–3), 101–107.

Miura, T., Huete, A.R., Yoshioka, H., Holben, B.N., 2001. An error and sensitivity analysis of atmospheric resistant vegetation indices derived from dark target-based atmospheric correction. Remote Sensing of Environment 78 (3), 284–298.

Narayan, U., Lakshmi, V., Njoku, E.G., 2004. Retrieval of soil moisture from passive and active L/S band sensor (PALS) observations during the Soil Moisture Experiment in 2002 (SMEX02). Remote Sensing of Environment 92 (4), 483–496.

Natural Resources Conservation Service (NRCS) Soil Climate Analysis Network (SCAN). Accessed May 2008 from ftp://ftp.wccr.nrcs.usda.gov/data/scan/.

Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC), 2009. MODIS subsetted land products, Collection 5. . Available on-line [http://www.daac.ornl.gov/MODIS/modis.html] from ORNL DAAC, Oak Ridge, Tennessee, U.S.A. Accessed January 2009.

Pavon, N.P., Briones, O., 2001. Phenological patterns of nine perennial plants in an intertropical semi-arid Mexican scrub. Journal of Arid Environments 49 (2), 265–277.

Roth, K., Schulin, R., Fluhler, H., Attinger, W., 1990. Calibration of time domain reflectometry for water-content measurement using a composite dielectric approach. Water Resources Research 26 (10), 2267–2273.

Rouse Jr, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS. Presented at Third Earth Resources Technology Satellite-1 Symposium, NASA, Washington, D.C, pp. 309–317.

Schaefer, G.L., Cosh, M.H., Jackson, T.J., 2007. The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). Journal of Atmospheric and Oceanic Technology 24 (12), 2073–2077.

Soil Climate Analysis Network (SCAN), http://www.wcc.nrcs.usda.gov/scan/, accessed December 2009.

Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J., Rango, A., 2002. Remote sensing in hydrology. Advances in Water Resources 25 (8–12), 1367–1385.

Senay, G.B., Budde, M., Verdin, J.P., Melesse, A.M., 2007. A coupled remote sensing and simplified surface energy balance approach to estimate actual evapotranspiration from irrigated fields. Sensors 7 (6), 979–1000.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for New Mexico and Texas. Available online at http://soildatamart.nrcs.usda.gov accessed May 2008.

Song, J., Wesely, M.L., Coulter, R.L., Brandes, E.A., 2000. Estimating watershed evapotranspiration with PASS. Part I: Inferring root-zone moisture conditions using satellite data. Journal of Hydrometeorology 1 (5), 447–461.

Sowell, J., 2001. Desert Ecology: an Introduction to Life in the Arid Southwest. The University of Utah Press.

Stevens Water Monitoring Systems, Inc. http://www.stevenswater.com/index.aspx. Accessed October 2008.

Terrestrial Biophysics and Remote Sensing Lab (TBRS), Theoretical Basis for the Enhanced Vegetation Index. Accessed March 2008 from http://tbrs.arizona.edu/cdrom/Index.html.

Topp, G.C., Zebchuk, W.D., Dumanski, J., 1980. The variation of insitu measured soil–water — properties within soil map units. Canadian Journal of Soil Science 60 (3), 497–509.

Vinton, M.A., Burke, I.C., 1997. Contingent effects of plant species on soils along a regional moisture gradient in the Great Plains. Oecologia 110 (3), 393–402.

Walker, J., Sharpe, P.J.H., Penridge, L.K., Wu, H., 1989. Ecological field-theory — the concept and field-tests. Vegetatio 83 (1–2), 81–95.

Wang, J., Rich, P.M., Price, K.P., 2003. Temporal responses of NDVI to precipitation and temperature in the central Great Plains. USA. International Journal of Remote Sensing 24 (11), 2345–2364.

Wang, X.W., Xie, H.J., Guan, H.D., Zhou, X.B., 2007. Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions. Journal of Hydrology 340 (1–2), 12–24.

Wilson, D.J., Western, A.W., Grayson, R.B., Berg, A.A., Lear, M.S., Rodell, M., Famiglietti, J. S., Woods, R.A., McMahon, T.A., 2003. Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment. New Zealand. Journal of Hydrology 276 (1–4), 254–274.

Wu, H.I., Sharpe, P.J.H., Walker, J., Penridge, L.K., 1985. Ecological field-theory — a spatial-analysis of resource interference among plants. Ecological Modelling 29 (1–4), 215–243.